

**A MANUFACTURING COST ESTIMATION BY UTILIZING A NOVEL
SENSOR BASED COST MODEL IN A KNOWLEDGE MANAGEMENT
SYSTEM**

HANDE TOKUCOGLU

**A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores
University for the degree of PhD/ Doctor of Philosophy**

AUGUST-2020

ACKNOWLEDGEMENT

Initially, I would like to thank my mom, dad and my sister who made the PhD process possible for me by giving me their love and strength. They believed in me and kept me motivated during my PhD journey as they have done since my childhood. Next, I would like to thank my husband to whom I will always be grateful for his love, patience and understanding during this journey. I am very lucky to have someone who supports my career and picks me up when I fall down because of struggles in life. Moreover, I would not have completed my project without the help and support of my supervisors. Therefore, I would like to thank to Prof Xun CHEN, Dr Tahsin OPOZ and Prof Abdennour EL RHALIBI. Besides, I would like to thank Drew BURNS and Louise GEEKIE who work in Croft Filters as well as Andy FLATT who works in Hi-Tech Steel Services for providing data for my project and easing my burden. Finally, I would like to dedicate this thesis to my aunt Serap GUREL who passed away suddenly one month before my submission date. She was the angel of our family who guided and protected us in this life. Now, she is sleeping peacefully and we are trying to learn how to fight against difficulties of life by ourselves. We miss you so much. Rest in peace.

Table of Contents

ACKNOWLEDGEMENT	2
Table of Contents	3
List of Tables	6
List of Figures	8
ABSTRACT	1
CHAPTER 1: INTRODUCTION	3
1.1. Aim of the Project:.....	4
1.2. Objectives of the Project:.....	4
CHAPTER 2: LITERATURE REVIEW.....	6
2.1. Production Process Planning, Monitoring and Control:	6
2.1.1. Manufacturing Process Planning:.....	6
2.1.2. Manufacturing Process Monitoring and Control:	7
2.2. Production Costing:.....	10
2.2.1. Some Detailed Explanations or Examples of Cost Estimation Methods:	15
2.3. Knowledge Management Systems:.....	18
2.3.1. Artificial Intelligence:	21
2.3.2. Data Mining:.....	33
2.3.3. Case Based Reasoning:	34
CHAPTER 3: METHODOLOGY OVERVIEW	37
3.1. Generic Cost Model Development:	37
3.2. Sensor Based Cost Modelling and Relevant Data Acquisition:.....	37
3.3. Assessment of The Manufacturing Process Efficiency:	37
3.4. Framework of The Knowledge Support System:	38
CHAPTER 4: GENERIC COST MODEL DEVELOPMENT	39
4.1. Step 1: Generic Cost Model Parameters:.....	39
4.2. Step 2: Sub-Formulas of Generic Cost Model:	42
4.3. Obtaining the Generic Cost Model for Calculation of Unit Manufacturing Cost of a Product:.....	44
4.3.1. Direct Process Level Cost (.....	44
4.3.2. Process Overhead Level Cost (£):	44
4.3.3. Company Level Cost (£):.....	45
CHAPTER 5: SENSOR BASED COST MODELLING AND RELEVANT DATA ACQUISITION	47
5.1. Introduction:	47
5.2. Experiment with Laser Sintering Machine in Company 1	48
5.2.1. Calculation of Electrical Energy Cost of the Laser Sintering Process	48

5.2.2. Analysis of Laser Sintering Process via Active Power Graph:	49
5.3. Experiment with Cutting Machine in Company 2:	51
5.3.1. Calculation of Electrical Energy Cost of the Laser Sintering Process:	51
5.3.2. Analysis of Cutting Process through Active Power Graph:	52
5.4. Experiment with Milling Machine in Liverpool John Moores University:	56
5.4.1. Calculation of Electrical Energy Cost of the Milling Process:.....	56
5.4.2. The Analysis of Milling Process through Active Power Graph:.....	56
CHAPTER 6: ASSESSMENT OF THE MANUFACTURING PROCESS EFFICIENCY.....	59
6.1. Classification of a Machine Status:.....	59
6.1.1. Company 1 Laser Sintering Machine Status Classification:	59
6.1.2. Application of Decision Tree Algorithm to Company 2 for Cutting Machine Status Classification:	71
6.1.3. Application of Multi-Class Support Vector Machine for Milling Machine Status Classification.....	77
6.1.4. Application of Decision Tree Algorithm for Classification of Status of Milling Machine:	85
6.2. Calculation of the Utilization Rate of the Machine:	93
6.2.1. Calculation of the Utilization Rate of the Laser Sintering Machine:.....	93
6.2.2. Calculation of the Utilization Rate of the Cutting Machine:.....	102
6.2.3. Calculation of the Utilization Rate of the Milling Machine:.....	105
6.4. Validation of the System of Calculation of the Utilization Rate of the Machine	110
CHAPTER 7: FRAMEWORK of the KNOWLEDGE SUPPORT SYSTEM	111
7.1. Designing Level of the Knowledge Support System:	111
7.2. Creation Level of the Knowledge Support System:.....	113
7.2.1. Explanation of the Flowchart of the Knowledge Support System Database:	114
7.2.2. Detailed Explanation of Knowledge Support System Database:	120
CHAPTER 8: CONCLUSION and FUTURE WORKS.....	127
8.2. Discussion and Conclusion:	127
8.2. Future Works:	131
References	132
APPENDIX 1: BINARY SUPPORT VECTOR MACHINE CODE of COMPANY 1 DATABASE	138
APPENDIX 2: MULTI-CLASS SUPPORT VECTOR MACHINE CODE of COMPANY 1 DATABASE .	140
APPENDIX 3: DECISION TREE ALGORITHM CODE of COMPANY 2 DATABASE	160
APPENDIX 4: MULTI-CLASS SUPPORT VECTOR MACHINE ALGORITHM CODE for MILLING MACHINE STATUS CLASSIFICATION	161
APPENDIX 5: DECISION TREE ALGORITHM CODE for MILLING MACHINE STATUS CLASSIFICATION	163
APPENDIX 6: The MATLAB CODE of the CALCULATION of the UTILIZATION	165
RATE of the MACHINE	165
APPENDIX 7: "FRONT PAGE" FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	167
APPENDIX 8: "REGISTRATION" FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE ...	168
APPENDIX 9: "LOG IN" FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	169
APPENDIX 10: "WELCOME" FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	170

APPENDIX 11: “NEW PRODUCT” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	171
APPENDIX 12: “PRODUCT TYPES” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	172
APPENDIX 13: “PRODUCT 2” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	173
APPENDIX 14: “PRODUCT DRAWING” FORM OF THE KNOWLEDGE SUPPORT SYSTEM DATABASE	174
APPENDIX 15: “PROCESSES OF PRODUCT 2” FORM OF THE KNOWLEDGE SUPPORT SYSTEM DATABASE	175
APPENDIX 16: “ADDITIVE MANUFACTURING PROCESS STEPS” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	176
APPENDIX 17: “FUSED DEPOSITION MODELLING PROCESS STEPS” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	177
APPENDIX 18: “ADDITIVE MANUFACTURING CASES” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	178
APPENDIX 19: “FUSED DEPOSITION MODELLING CASES” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	179
APPENDIX 20: “MANUFACTURING COST CALCULATION” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE	180
APPENDIX 21: “WELCOME” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	181
APPENDIX 22: “OVERHEAD COSTS” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	182
APPENDIX 23: “PROCESS OVERHEAD COST” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	183
APPENDIX 24: “COMPANY OVERHEAD COST” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	184
APPENDIX 25: “TOTAL OVERHEAD COST” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	185
APPENDIX 26: “DIRECT PROCESS COST” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	186
APPENDIX 27: “TOTAL UNIT MANUFACTURING COST” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	187
APPENDIX 28: “WELCOME” FORM of the CASE BASED REASONING SYSTEM SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	188
APPENDIX 29: “CASE BASED REASONING” FORM of the CASE BASED REASONING SYSTEM SUB- DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	189
APPENDIX 30: “PREVIOUS CASES” FORM of the CASE BASED REASONING SYSTEM SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE	190

List of Tables

Table 2.1: Cost Estimation Methods.....	11
Table 2.2.: Machine Learning Methods.....	22
Table 2.3: Data Mining Operations.....	33
Table 4.1.: Cost Parameters of Generic Cost Model.....	39
Table 4.2.: The Cost Parameters of Company 1.....	40
Table 4.3.: The Cost Parameters of Company 2... ..	41
Table 4.4.: Cost Parameters of Generic Cost Model that Gathered Under Five Groups.....	42
Table 5.1: Minimum, Maximum and Average Value of Power Types of Laser Sintering Machine.....	49
Table 5.2.: Minimum, Maximum and Average Value of Power Types of Cutting Machine.....	51
Table 5.3.: Minimum, Maximum and Average Value of Power Types of Milling Machine.....	56
Table 6.1: Power Data of Laser Sintering Machine Experiment of Company 1.....	60
Table 6.2: Prediction Results of Binary Support Vector Machine Code.....	66
Table 6.3: Prediction Results of Multi-Class Support Vector Machine Code.....	70
Table 6.4: Power Values of Cutting Machine of Company 2 and Their Status.....	71
Table 6.5: Prediction Results of Decision Tree Code of Company 2.....	76
Table 6.6: A Part of the Dataset of the Milling Machine.....	78
Table 6.7: Prediction Results of Week 1 of Milling Machine Experiment.....	82
Table 6.8: A Part of the Prediction Results of Week 2 of Milling Machine Experiment....	83
Table 6.9: A Part of the Prediction Results of Week 3 of Milling Machine Experiment....	84
Table 6.10: A Part of the Dataset of Milling Machine Power Values for Decision Tree....	86
Table 6.11.: A Part of the Prediction Result of Week 1 Milling Machine Experiment.....	89
Table 6.12: A Part of the Prediction Result of Week 2 Milling Machine Experiment.....	90
Table 6.13: A Part of the Prediction Result of Week 3 Milling Machine Experiment.....	91

Table 6.14: A Part of the Dataset of Laser Sintering Machine for Time Calculation.....	94
Table 6.15: A Part of the Dataset of Cutting Machine for Time Calculation	103
Table 6.16: A Part of the Dataset of Milling Machine for Time Calculation	106

List of Figures

Figure 2.1: Bottom-Up CBA Process.....	17
Figure 2.2: The Relationship between Data, Information and Knowledge	18
Figure 2.3.: Relationship between the vectors	26
Figure 2.4: Support Vector Machine Techniques	27
Figure 2.5.: Example Decision Tree	31
Figure 2.6.: Summary of CBR Steps.....	35
Figure 5.1.: Active Power Graph of the Experiment of Company 1.....	50
Figure 5.2.: Active Power Graph of the Experiment of Company 2.....	53
Figure 5.3.: Active Power Graph of the Experiment of Laboratory of Liverpool John Moores University.....	57
Figure 6.1: Classification Tree of Status of Cutting Machine	74
Figure 6.2: Classification Tree of Status of Milling Machine	88
Figure 6.3: The Confusion Matrix of the Binary SVM and Decision Tree Algorithm.....	92
Figure 6.4: The Confusion Matrix of the Decision Tree Algorithm of “Shifting” Filter Method	92
Figure 7.1: Flow Diagram of the Base of the Knowledge Support System Database	112
Figure 7.2.: The Flowchart of the Summary of the Whole Database System.....	114
Figure 7.3.: Beginning of the Database.....	115
Figure 7.4.: Part 1 of the Flowchart of the Database	116
Figure 7.5.: Part 2 of the Flowchart of the Database	117
Figure 7.6.: Part 3 of the Flowchart of the Database	118

ABSTRACT

Nowadays, many small or medium size manufacturing companies are struggling to identify the right solution to tackle the problems of long production cycle time, poor quality and expensive cost in their manufacturing processes. In order to facilitate overcoming these struggles for small or medium companies, this thesis presents a research that intends to develop an effective knowledge management system, which includes a flexible sensor based cost model for calculation of unit manufacturing cost of a product, for small or medium size manufacturing companies with the potential of taking the advantages of Industry 4.0 which is an approach of generation of smart facilities with smart cyber-physical systems. (Thames & Schaefer, 2016) In this research, it is identified that the facilitation of suitable sensors to observe and monitor manufacturing processes could assess the efficiency of the manufacturing processes in real time and help the manufacturing process cost calculation to be more accurate. Besides, the cost model is converted from the traditional cost model to a dynamic one through adding flexibility to the model via two ways: Transferring the data of the cost parameters from the sensor on a machine to the knowledge support system directly via the internet and measuring the efficiency of a manufacturing process by calculating the utilization rate of the machine through utilizing the power values of the machine that comes from the sensor. In order to achieve flexibility in the cost model, three case studies were done. In each case study, manufacturing processes are observed via a power meter that can be easily assembled on the machines of interests. Subsequently, the efficiency measurement of the manufacturing process is provided through classifying the power values of the machine via an artificial intelligence method and the value of the utilization rate of the machine is obtained via utilizing these classified power values in Matlab Software. Lastly, the efficiency measurement of the manufacturing process and calculation of the unit manufacturing cost of a product are gathered into a knowledge management system that is created by utilising Microsoft Access Software. In this project, it is aimed to create a knowledge management system that has some benefits to small or medium companies. These benefits are presented below:

- The knowledge management system has a capability to collect data via the internet and this capability provides a foundation for establishing or improving Industry 4.0. approach in a facility.
- Another key beneficial feature of the database is the calculation of the manufacturing cost taking account of the real utilization rate of the machine and the data that comes

from the sensors on the machine. Therefore, the cost model becomes dynamic which means that the user can reflect the changes in a manufacturing process to the cost model and also it can monitor the efficiency of the manufacturing process via the value of the utilization rate of the machine.

In this project, how the real utilisation rate is monitored and integrated into the general cost model is presented together with the application cases of two industry companies and Liverpool John Moores University laboratory.

CHAPTER 1: INTRODUCTION

One of the main indicators of a successful small or medium company is customer satisfaction through serving high quality products that match customers' requirements. However, sometimes, companies have struggles about fulfilling customer needs in a cost-effective way. According to the industrial surveys in this research project, one of the reasons for these struggles is that the companies cannot notice cost related troubles in a manufacturing process effectively due to utilizing traditional cost estimation methods. In literature, several research papers mention disadvantages of traditional cost estimation methods: the traditional costing methods, which are fundamental techniques for costing, use only one cost driver in order to allocate overhead costs. This feature of the traditional costing methods can cause big differences between estimated and real cost prices (Kowsari, 2013). Furthermore, traditional costing methods allocate costs according to labour or machine hours which means that it is hard to uncover the actual reason and efficient relationship between overhead costs and each product in a facility (Rasiah, 2011). Lastly, it is a fact that traditional costing systems have struggles about fulfilling the requirements of managers; and also the results of these costing systems sometimes misdirect companies and cause them to make inaccurate decisions (Haddadi & Seyednezhad, 2015).

During the process of industrial surveys and literature review, several research questions were occurred: How can a traditional cost model provide updated information about a manufacturing process as well as give accurate results to the operator for making accurate decisions about the manufacturing process and quality control of the process? and How can industrial 4.0 approach be integrated to the cost model?

After having meetings with industrial partners, examining the literature and creating research questions, it is thought that generation of a cost model, that covers almost all direct and indirect cost parameters of a facility and transformation of the related cost model from traditional to dynamic model through adding data from sensors on a machine, can facilitate control and optimization of the manufacturing processes of SMEs. It is thought that transferring the data, from the sensors on a machine to the knowledge support system that includes the cost model, can provide not only flexibility but a base for Industry 4.0. approach in a facility. Therefore, this project offers a knowledge support system that includes a novel sensor-based cost model for calculating the unit manufacturing cost of a product. The sensor-based cost model covers direct and overhead costs of a product as well as a parameter, as an efficiency indicator, that is called the utilization rate of a machine in a manufacturing system of a product as a part of the manufacturing process.

The potential methodology of the project covers creating a sensor-based cost model with a suitable cost estimation method, adding utilization rate of the machine as an efficiency parameter to the model in order to convert the traditional cost model to dynamic cost model and calculating the efficiency parameter by classifying power values of the machine, which can be obtained from the sensor on the machine, through a suitable machine learning method and calculating the utilization rate of the machine automatically through using MATLAB software.

1.1. Aim of the Project:

The aim of this project is combining a knowledge support system with a novel sensor-based cost model, which is capable of calculating unit manufacturing cost more accurately, presenting the efficiency of the manufacturing process more clearly and dynamically to the user. As a result the knowledge support system can give essential information for the operator to rationally select manufacturing process steps and manufacturing techniques, in order to improve the performance of a manufacturing process, customer satisfaction and enabling the Industry 4.0 approach in a facility.

1.2. Objectives of the Project:

- In order to calculate unit manufacturing cost of a product, a generic cost model will be generated through determining direct and overhead cost parameters with the capability of taking account of machine utilisation dynamically.
- Experiments will be done through utilizing the machines of industrial partners and Liverpool John Moores University with a power meter in order to add the parameter of electricity energy consumption of the manufacturing process and utilization rate of the machine for measuring efficiency of the manufacturing process via the generic cost model.
- Develop and utilise Matlab codes for the automatic calculation of the utilization rate of the machine.
- Establish a knowledge support system database that includes the generic cost model and process steps information and previous cases (products) of some manufacturing techniques.
- A Case Based Reasoning (CBR) system will be generated in the knowledge support system in order to estimate utilization rate of a machine in a manufacturing process of a product based on similar previous cases.

Therefore, the presentation structure of the thesis is given as following:

- Chapter 2 constitute the literature review of the project. This chapter covers some example research about production process planning from literature and explanation of one of the production planning monitoring and control methods that is Industry 4.0 and a table presentation of the most commonly used cost estimation methods. Furthermore, it presents a detailed explanation of the knowledge support system that includes the examination of artificial intelligence, machine learning methods, data mining operations and Case Based Reasoning method.
- Chapter 3 presents a methodology overview and summarizes the content of each chapter of the methodology.
- Chapter 4 is created in order to explain the experiments with a laser sintering cutting and milling machine in Company 1, Company 2 and Liverpool John Moores University through assembling the power meter on the machines and presenting the results of the experiments.
- Chapter 5 includes a sensor-based cost model that is generated in order to calculate unit manufacturing cost. In this chapter, initially, the cost parameters of the model are explained. Afterwards, the cost parameters that are considered by Company 1 and Company 2 are given through two different tables. Finally, the sub-formulas of the cost model are presented and the general cost model is obtained by gathering these sub-formulas.
- Chapter 6 presents the assessment of manufacturing process efficiency. This chapter covers the application of some machine learning methods to the results of the experiments, which are given in Chapter 4, in Matlab Software in order to classify the power values of the machine based on machine status (stop, idle or working) and calculate the utilization rate of these machines.
- In Chapter 7, design and creation level of the knowledge support system is explained. Design level of the knowledge support system is stated via a flowchart and creation level of the knowledge support system is explained not only with the flowchart but also with the screenshots of the forms that were taken from the knowledge support system.
- The last chapter of the thesis is Chapter 8 that covers discussion, conclusion and future works of the project.

CHAPTER 2: LITERATURE REVIEW

In this chapter, there will be a literature research about current production process planning, monitoring and control, subsequently, there will be a review of production cost modelling and its techniques for selecting the most adequate one for the specified purposes, and the last sub-chapter is the definition of industry 4.0 because it is aimed to achieve Industry 4.0 approach in a facility through utilizing a sensor-based cost model that is created in this project.

2.1. Production Process Planning, Monitoring and Control:

Realistic planning and effective monitoring of a manufacturing process of a product can improve the quality of the product and shorten the cycle time significantly. Therefore, there are quite a few research papers about production process planning, monitoring and control in the literature. In this section, some of the works in the literature were researched and presented.

2.1.1. Manufacturing Process Planning:

First of all, a hybrid manufacturing system, which combines computer numerically controlled (CNC) machining continuum and layered deposition continuum and obtains the advantages of both continuums, is familiar and a joined process planning framework is searched. The related framework covers all modules of the hybrid production continuum and aims to automate the hybrid production. In order to obtain an automated hybrid production continuum, important constituents of process planning, for instance, the analysis of the computer aided design (CAD) model, development of the toolpath creation pattern and collision detection algorithms, are discussed and some experiments are carried out for confirming the practicality and reliability of the joined process planning framework (Ren, 2010). The next paper aims to create an Internet and Web-based service-oriented system through cloud manufacturing in order to monitor machine availability and plan the manufacturing process. Therefore, a tiered system architecture is offered and IEC 61499 function blocks, which transfer the machining technological information via various system modules until the end of the machining process, are presented in order to apply the prototype. This system provides real-time machine availability and execution status monitoring during metal-cutting operations, both locally or remotely (Wang, 2013). The next research searched for useful models of cloud-based systems as well. Subsequently, in the research, distributed process planning services are created and validated via the evaluation of the structure of the process planning services through a case study that the industrial users applied as a specific category of process planning service in the cloud. According to the research, distributed process planning (DPP) structure consists of two

basic hierarchy levels: Supervisory Planning that is in charge of creating a standard and machine-neutral continuum plan by grouping machining properties based on their datum references, tool access directions and some sequencing rules; and Operation Planning Level in which a detailed continuum plan is developed through determining all the decisions about the tool and fixture options, tool path planning and tool parameter option. Traditionally, altering machines or cutters requires hours or even days of extra effort. Under favour of the Cloud – DPP system, new continuum plans can be developed in minutes (Wang, 2017). In the last paper, according to the authors, there is a relationship between process planning and process scheduling and combining these two terms is critical in order to increase the flexibility of scheduling and accomplishing a universal advancement of the performance of the production system. Hence, a mathematical model for the integrated process planning and scheduling (IPPS) is created and an improved genetic algorithm (IGA) is offered in order to ease the optimization of continuum planning and scheduling at the same time. In the verification process of the proposed approach on IPPS, experimental studies are done and this approach is compared with the other approaches with the makespan and mean flow time performance measures. According to the verification results, the proposed approach on IPPS has accomplished remarkable enhancement about minimizing makespan and taken good results for the mean flow time performance measure with high efficiency (Lihong & Shengping, 2012).

2.1.2. Manufacturing Process Monitoring and Control:

Manufacturing process monitoring and control can be important for the efficiency of the manufacturing process as much as manufacturing process planning. In this section, Industry 4.0, which is one of the latest manufacturing monitoring and controlling approaches will be explained.

2.1.2.1. Industry 4.0:

Industry 4.0 concentrates on the development of the smart products and manufacturing continuums (Brettel, 2014). However, alteration from the current position to smarter machines needs more progression through handling several basic points. These points can be separated into five different groups:

- **Manager and Operator Communication:** Even though assigned duties are generally optimized by professional operators and managers, a very significant aspect is lacking in these choices: how sturdy the machine components are.
- **Machine Fleet:** In the current situation, achievable prognostic and health management techniques are not useful considering same machines, which have dissimilar working environments for dissimilar operations, as a fleet through collecting worthwhile information from different cases.
- **Product and Continuum Quality:** A feedback that can be utilized to enhance manufacturing planning can be ensured by product quality. Nowadays, there is no this kind of feedback loops.
- **Big Data and Cloud:** The significance of leveraging further flexibility and abilities that are presented by cloud computing is unavoidable, however transcribing prognostics and health management algorithms to data management technologies, which are currently applied in an efficient way, needs more research and improvement.
- **Sensor and Controller Network:** Physical environment of a machine is identified by machine sensors. But sensor defects and breakdown can cause incorrect and inaccurate readings to decision-making algorithms and this situation can give rise to a wrong outcome (Lee, 2014).

There are some process monitoring and control processes in which artificial intelligence methods were utilized in the literature:

Abellan-Nebot and Subiron (2010) considered six main issues of the generation of intelligent machining systems in order to introduce a generic overview of a machining monitoring system and ease its utilization. (1) Different sensor systems were implemented during monitoring machining continuums. (2) The most efficient signal processing methods were taken into account. (3) The most common sensory properties were implemented during modelling machining continuums (4) The sensory property picking and removal techniques were considered in order to utilize accurate sensory information. (5) the experiments were designed based on the model of a machining operation that needs a minimum amount of experiment data. (6) The main features of several artificial intelligence methods were utilized for easing the application or selection of sensor systems (Abellan-Nebot & Subirón, 2010). Moreover, in another paper, the utilization of artificial intelligence methods, which can provide evidence of the potential of modern CNC machines that can monitor main drive power, was suggested for real- time forecasting surface roughness deviations in a machining process. In order to obtain the surface roughness, three parameters should be estimated: maximum tool wear,

machining time and cutting power. Several AI methods were tested for estimation of these parameters: Random Forest (RF), Standard Multilayer Perceptrons (MLP), Regression Trees, and radial- based functions. In conclusion, it is seen that Random Forest has the most model accuracy compared with other methods (Pimenov, 2018). The next paper presents the principles of utilizing neural networks and elements of AI in the continuums of generation, distribution and consumption of electricity. As a conclusion, it has been seen that failures in power supply can be minimized via the utilization of artificial intelligence in the electric power industry (Sozontov, 2019). Moreover, an examination is done about industry 4.0. approach in smart manufacturing systems in order to make a deep research on the usage of Industry 4.0. This paper investigates industry 4.0. approach in smart manufacturing systems through creating four steps: Firstly, theoretical structure of smart manufacturing systems for industry 4.0. is offered. Next, indicative scenarios, are which related to smart design, smart machining, smart control, smart monitoring, and smart scheduling, are given. Subsequently, based on these indicative scenarios, key technologies and their probable implementations to Industry 4.0 in smart manufacturing systems are evaluated. Lastly, possible problems and future evaluations are decided and discussed (Zheng, et al., 2018). Another paper purposes to comprehend the adaptation patterns of Industry 4.0 technologies in manufacturing companies. The paper offers a theoretical structure for these technologies that are separated into two parts: Front-end and Base Technologies. Front-end technologies include Smart Manufacturing, Smart Products, Smart Supply Chain and Smart Working as well as Base Technologies cover internet of things, cloud services, big data and analytics. A survey is applied to 92 manufacturing companies in order to analyse the application of these technologies. The result indicates that Industry 4.0 is associated with a systemic adoption of the front-end technologies of which smart manufacturing is in the center. On the other hand, the application of Base Technologies is struggle for the companies because big data and analytics are still low applied in the sample studied (Frank, 2019). Last paper aims to detect and search the Industry 4.0 technologies that can be implemented to production logistics and how the manufacturing environment affects the practicality of these technologies. These aims are achieved by completing multiple case studies of four Norwegian manufacturing companies. The results of the work show that the practicality of industry 4.0 in manufacturing logistics is based on production environment. Companies that have low degree of manufacturing iteration think that application of Industry 4.0 technologies in manufacturing logistics have less possibility. On the other hand, companies that have high degree of manufacturing iteration see high possibility in implementing Industry 4.0. technologies in manufacturing logistics (Strandhagen, 2017).

2.2. Production Costing:

The types and definitions of the term “Cost” were made by Ruth (2006) in the literature and the definitions were given as follows:

What is “Cost”?: It is total money, time, and resources that are related to a purchase or activity.

Fixed Cost: It covers all costs that any accounting period activity does not change. Fixed costs are, at any time, certain costs that have to be paid without considering the amount of output and of the resources used.

Variable Cost: It is the costs that are some function of activity. They are generally linear function because the unit cost is calculated through dividing the total other costs for a period, or event, by the amount of activity in the period. Combination of fixed and variable costs create the total cost. The total variable cost alters according to increased usage, on the other hand, the total fixed cost stays constant.

Total costs = Fixed + Variable Total Cost (Ruth, 2006)

All costing systems can be defined as multi-purpose tools which means that costing systems consist of an enterprise model and, on the other hand, they calculate cost of the objects. Beginning of the any calculation is ensured by hard figures from the cost model (allocation bases) independent from their title. Subsequently, these hard figures are joined to technical data which is about cost objects that are gathered in bills of materials. The term of “Bill of Materials” comes from manufacturing management and it emphasises that every object is a complicated group of tangible and intangible constituents of work presented in an establishment reflected entirely or partially by the cost model (Mevellec, 2009).

Although there are many different cost prediction methods in literature, many authors approve that these methods can be grouped into two basic topics: Qualitative and Quantitative Methods.

Table 2.1: Cost Estimation Methods

	Method	Explanations	Commentary notes	Reference
Qualitative Methods	Intuitive	Intuitive Cost Estimation is based on expert judgement that only depends on the experience of one or more than one expert.	In order to forecast the costs and benefits in a product line improvement organization, the Structured Intuitive Model of Product Line Economics (SIMPLE) is generated. Through SIMPLE, some decisions can be made; for example: utilization of a product line strategy in a particular situation, the particular strategy to implement and the suitability of obtaining or generating a particular asset.	(Angelis & Stamelos, 2000) (Clements, 2005)
	Analogical	Analogical Cost Estimation depends on comparison that means comparing physically similar features of a new product and current products. This method evaluates the cost of a product through comparing it with the costs of other existing similar products. Analogical Cost Estimation includes identifying coding of the components, frequently a morpho-dimensional coding and comparison physical properties and functions of new product and existing products.	In design process for Additive Manufacturing (AM), the decision of utilization of AM is required to be taken in early level of the design process. Therefore, it is aimed to generate suitable cost forecasting models that concentrate on purchase cases because a lot of organizations prefer to purchase components from service providers. In order to analyse data for the cost model, analogical cost estimation methods were implemented to a data set of market price for laser sintering and laser melting components.	(Camargo, 2003) (Baldinger, 2016)
Quantitative Methods	Parametric	Parametric Cost Estimation technique is based on the utilization of cost estimating relationship (CER). That is: mathematical expressions or formulas about cost consist of appendant selected variable and independent cost-driving variables. The relationship between appendant and independent variables depends on regression of historical data. Practically, many parametric cost estimation methods are close to top-down techniques due to selected independent variables and the utilization of parametric forecasting in early level of a design of the product.	In the paper, there is comparison between several cost estimation methods for the textile area: their advantages, disadvantages and practicability in the product life cycle. According to the result of the comparison, it is seen that parametric cost estimation technique is the most suitable method in the earliest level of the product process from design stage to cost stage. Besides, the paper presents several ways to implement parametric cost estimation techniques that include the processes and tools of the techniques in the textile and garment industries. Lastly, parametric cost estimation method has been implemented for forecasting the unit cost of a specific family of wool textile fabrics.	(Shetelig, 2013) (Camargo, 2003) (Smart & Culver, 2009)

Activity- Based Costing	<p>Activity-Based Costing (ABC) is a complicated cost estimation method because it eases the manager's job through making significant strategic business decisions. According to the literature, ABC has different stages: First one is Activity Analysis (AA) that is the initial level to define activities and processes in order to generate the final products or services. Second stage of Activity Cost Analysis is the identification of the cost of each activity and cost drivers and last stage is Activity-Based costing that is checking costs of activities to products or services.</p>	<p>A research approach concentrated on the innate advantages and disadvantages of activity-based costing method for cost accounting inside of a company. The idea of inductive hypothesis is that through obtaining theoretical insights, the current model can be imitated, validated or changed by practical insights that are the result of a new comprehensive structure.</p>	<p>(Askarany, 2010) (Schulze, 2012) (Javid, 2016)</p>
Top-Down Costing	<p>In summary, through the top-down costing approach, overall expenses are covered for each input at a central stage. Subsequently, the costs are divided by utilizing the formula, which depends on allocation determinants, in order to forecast unit process or service cost.</p> <p>Top-down macro costing, generally created based on econometric models that are appendages of the Harrod-Domar model, is utilized in order to calculate the needed investments that are for achieving a target growth rate. The financing gap between needed investment and accessible assets, which occurs after the utilization of the model, is generally presumed to be filled with overseas help. A top-down model would frequently be pursuant to a given incremental capital output ratio (ICOR) that is an extensive measurement of the productivity of investments in each country. It would get results in annually gathered predictions of investment costs, which in turn could be divided into budget classification codes utilized in the planning and accounting systems. More frequently, however, investment costs, created by top-down econometric models are utilized as total predictions of total costs, which in turn are used as checks against more detailed cost estimates.</p>	<p>Users of energy-economy policy models generally implement top-down and bottom –up cost approaches and different hypotheses about future technologies and the choices of companies and managers. In this paper, a hybrid energy-economy policy model has high technology, similar to a bottom-up model, and obtains empirically forecasted behavioural parameters for risk and technology choices like a top- down model. However, dissimilar to top-down approach, it replicates technological alterations endogenously through functions that are about the financial costs of technologies in order to ensure accretive production and arrange technology choices as market shares alteration.</p>	<p>(Cunnama, et al., 2016) (Jaccard, 2004) (Kim & Park, 2006)</p>

	Bottom- Up (Analytical) Costing	<p>Conventionally, bottom-up costing (process-based cost estimation) is for forecasting costs that cover the cost of each process that is required in order to manufacture the final product. The bottom-up part of this method can be done through forecasting each process cost individually as a subcomponent and subsequently, the cost of each subcomponent is gathered and summed up in order to generate a total cost estimation.</p>	<p>For the last four years, Reichhold, which is a divisional-structured company producing chemical adhesives, resins, and polymer systems, has been applying activity-based costing (ABC) with bottom-up approach in order to examine the cost implications of the alterations happening in manufacturing process as a result of the shift to specialty products.</p> <p>On the other hand, a study in the literature compares cost forecasting of the treatment of acute myeloid leukaemia between September 2004 and August 2007 in the Haematological Malignancy Research Network (HMRN), from top-down and bottom – up costing approaches and analyses their affects.</p> <p>Lastly, one of the papers searched three cost estimation models, which are multiple regression analysis as an analytical method, neural networks and case-based reasoning, through utilizing the data of 530 historical costs. In the sequel of the examination and application of the cost estimation method, it is seen that neural network forecasting model gives the most accurate result of the three; However, the case-based reasoning forecasting model had better performance than neural networks method corresponding with long-term utilization, obtainable information from the results and time-accuracy trade-offs.</p>	<p>(Toth, 2006)</p> <p>(Blocher, 2002)</p> <p>(Hi, 2015)</p> <p>(Carolus, 2018)</p> <p>(Kim, 2004)</p>
--	--	--	--	--

2.2.1. Some Detailed Explanations or Examples of Cost Estimation Methods:

2.2.1.1. Analogical Cost Estimation:

As an example of utilization of this method, analogical cost estimation of National Aeronautics and Space Administration (NASA) can be given:

Predictors of NASA utilize this technique during the early stage of a new program or system that has inadequate and undeveloped cost data. Cost data of the current system, which is technically similar to the predicted cost model of the new system, is used as the Basis of Estimate (BoE). Afterwards, cost data are arranged upward or downward according to their subjects based on whether the subject system is seen to be more or less complicated than the analogous system. Linear extrapolations from the analogous system are suitable alterations, when it is presumed that a valid linear relationship exists. There is an example of analogical cost estimation given below:

	Previous System	New System
Solar Array	A	B
Power	2.3 KW	3.4 KW
Solar Array Cost	\$10M	?

Presuming a linear relationship between power and cost, and presuming also that power is a cost driver parameter of solar array cost, the single-point analogy calculation can be given in following (NASA, 2015):

$$\text{Solar Array Cost for System B} = 3.4/2.3 * \$10M = \$14.8M \quad (2-1)$$

2.2.1.2. Parametric Cost Estimation:

The stages of the Cost Estimating Relationship (CER), which is mathematical expressions or formulas that are the foundation of parametric cost estimation are presented below:

- First stage starts with experience and data of the analyst. According to the experience and data, the most suitable model is selected by the analyst.
- In the second stage, after the model selection, the most suitable technique for calculating the parameter of the model is chosen by the analyst. Examples of this technique are statistical methods, neural network and bottom-up technique approaches.

- After a technique is selected and the parameters are calculated, the model generation process is completed for most of the analysts. Subsequently, CER is implemented through considering the suitability of CER for the data or whether the model theories have been validated (Datta & Roy, 2010).

2.2.1.3. Activity-Based Costing:

The utilization steps of ABC Costing from articles in literature, that include calculating the unit cost of medical services, and to evaluate their applicability in a hospital in Iran, is given below as an example:

- **Step 1: Hospital Analysis and Cost Centres' Classification:** The hospital was separated into numerous patient care cost centres (PCCs), which represented services to patients and covered staff departments, laboratory and heart surgery, and supportive cost centres (SCCs) that are responsible for providing facilities and services to all units for instance management and accounting.
- **Step 2: Defining Main Activities:** In this step, initially, extensive division was done between two groups of activities: treatment-related activities that were directly related to staff units; and supporting activities that support the treatment continuum and the department.
- **Step 3: Describing Activity Cost Drivers:** The cost driver for treatment-related and supporting activities is decided in this step. Subsequently, cost driver rates for each activity are calculated.
- **Step 4: Appointing Costs to Cost Centre:** The first level of the division continuum within the ABC system is appointing costs to cost centres. In order to divide indirect costs, a resource cost driver must be decided. In this level, costs were allocated to activity centres through utilizing the related cost drivers: staff workload, quantity of equipment, floor area, and estimation.
- **Step 5: Calculation the Unit Cost of Activity and Services:** The indirect costs were appointed to particular activities, and the activity costs were appointed to cost objects. The needed materials along with labour time were decided based on the planned interviews with the person involved for each activity. The costs of each part of the activity were gathered into the cost of the activity:

Activity cost = cost of floor area + cost of materials + staff costs + cost of capital equipment + cost of hospital infrastructure + other indirect costs (Javid, 2016).

2.2.1.4. Bottom-Up Costing:

In literature, there is an example of the utilization of the bottom-up approach in Cost Benefit Analysis (CBA):

Cost Benefit Analysis was applied on two cases: The Helge River in Sweden and the Berze River in Latvia. Both cases have experienced struggles about water quality that creates focal environmental problem. The bottom-up CBA generation continuum covers numerous stakeholder workshops in both case areas that assessed the costs and benefits of offered strategies, the strategies themselves, and the stakeholder arrangement were negotiated, arranged and validated. The whole bottom-up CBA process is shown below:

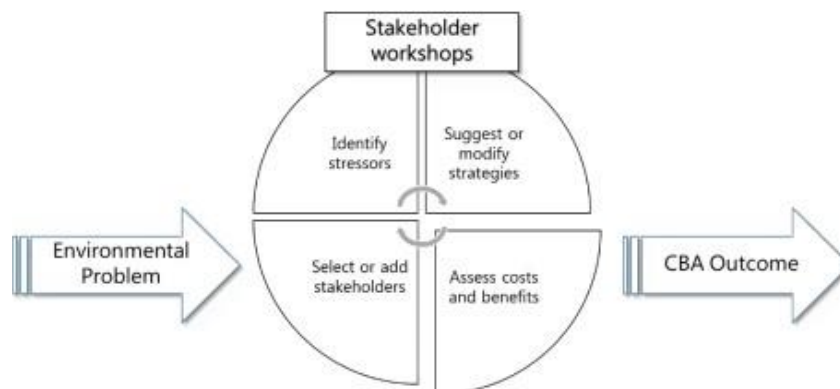


Figure 2.1: Bottom-Up CBA Process (Carolus, Hanley, Olsen, & Pedersen, 2018)

After explaining cost estimation techniques in Table 2.1., some examples from literature, which covers cost estimation techniques that includes artificial intelligence (AI), are given in the following:

In the literature, a research paper examined the reliability of multiple regression analysis (MRA), artificial neural networks (ANNs), case-based reasoning (CBR), and hybrid intelligence (HI) in order to estimate the cost of thin-film transistor liquid-crystal display (TFT-LCD) equipment. After analysis of the performance, reliability and prediction capability of the models that were created by using these four artificial intelligence methods, the results of the analysis show that the hybrid intelligence method (HI) is better than the other three AI methods in cost estimation during the design stage (Chou, 2010). Another paper created and tested a model based on cost forecasting of piping elements in the early design level for easing the decision-making process of the designers by utilizing

artificial neural networks (ANN). Subsequently, the created model is compared with other models that were created via applying conventional regression models. The comparison indicated that ANN is able to decrease uncertainties about the cost forecasting of shell and tube heat exchangers (Duran, 2012). The last example in the literature utilized the ‘least squares support vector machines’ (LS-SVM) technique in order to estimate the production cost of airframe structural projects. Before applying LS-SVM method, the research evaluated forecasting performance of the other two cost estimation models through applying back-propagation neural networks and statistical regression analysis methods. In the end of the research, it is seen that the LS-SVM model can ensure reliable forecasting performance and is better than other methods (Deng & Yeh, 2011).

2.3. Knowledge Management Systems:

Data: Pure, unformed or raw realities can be the definition of the term of data. These realities can be measurements, statistics, numbers or alphabets.

Information: Treated data is called “Information”. In other words, information is an output that is obtained from data.

Knowledge: Knowledge can be described as an idea that is believed to be justifiable and right. On the other hand, some people in literature state that knowledge is the implementation of information.

The relationship between data, information and knowledge is presented in Figure 2.2:

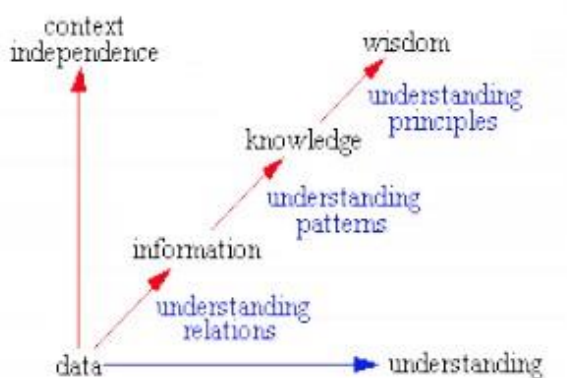


Figure 2.2: The Relationship between Data, Information and Knowledge (Edosio, 2014)

Knowledge Management: Knowledge Management (KM) is referred to as the continuum of knowledge generation, confirmation, presentation, deployment and implementation. These five stages of Knowledge Management, which allow a company to learn, externalize, unlearn and learn again, are generally accepted significant in order to establish, protect and refill core-proficiencies (Bhatt, 2001).

Advanced Information Technologies, such as the Internet, intranets, extra-nets, browsers, data storage devices, data mining methods and software agents, could be utilized in order to systematize, increase and accelerate large scale intra- and inter-company knowledge management. Therefore, organizational and administrative application have become more knowledge-oriented these days. For instance, benchmarking, knowledge controls, best application transfer, and staff improvement indicate the perception of the significance of organizational knowledge and intangible entities generally (Alavi & Leidner, 2001).

Four basic parts of Knowledge-Based Systems are generally recognised: a knowledge base, a deduction engine, a knowledge engineering tool and a particular user interface. Moreover, the locution of Knowledge Based Systems covers all the organisational information technology implementations in order to control the knowledge entities, for instance expert systems, rule-based framework, groupware, and database management systems (Liao, 2003).

One example of a knowledge management system is IDeA. The IDeA has a continuing programme of business with regional government co-workers in order to facilitate knowledge flow through regional government, implementing tests to Knowledge Management tools and methods for their suitability in the sector (Leask, 2008).

However, there are some limitations of knowledge management systems that are mentioned below:

- A precise and empirical investigation is required in order to find the suitable knowledge management tools for global software development because there are a lot of types of knowledge management tools and techniques in literature.
- It can be compelling to constitute knowledge management tools with different cultural settings because shaping, structuring and using the knowledge can change from country to country (Dingsøy, 2004).

There are many data management software packages in the market in order to create a knowledge-based system for an organization, such as Oracle Relational Database Management System, Microsoft SQL Server, SAP Adaptive Server Enterprise (ASE), Teradata, FileMaker and Microsoft Access.

There are some papers, of which knowledge management systems is used or searched inside, in literature. For instance, a framework that has four steps is offered for self-progressing additive manufacturing (AM) knowledge management. This framework has two continuums: Bottom-up data-driven knowledge engineering and top-down goal-oriented active data creation. These continuums are working parallel and they are attached by the users. Hence, a lose loop are created so that AM knowledge can develop without interruption and automatically (Lu, 2018). Furthermore, another article presents the analysis of the performance of Knowledge Management (KM) Dashboard. KM Dashboard was created, and the prototype was applied in order to facilitate manufacturing and operational efficiency observing in a Manufacturing SME when a resource is limited. The analysis of the performance of the dashboard had been conducted in two small-medium dairy processing plants in South Western Uganda for 6 months. An observational study and a survey that focused on usefulness (the amount of the contribution of the Dashboard to decision making process), usability (the amount of the flexibility of the utilization of the Dashboard) and usage (the real application of the dashboard according to the user.) of the Dashboard was applied to 8 attenders as key informers. All the participants think that the Dashboard is very useful. 88% of the attenders say that the Dashboard showed its usage and lastly, 100% of the attenders think that the Dashboard is usable (Twongyirwe & Jude, 2018). Last example is the paper that offers some solutions to a company which has some problems about the process of the engineering design and manufacturing of plastic injection molds. The company took an approach of presenting knowledge management system in order to decrease the risk and prevent the challenges that had already been detected in the past. Knowledge management system model has been created and possible remedies have been tested according to identified requirements through preliminary analysis. Subsequently, the most appropriate system was applied in the company (Marjanović, 2018).

One type of knowledge management is Artificial Intelligence (AI). In AI, knowledge management covers the detection and examination of accessible and needed knowledge resources and continuums that are about knowledge resources, and the following planning and control of actions in order to improve the resources as well as the continuums for the purpose of achieving organizational goals (Girard & Girard, 2015). A detailed definition of artificial intelligence is given in the following:

2.3.1. Artificial Intelligence:

Artificial Intelligence (AI) is a program that runs in a way that is similar to a human behaviour (Lozano-Pérez & Kaelbling, 2002).

The AI research contains wide kind of problems. Some of them are given below:

- In order to do reasoning, pattern identification, learning or some other format of illation, computers are utilized.
- Problems that cannot be solved with algorithmic techniques are concerned. This situation is the basic reason for the trust in heuristic search in AI problem-solving methods.
- Utilization of incorrect, missing or not clearly defined information and the usage of representational formalisms, which make compensating possible for these problems, are the limitations in the problem-solving methods.
- When ideal or definite results are excessively expensive or impossible, there is the result of significant trust in heuristic problem-solving techniques: Not definite or ideal answers, however, they are adequate in some cases (Luger, 2005).

On the other hand, there are some advantages of AI approaches:

- The implementation of a continuum of knowledge acquirement can clarify the complex interrelationships between a series of system elements and activities.
- Creating an automated system, letting incorporation of implicit and frequent qualitative opinions that are counted significant by engineers and/or operators can fill the gap between results, that is created from elaborate modelling endeavours and feasibility of those results to a practical case (Chan & Huang, 2003).

2.3.1.1. Machine Learning:

Machine learning comprises of three components: Presentation, Assessment and Optimization. Presentation means representation of a classifier in an official language that a computer can cope with. Generation of a group set of classifiers, which the learner can understand, is very critical. Moreover, Assessment Component shows that an analysis is required for separating useful classifiers from unserviceable classifiers. Lastly, Optimization Function is applying a technique in order to investigate among the classifiers to find the classifier that has the highest score. It is critical to choose the most suitable optimization technique because it is the base of efficiency of the machine learning algorithm (Domingos, 2012).

Some machine learning methods are presented in Table 2.2.:

Table 2.2.: Machine Learning Methods

Methods	Explanation	Reference	Advantages and Disadvantages
Hidden Markov Model	A Markov model is basically a mathematical demonstration of a Markov continuum and it utilizes the language of probability theory. Hidden Markov Model works on two levels: observed levels and hidden levels. A system can only provide an observable series of events that are controlled by a hidden Markov chain that cannot be observed instead of being causally related. (Leung, 2014)	(Leung, 2014) (Holmes & Huckvale, 1994) (Degirmenci, 2014)	Based on literature, one of the advantages of Hidden Markov Model is that the method ensures a complete model that can be applied in higher levels through determining a single decision about finding the best match of the model parameters at all levels. Another advantage is that Hidden Markov Model Method presents a mathematical foundation that can be examined and controlled easily and analytically. On the other hand, there are some limitations of HMM; For example, in the problems that have wide parameter spaces, the transition matrix can be very large and this situation can cause significant overfitting.
Decision Trees	The problem of generating decision trees can be revealed repetitively. Initially, an attribute is chosen in order to locate at the root	(Witten, 2011) (Westreich, 2010)	According to the one of the papers in the literature, decision trees are well-researched and they are easily adapted and applied to a problem. Besides, decision trees are applicable

	<p>node and, for each possible value, one branch is created. This branch divides example set into subsets and each subset is for every value of the attribute. After this step, the continuum can be renewed repetitively for each branch through utilizing only the cases that actually reach the branch. If all cases that are located at a node have the same classification at any time, development of the related part of tree should be stopped.</p>	<p>(Joshi, 2019)</p> <p>(Rokach, Decision forest: Twenty years of research, 2016)</p>	<p>for standard software packages. Lastly, the equations of other machine learning methods, such as neural network or support vector machine, were created to run only for numerical data. However, decision tree algorithm does not count on only the presumption of numerical data. On the other hand, there are some disadvantages of the decision tree method: If the attributes are not so many and they are highly relevant, decision tree algorithm runs effectively. However, decision tree algorithms may not work well if there are complex relationship between the attributes. Another limitation of decision tree is that the algorithm can consider only one level ahead. Lastly, decision tree algorithm is highly sensitive to the noises of the training data set. The noises make the algorithm unstable.</p>
<p>Bayesian Networks</p>	<p>Bayesian Networks consist of the combination of graph theory and probability theory which means that modelling stochastic and causal connections can be possible for many kinds of decision – support problems. A Bayesian Network is comprised of a directed acyclic graph (DAG) which presents relationships between related nodes (variables) a set of local probability distributions that are connected to each node</p>	<p>(Zhou, 2014)</p> <p>(Uusitalo, 2007)</p>	<p>Bayesian Networks have many properties that make this method preferable for users. For instance, data combinations with information field are allowed in Bayesian Networks, a method is ensured in Bayesian Networks in order to prevent overfitting as well as Bayesian Networks can provide an efficient forecasting accuracy even with small data group. However, Bayesian Networks may not be suitable for continuous variables. The values should be discretized for obtaining a</p>

	that measures the strengths of these relationships.		good result. Besides, although Bayesian Networks are a practical method for modelling expert knowledge, it may be a struggle to obtain the knowledge out of the experts as probability distributions.
Artificial Neural Network Algorithm	Artificial Neural Networks are complicated non-linear method that create new features, which are called linear compositions of the beginning predictors and they model the output variable that is a non-linear function of related newly created features.	(Loyer, 2016) (Mijwel, 2018)	There are some advantages and limitations of Artificial Neural Network Algorithm. Some advantages of Artificial Neural Network (ANN) Algorithm are that, in ANN that is one of conventional programming methods, information is gathered in the entire network, not in the database. Therefore, even if a few pieces of information disappeared, the quality of functions of the network does not drop. Besides, after ANN is trained, the data may present an output even via incomplete information. However, one of the most important disadvantages of ANN is that when a detailed solution is generated by ANN, the algorithm does not give a hint about why and how it produced this solution. This situation decreases the reliability of this method. Moreover, a specific rule does not exist in the process of deciding the structure of ANN networks. Suitable network structure is obtained via experience and trial and error.
Multiple Linear Regression Analysis	If more parameters are added in linear regression analysis in an	(Pandis, 2016)	Multiple Linear Regression Analysis is a useful method because the relationship

	<p>effort in order to obtain factors that forecast the outcome better, it is called multiple linear regression analysis.</p>	<p>(Jeon, 2015) (Sharma & Kumar, 2017)</p>	<p>between dependent and independent variables can be modelled statistically through this method. However, the reliability of the model, which was created via utilizing Multiple Linear Regression analysis, may decrease, if all variables that are in the model are not independent of each other and have multicollinearity among themselves.</p>
<p>Support Vector Machine</p>	<p>A support vector machine (SVM) is a non-probabilistic binary linear classifier. The SVM divides the whole data set in a feature space into two groups. The data set should have features $\{x_1... x_n\}$ and a group label, y_i. SVM treats each data object as a point in feature space for example the object is allocated into one group or the other group.</p>	<p>(Awad & Khanna, 2015) (Auria & Moro, 2008)</p>	<p>SVM can be a practical method for insolvency analysis of which the data is not regular. For instance, when the distribution of the data is not regular or the distribution of the data is unknown, SVM can facilitate analysis of the information. Besides, as long as the optimality problem is convex, a unique solution can be obtained from SVM analysis. On the other hand, because of SVM is one of the non- parametric techniques, the results of SVM are not completely transparent. For example, SVM cannot present all companies' score as simple parametric function of the financial ratios because of very high dimensions.</p> <p>After the literature review about SVM, it can be seen that this method is suitable for binary classification and it is not very useful for large data sets.</p>

After the evaluation of the machine learning techniques and their advantages and limitations, some machine learning methods that are considered for utilizing in this project are given below:

- Support Vector Machines (SVM) Methodology:

According to the definition of the original term of “Support Vector Machine”, the technique includes a group of labelled data examples and the goal of the SVM algorithm is to discover a hyperplane, which distinguishes the data set into a number of independent classes, cohere with the training examples. The decision limit that minimizes misclassifications is called optimal separation hyperplane and it is gained in the training step. Moreover, a repetitive continuum of investigating a classifier though optimal decision boundary in order to differentiate the training patterns, which are generally located in high dimensional space, and, subsequently, distinguish simulation data under the same configurations (dimensions) is called Learning (Mountrakis, 2011).

Definitions of Support Vector Machine Approach

- Length of a Vector: The length of a vector is called x:

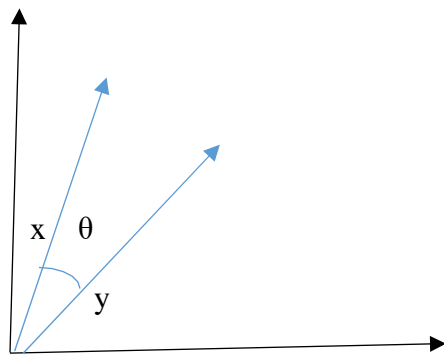
$$||x|| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}. \quad (2-2)$$

- Direction of a Vector: The direction of a vector, which is $x=(x_1, x_2)$, is shown as w:

$$w = [(x_1/||x||), (x_2/||x||)]. \quad (2-3)$$

The other name of the direction of the vector is unit vector.

Dot Product: This term presents the relationship of two vectors. There are two vectors (x,y) and the angle between them is θ :



$$x \cdot y = ||x|| ||y|| \cos(\theta) \quad (2-4)$$

Figure 2.3: Relationship between the vectors (Fan, 2018)

Support Vector Classification: If SVM is utilized for classification, it is called Support Vector Classification (SVC) and the aim of SVC is to investigate a classification criterion, which can be a decision function, in order to distinguish hidden data with a good generalization skill in a testing level.

Support vector machine includes 2 techniques: Linearly separable data in a case and non-linearly separable (linearly non-separable) data in a case. These techniques are presented as graphs below:

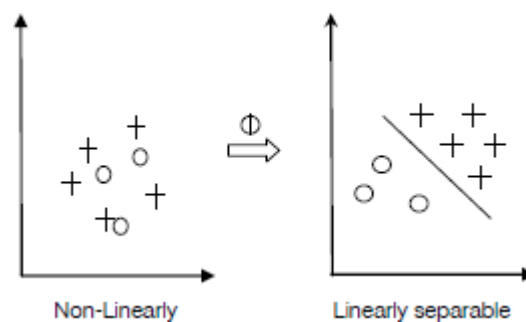


Figure 2.4.: Support Vector Machine Techniques (Ahuja & Yadav, 2012)

- a) Linearly Separable Case (Hard Margin): This case can be a linear straight line (linear hyperplane) that has maximum margin (distance) from the data of each class in two-class data classification. The linear hyperplane should have two features:
 - The hyperplane should have the minimum error possible in distinguishing the data.
 - The hyperplane should have maximum distance from the closest data.
- b) Linearly Non-separable Case (Soft Margin): The data of some cases cannot be separated linearly because of similarity of few properties in the database. However, linear SVM can be useful for this case if a penalty function could be identified in which the distance between the inadequate graded data of each class from the margin of related class could be gauged and minimized (Gholami & Fakhari, 2017).

Loss Function:

- a) Hard Margin Loss Function: In SVM, which can do binary classification, when the target value is $y_i \in \{+1, -1\}$

The loss function (ℓ_h) formula is:

$$\ell_h(y_x \cdot f_x) = \begin{cases} 0 & \text{if } y_x \cdot f_x \geq +1; \\ +\infty & \text{otherwise} \end{cases} \quad (2-5)$$

Hard margin loss function is suitable for the datasets that have not any noise.

- b) Soft Margin Loss Function: Soft margin loss function is preferred more frequently. The loss function (ℓ_ρ) formula is:

$$\ell_\rho(y_x \cdot f_x) = \begin{cases} 0 & \text{if } y_x \cdot f_x \geq +1; \\ \frac{1}{\rho}(1 - y_x \cdot f_x)^\rho & \text{otherwise,} \end{cases} \quad (2-6)$$

ρ = Any positive integer

- c) L_e Loss Function: As loss function leads to a convex programming problem for any positive integer ρ ; for L1 ($\rho = 1$) or L2 ($\rho = 2$) soft margin, it is a convex quadratic programming problem.

When L1 and L2 are generalized, ℓ_ϵ loss function is obtained. It is given in the following (Chu, 2002):

$$\ell_\epsilon(y_x \cdot f_x) = \begin{cases} 0 & \text{if } y_x \cdot f_x > 1; \\ \frac{(1 - y_x \cdot f_x)^2}{4\epsilon} & \text{if } 1 \geq y_x \cdot f_x \geq 1 - 2\epsilon; \\ (1 - y_x \cdot f_x) - \epsilon & \text{otherwise,} \end{cases} \quad (2-7)$$

where $\epsilon > 0$.

Although machine learning techniques are more reliable, have less fault and the techniques are more inclined to give its own decisions and solve the problem compared to data mining methods, data mining process is required likewise the machine learning because it will identify the problem of a particular case and solve the related problem likewise machine learning techniques.

Mostly, data mining is utilized for forecasting the result from historical data or finding a new solution from current data. (EDUCBA, 2019)

Multi-Class Support Vector Machine:

The standard Support Vector Machine (SVM) is created in order to classify the data into two classes only. This kind of classification is called binary classification. On the other hand, decomposition of several binary problems, which can be solved by standard SVM, can generally solve multi-class classification problems. For example, one-against-all decomposition is utilized frequently. In this situation, the classification problem to k classes is divided into k dichotomic decisions $f_m(x)$, $m \in K = \{1, \dots, k\}$, where the rule $f_m(x)$ distinguishes training data of the m-th class from the other training patterns. The classification of a pattern x is performed based on maximal value of functions $f_m(x)$, $m \in K$, for example:

the label of x is computed as $\operatorname{argmax}_{m \in K} f_m(x)$ (Franc & Hlavac, 2002).

Multi-class support vector machine has two popular algorithms in order to generate and gather several SVMs for multi-class problems:

- **One-Against-All (One-Against-Rest)**: It is also called standard technique and it contains N different classifiers where N represents number of classes. The i th classifier is trained during referring all the samples in the i th class as positive and the others as negative (Vural & Dy, 2004).

If the number of training data in each class is balanced, the sub-problem is an unbalanced binary classification problem, and can be presented as,

$$\begin{aligned} \min_{w_j, b_j} \quad & \frac{1}{2} \|w_j\|_2^2 + C \sum_{i=1}^n \varepsilon_i^j \\ \text{s. t.} \quad & w_j^T x_i + b_j \geq 1 - \varepsilon_i^j, \text{ if } y_i = j \\ & w_j^T x_i + b_j \leq -1 + \varepsilon_i^j, \text{ if } y_i \neq j \\ & \varepsilon_i^j \geq 0 \end{aligned} \quad (2-8)$$

Definition of parameters of the formula above are given below:

- The training data is in the form of (x_i, y_i)
- c : number of classes (therefore, c binary svm models are required.)
- Class j defined as positive during the training of the related class.

The class j that has the largest decision function value contains a new sample x_i (Xu, et al., 2017),

$$\bar{y}_i = \operatorname{arg max}_j w_j^T x + b_j \quad (2-9)$$

- **One-Against-One**: This algorithm generates $N \times (N - 1) / 2$ classifiers through utilizing all the binary pairwise combinations of the N classes. In order to gather these classifiers, a lot of methods were proposed in literature, such as; AND gate, Max Wins Algorithm (it finds the outcome result through voting the classes based on the result of each classifier and subsequently, selecting the class that has the highest vote) and Directed Acyclic Graph (it is utilized for gathering the results of one-against-one classifiers.) (Vural & Dy, 2004).

For class j and k , the maximum-margin hyperplane between them is $w_{jk}x_i + b_{jk} = 0$, and it can be learned through the problem as follows,

$$\begin{aligned}
& \min_{w_{jk}, b_{jk}} \quad \frac{1}{2} \|w_{jk}\|_2^2 + C \sum_{i=1}^n \varepsilon_i^{jk} \\
& \text{s. t.} \quad w_{jk}^T x_i + b_{jk} \geq 1 - \varepsilon_i^{jk}, \text{ if } y_i = j \\
& \quad \quad w_{jk}^T x_i + b_{jk} \leq -1 + \varepsilon_i^{jk}, \text{ if } y_i = k \\
& \quad \quad \varepsilon_i^j \geq 0
\end{aligned} \tag{2-10}$$

Max Wins Algorithm is utilized in testing stage, if $\text{sign}(w_{jk}x_i + b_{jk}) = 0$ shows that x_i is in class j , subsequently the vote for class j is added by one, otherwise vote for class k is added by one. Final prediction class is the class that has the largest vote (Xu, et al., 2017).

- Decision Tree:

A directed tree, which has a root node that does not include incoming edges and all other nodes that have at least one incoming edge (decision nodes), is called a decision tree (Dai & Ji, 2014).

Another explanation of decision tree nodes is given in the following:

- **Root Node:** It has neither incoming nor outgoing edges.
- **Internal Nodes:** It has only one incoming edge and one or more than one outgoing edge.
- **Leaf or Terminal Node:** It has one incoming edge but it does not have any outgoing edges (Tan, Steinbach, Karpatne, & Kumar, 2005).

Each internal node divides an instance space to minimum two parts at the training level in order to improving the performance of the classifier. Afterwards, a decision rule is taken a shape by every way from the root node to the leaf node for deciding the class to which a new instance belongs (Dai & Ji, 2014).

There is an example decision tree in the following:

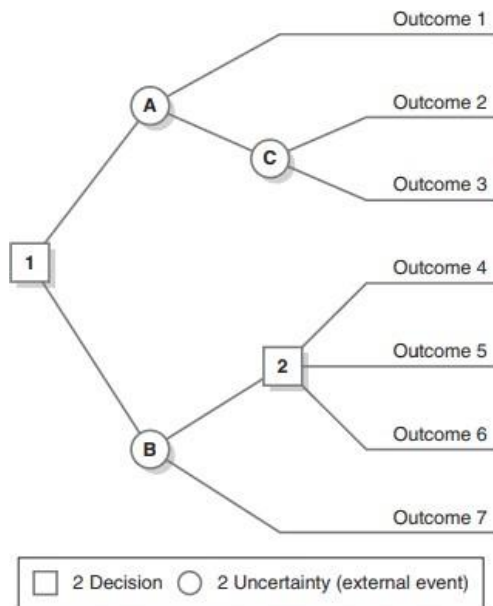


Figure 2.5: Example Decision Tree (Ishwaran & Rao, 2009)

As can be seen in Figure 2.5., Node 1 is a root node and Node A, Node B and Node C are internal nodes.

Algorithmic Framework of Decision Tree

In the literature, the results show that utilizing optimal decision tree algorithms is adequate only for small problems. As a result, heuristic techniques are needed in order to solve more complicated problems. These techniques can be split into two groups: top-down and bottom-up approach.

There are a lot of top-down decision tree inducers for example ID3, C4.5 and CART. Some of the inducers have two cognitive levels: Growing and pruning. (C4.5 and CART). Other ones have only the growing level (Rokach & Maimon, Decision Trees, 2005).

Decision Tree Algorithms:

- Hunt’s Algorithm: This algorithm creates a decision tree through utilizing the top-down approach or dividing and conquering move toward. The sample (raw data) includes a modifier group that utilizes an attribute test in order to divide the data into smaller subsets. Hunt’s algorithm provides ideal division for every group based on some threshold value.

- C4.5 Algorithm: It is developed according to Hunt's and ID3 algorithm and pruning figures in it through subrogating the internal node with a leaf node in order to decrease the error rate that happens because of noise and too many details in the training dataset.
- ID3 Algorithm: Iterative Dichotomiser 3 is simple decision tree algorithm that Quinlan Ross presented in 1986. It is developed based on Hunt's algorithm and serially utilized. The basic principle of ID3 Algorithm is that it creates a decision tree through utilizing a top-down, detailed search by the given sets to test each attribute at every tree node.
- SPRINT Algorithm: SPRINT Algorithm represents a scalable parallelizable induction of the decision tree algorithm. The difference between Hunt's and SPRINT Algorithm is that it segments the training dataset in a repetitive way by utilizing breadth-first detailed method until each segment belongs to the same leaf node or class. It can be used in both serial and parallel pattern in order to locate data effectively and balance the load.
- CART Algorithm: CART Algorithm represents both classification and regression trees. It generates the classification tree according to binary segmenting of the attributes. CART Algorithm is based on Hunt's Algorithm and it is serially utilized. However, the difference of CART Algorithm from other Hunt's based algorithms is that CART Algorithm is implemented for regression analysis through regression trees. The property of the regression analysis is utilized in estimation of dependent variable as presented a set of predictor variables over a presented period of time.
- SLIQ Algorithm: SLIQ Algorithm is a quick scalable decision tree algorithm that can be utilized in serial and parallel pattern. It was not generated according to Hunt's Algorithm for decision tree classification. It segments a training data set iteratively using breadth-first detailed strategy that is integrated with the pre-levelling method in the tree generation level (Priyam, 2013).

2.3.2. Data Mining:

The continuum of finding attracted patterns and knowledge from a large amount of data is called data mining. The data resources can consist of databases, data warehouses, the Web, other information repositories or data that comes to the system directly. Data mining can be implemented for any type of data if the data are suitable for aimed application (Han, 2012).

Data Mining operations is given in Table 2.3.:

Table 2.3.: Data Mining Operations

Method	Explanation	Application Areas	Reference
Classification	The problem of data mining is defining the class label for the test examples, which are unlabelled, through a current set of training data points along with related training labels. Classification algorithm includes two stages: Training Stage (a model is generated from the training examples.) and Testing Stage (the model is utilized in order to allocate a label to an unlabelled test example.) The output of the classification algorithm is introduced for a test example in two ways: Discrete Label (the label is retrieved for the test example.) and Numerical Score (a numerical score is retrieved for each class label and test example combination.)	Classification method has so many application areas because various data mining methods use the classification technique in order to analyse a database. For example: Decision Tree Method can be utilized in many fields such as remote sensing image, disaster weather forecasting and correlation analysis of environmental variables. On the other hand, Support Vector Machine can be applied to recognize handwriting, identify fraudulent credit cards, distinguish a speaker, as well as detect a face.	(Aggarwal, 2015) (Geetha, 2015) (HUANG, et al., 2017)
Clustering	Data can be analysed deeply by the clustering method through sectioning the objects into clusters of objects, therefore objects of each cluster are more congener in themselves than the objects in other clusters. External information, such as class labels, is not utilized in clustering; as a result of this, cluster analysis is counted as unsupervised learning in some conventional areas for example machine learning and pattern recognition. Generally, clustering analysis has two aim: understanding (applying the clustering analysis in order to discover conceptually meaningful groups of objects, which have common features, automatically.) and utility (attempting to separate the prototypes or symbolic objects from the remaining objects in the same cluster).	Clustering method in data mining can be utilized in many different fields. For example: Medicine (it can be used for finding differences between different types of tissues and blood (PET Scans).), Business and marketing, World wide web (in social networks, clustering can be utilized for identifying communities between large groups of people.), Computer science (Clustering can be beneficial for image segmentation), Social science (It may be applied in crime analysis for detecting fields where the crimes occurred.), Educational data mining and Climatology.	(Wu, 2012) (Neha & Vidyavathi, 2015)

Regression	<p>In order to forecast the future value of a variable according to the linear relationship that is between the related variable and other variables, linear regression, which is a common data mining method, can be used. Mainly, it is supposed that there is a straight line that approximates the data set and is based on the predictions.</p> <p>In linear regression, there is only one detached variable and the formula presents the relationship between the detached variable and other variables and states the straight line at the same time: $y = a+bx$ (y is dependent variable, x is detached variable, a and b are the line's coefficients.)</p>	<p>Profit, sales, mortgage rates, house values, square footage, temperature, or distance can be forecasted by the regression technique in data mining. For instance, a regression model can forecast the value of a data warehouse according to web-marketing.</p>	<p>(Olaniyi, 2011) (SINGH, 2012)</p>
Pattern Mining	<p>Pattern mining includes finding absorbing, effective and unforeseeable patterns in a data set. Pattern mining is a useful method because of its ability to reveal patterns that can be unseen in large data sets. These patterns can be interpreted by humans and beneficial for comprehending the data and for decision-making. Pattern mining became very popular because of its application in so many areas. However, various pattern mining methods, for example frequent itemset mining and association rule mining, target analysing the data that does not have the sequential ordering of events. Therefore, in order to analyse sequential events of a data, sequential pattern mining is presented</p>	<p>Pattern mining technique searches for repeating patterns through mining the sequence database in order to be utilized in future by end users or management for finding relationships between the different items or events in their data for purposes such as marketing campaigns, business reorganization, prediction and planning.</p>	<p>(Fournier-Viger, 2017) (Wright, 2015) (MABROUK, 2010)</p>

2.3.3. Case Based Reasoning:

The continuum of withdrawing previous cases, which is similar to a new problem, obviating the problem through adjusting solutions of the previous similar problems and recording the new solution in order to utilize in the future is called Case Based Reasoning (Kim & Kim, 2010).

Case Based Reasoning has four main levels: Case Retrieval, Case Reuse, Case Revision and Case Retain. Definitions of the case based reasoning levels are given in following:

- **Case Retrieval:** The objective of this level is identification of similar cases and to facilitate the decision-making process for current cases through creating a case database. Case Retrieval process needs multi-directional information with a modular structure for enhancing retrieval productivity and making case storage easier.
- **Case Reuse:** Case reuse (case adaptation) aims to plan the solution based on former cases for the related problem. The main duty of this process is the calculation of the

similarity between current cases and past cases through following these steps: deciding the weight of case attributes, calculating the similarity of case attributes, and calculating the global similarity of compared cases.

- **Case Revision:** Case revision is before testing the new solution in the reality, including planned former solutions for the current situation. If the problem characteristic of a similar case is convenient with the case that is about to be analysed, at that time the solution characteristics of the similar case can be suitable for the direct decision-making process. On the other hand, if the similar case is not adequate for the new background or scenario, the retrieved case needs to be rearranged.
- **Case Retention:** After the solution has been effectively fitted to the current problem, it is essential to stock the experience that is gained at the end of the previous step as a new case in the database. Case retention is a flexible continuum of attaching and ejecting cases, with the purpose of enhancing the efficiency of the CBR model.

There is graphical summary of the steps of Case Based Reasoning in the following:

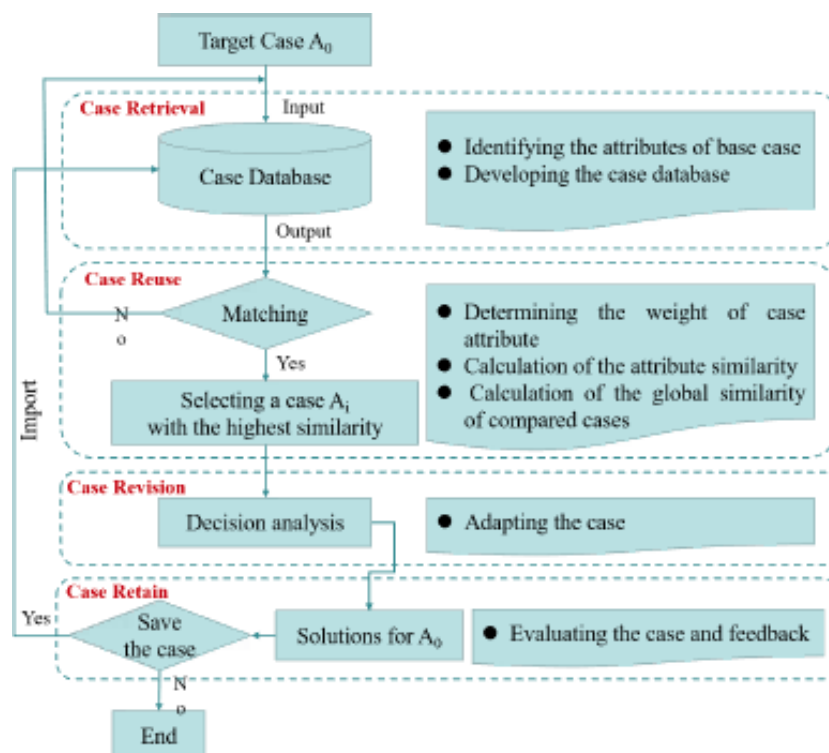


Figure 2.6.: Summary of CBR Steps (Su, 2019)

There are some significant features of the case based reasoning model:

- It is critical to reference to old similar cases in order to handle the complications of the

current situation.

- Comprehending and paraphrasing a problem is a critical part of the case based reasoning process as well as a precondition and corequisite of the problem solving approaches.
- It is generally essential to adjust the old case in order to fit the new case because there is no old case that is exactly same as the new one.
- Learning is the natural conclusion of the reasoning. If a new method comes into existence during solving a complicated problem and the method increases the efficiency of the entire reasoning process, the new method is learned in order to handle the related new group of cases.
- The last important part of the case based reasoning process is evaluating feedback by analysing follow-up procedures and explanatory reasoning (Kolodner, 2014).

There are some works that includes utilization of Case Based Reasoning (CBR) Method in literature. For example, one of the paper analysed CBR Method are which used in different processes of injection molding, for instance continuum design, continuum parameters, fault detection, and improvement of quality control. Moreover, the paper researched and conferred usage trends of CBR Method in different steps of injection molding process as well as the most important problem of the implementation of CBR to injection molding. Subsequently, the analysis is completed through planning on some open research fields and future forecasts (Khosravani & Nasiri, 2019). Another paper covers a novel theoretical approach in order to design a CBR system for facilitating data-based complexity management systematically in Industry 4.0. in production systems. Besides, the suitability of the CBR system is presented for ensuring appropriate knowledge in order to decrease and conduct complexity in business practice. However, the research was done according to a single case study so that the approach may not be possible to generalized. Therefore, future work of the paper is to test the approach and improve the research influence (Schott, 2020). Last example is the paper that aims to offer a structure in order to discover, isolate and propose suitable maintenance jobs for large-scale complex machinery in a simplistic and organised way through using the earlier fault histories accessible with the organization in combination with case-based reasoning (CBR) method. Besides, in the paper, it is showed that the offered framework can be applied through utilizing modern graphical user interface tools for example Microsoft Visual Basic or similar programs (Boral, 2019).

CHAPTER 3: METHODOLOGY OVERVIEW

This chapter presents a summary of the methodology of the project that will be presented in Chapter 4, Chapter 5, Chapter 6 and Chapter 7. The explanation of the content of each chapter of the methodology is given in following:

3.1. Generic Cost Model Development: The generic cost model for calculating unit manufacturing cost will be explained part by part in Chapter 4. In summary, the aim of this chapter is creating a generic cost model that covers direct fixed and variable cost parameters as well as fixed and variable overhead cost parameters of a facility. Therefore, the generic cost model consists of three parts: Direct Process Level Cost that includes variable direct process cost parameters, Process Overhead Level Cost that covers fixed and variable process overhead cost parameters and Company Level Cost that has fixed and variable company overhead cost parameters.

3.2. Sensor Based Cost Modelling and Relevant Data Acquisition: Several experiments that will be done for obtaining data for calculating electrical energy cost of a manufacturing process in Chapter 5. The experiments will be done in Company 1, Company 2 and Liverpool John Moores University through assembling the power meter on laser sintering, cutting and milling machines in order to obtain active power data for calculating electrical energy cost as well as analysing active power graph against time.

The electrical energy cost formula, which will be used for calculating electrical energy cost in this chapter, will be added to the total direct process cost of the generic cost model of the project, which will be explained in Chapter 4, in order to calculate unit manufacturing cost.

3.3. Assessment of The Manufacturing Process Efficiency: The aim of this chapter that is Chapter 6 is to classify the power values of the milling machine, which were obtained through three weeks experiment with a power meter in the Liverpool John Moores University Laboratory, as “idle” (the machine is idle and it is not producing a product) and “working” (the machine is producing a product.). Subsequently, total experiment time and total working time of the milling machine is aimed to be calculated and total utilization rate of the machine, which will be used in the unit manufacturing cost model, is aimed to be obtained based on these time calculations. Therefore, in order to achieve these goals, binary support vector machine (SVM) and decision tree algorithms is decided to be applied to the power values of the machine in Matlab Software. However, initially, it is decided to apply these algorithms to Company 1 and Company 2 experiment results for obtaining a Matlab code that can

estimate the status of the power values successfully. The utilization rate of the machine can be a strong indicator about the efficiency of the manufacturing process and it can also be a useful parameter for calculating manufacturing process overhead cost accurately. Therefore, it was decided to use the utilization rate of the machine parameter in the manufacturing cost calculation model.

3.4. Framework of The Knowledge Support System: The knowledge support system database of the project was explained in a detailed way in Chapter 7. In summary, the database has three parts: Part 1 is new and existing product types information that has unit manufacturing cost estimation of the new product and process and product information of existing product types. Three types of product (Product 1, Product 2 and Product 3) and their process and previous product information are given in the database as an example. Part 2 of the database is the manufacturing cost calculation sub-database. In this sub-database, the unit manufacturing cost of the product is calculated through the generic cost model that is presented in this project. Lastly, Part 3 of the database is the case based reasoning system sub-database.

The knowledge support system database is designed to facilitate optimization of manufacturing processes and calculation of the unit manufacturing cost of a product in small or medium size enterprises (SMEs).

CHAPTER 4: GENERIC COST MODEL DEVELOPMENT

This chapter presents the development of a generic sensor-based cost model of the knowledge support system database that can be utilized for calculating unit manufacturing cost of a product. The generation process of the generic cost model includes three steps: 1) Determining cost model parameters and grouping them under four groups, 2) Presenting sub-formulas of some cost model parameters and combining the cost parameters and 3) Obtaining a generic cost model for calculating unit manufacturing cost. The steps are explained in detail below:

4.1. Step 1: Generic Cost Model Parameters:

The generic cost model should include direct cost parameters as well as overhead cost parameters in order to give a reliable result to the user of the knowledge support system database. Therefore, the parameters were determined through having meetings with Company 1 and Company 2. The parameters of the cost model and explanation of some of the parameters are presented in Table 4.1.

Table 4.1.: Cost Parameters of Generic Cost Model

Cost Parameters	Explanations
Computer Software Cost	The cost of the software that can be used during the process.
Operator Cost per Process	The cost of the operator that directly operates the machine of manufacturing process.
Raw Material Cost	The cost of the material of the product.
Electricity Cost (used only for manufacturing process)	The cost of the electricity consumption of the machine of the manufacturing process. It is calculated according to the active power of the machine that can be obtained from the power meter and cycle time of the process.
Storage Cost	The storage cost of the product.
Packaging Cost	The packaging cost of the product after the manufacturing process is completed.
Administrative Personnel Cost	The cost of the personnel that is in the management level of the company.
Standing Electricity Charge	The cost of the electricity that is constant even there is no electricity consumption in the facility.
Depreciation of the machine	The cost of the depreciation of the machine.
Pension Cost and other post-retirement benefits	The amount of the payments which the company makes to the people that are retired from the company.
Security Charges	The cost of the security measures of the company.
Staff Welfare	The cost of staff welfare.
Trade Subscriptions	The monthly or annual payments of organizations or institutions to which affiliated.
Courier Service	The cost of the courier service that is used in the facility.
Medical Insurance	The cost of the medical insurance of staff of the company.
Accountancy	The amount of the payment that is given to the accountant.
Legal and Professional Fees	The amount of the payment that is given to the lawyer.

Estate Taxes	The cost of the estate taxes that are paid by the company.
Rent of the Facility	The cost of the rent of the facility that is paid by the company.
Maintenance Cost	The cost of the maintenance of the machines in the facility.
Insurance	The cost of the insurance payments of the equipment of the facility.
Post and Stationery	The cost of post and stationery that can be used during or at the end of the manufacturing process as well as in the management level.
Telephone	The cost of the telephone usage.
Fluid Cost per Process	The cost of the fluid consumption during the manufacturing process.
Gas Cost per Process	The cost of the gas consumption during the manufacturing process.
Advertising	The amount of the payment that is allocated for advertising.
Staff Training	The amount of the payment that is allocated for staff training for a specific manufacturing process.
Light and Heat	The cost of light and heat consumption of the facility.

After identifying the parameters of the generic cost model, according to the meetings with the industrial partners, the cost parameters that are considered by Company 1 and Company 2 are gathered in four groups:

- **Fixed Cost:** This group includes cost parameters of which the value does not change with the amount of a product that it is planned to produce.
- **Variable Cost:** In this group, the alteration of the value of the cost parameters is dependent on the amount of a product that it is planned to produce.
- **Fixed Overhead Cost:** Fixed cost parameters that do not affect the manufacturing process.
- **Variable Overhead Cost:** Variable cost parameters that do not affect the manufacturing process.

The cost parameters that are considered by Company 1 are presented in Table 4.2.:

Table 4.2.: The Cost Parameters of Company 1

Fixed Cost	Variable Cost	Fixed Overhead Cost	Variable Overhead Cost
Computer Software Cost	Raw Material Cost	Pension Cost and other post-retirement benefits	Maintenance Cost
Depreciation of the machine	Electricity Cost (used only for manufacturing process)	Security Charges	Telephone
	Storage Cost	Staff Welfare	Light and Heat
	Packaging Cost	Trade Subscriptions	Courier Service

	Fluid Cost per Process	Medical Insurance	
	Gas Cost per Process	Accountancy	
	Operator Cost per Process	Legal and Professional Fees	
		Estate Taxes	
		Rent of the Facility	
		Insurance	
		Standing Electricity Charge	
		Administrative Personnel Cost	

On the other hand, the cost parameters of Company 2 are given in the following under these four groups:

Table 4.3.: The Cost Parameters of Company 2

Fixed Cost	Variable Cost	Fixed Overhead Cost	Variable Overhead Cost
Computer Software Cost	Raw Material Cost	Medical Insurance	Maintenance Cost
Depreciation of the machine	Electricity Cost (used only for manufacturing process)	Accountancy	Telephone
	Packaging Cost	Legal and Professional Fees	Light and Heat
	Fluid Cost per Process	Rent of the Facility	
	Gas Cost per Process	Insurance	
	Operator Cost per Process	Standing Electricity Charge	
		Administrative Personnel Cost	

After investigating and specifying the cost parameters of Company 1 and Company 2 as a result of the meetings and presenting these cost parameters in Table 4.2. and Table 4.3., the cost parameters of these two companies are combined together and they are gathered under two main groups (Product Level Cost and Company Level Cost) and three sub-groups which are Direct Process Cost, Company Overhead Cost and Process Overhead Cost.

- **Direct Process Cost:** This group includes fixed and variable costs of the manufacturing process that are directly related to the process.
- **Process Overhead Cost:** Process Overhead Cost covers fixed and variable overhead costs of the manufacturing process that are indirectly related to the manufacturing process. Total Direct Process and Process Overhead Costs are called Product Level Cost.

- Company Overhead Cost (Company Level Cost): This group has fixed and variable overhead costs of the company that are not related to the manufacturing process. Company Overhead Cost is under Company Level Costs.

Table 4.4. shows the cost parameters that are gathered under these five groups:

Table 4.4.: Cost Parameters of Generic Cost Model that are Gathered Under Five Groups

PRODUCT LEVEL COST			COMPANY LEVEL COSTS	
<u>Variable Direct Process Cost</u>	<u>Fixed Process Overhead Cost</u>	<u>Variable Process Overhead Cost</u>	<u>Fixed Company Overhead Cost</u>	<u>Variable Company Overhead Cost</u>
Raw Material Cost	Computer Software Cost	Staff Training	Pension Cost and other post-retirement benefits	Maintenance Cost
Electricity Cost (used only for manufacturing process)	Depreciation of the machine	Post and Stationery	Security Charges	Telephone
Storage Cost			Staff Welfare	Light and Heat
Packaging Cost			Trade Subscriptions	Courier Service
Fluid Cost per Process			Medical Insurance	
Gas Cost per Process			Accountancy	
Operator Cost per Process			Legal and Professional Fees	
			Estate Taxes	
			Rent of the Facility	
			Insurance	
			Standing Electricity Charge	
			Administrative Personnel Cost	

4.2. Step 2: Sub-Formulas of Generic Cost Model:

After determining the parameters of the generic cost model and grouping them under four categories, the sub-formulas of some of the parameters were decided. The generic cost model is time-based; therefore, the sub-formula of the parameters includes time values that can be cycle time of the process or time of the process of the parameter. The parameters and their sub-formulas are given in the following:

- Electricity Cost (only used for manufacturing process) (£):

$$EC = TAE \text{ (kW)} * EP \text{ (£/kWh)} * CT \text{ (h)} \quad (4-1)$$

Where *EC* (£): Electricity Cost (used only for manufacturing process)

TAE (*Total Active Energy* (kW)): Total active energy that the machine consumes. This value can be obtained from the sensor of the machine that is a power meter.

EP (*Electricity Price* (£/kWh)): This parameter indicates electricity price of the company per hour.

CT (*Cycle Time* (h)): It is the cycle time of the manufacturing process of the product from which manufacturing cost will be calculated.

- Fluid Cost per Process (£):

$$FC = TFC \text{ (lt)} * UFP \text{ (£/lt h)} * CT \text{ (h)} \quad (4-2)$$

Where *FC* (£): Fluid Cost per Process

TFC (*Total Fluid Consumption* (lt)): The amount of total fluid consumption during the manufacturing process of the product.

UFP (*Unit Fluid Price* (£/lth)): This parameter shows fluid consumption cost of the company per hour.

- Operator Cost per Process (£):

$$OC = HSO \text{ (£/h)} * NOMP * CT \text{ (h)} \quad (4-3)$$

Where *OC* (£): Operator Cost per Process

HSO (*Hourly Salary of the Operators* (£/h)): The salary of an operator of a machine of a manufacturing process per working hour.

NOMP (*Number of Operators for a Manufacturing Process*): The number of operators that are required for a manufacturing process.

- Administrative Personnel Cost (£):

$$\text{Administrative Personnel Cost} = HSAP \text{ (£/h)} * NAPC * TWH \text{ (h)} \quad (4-4)$$

Where *HSAP* (£/h): Hourly Salary of the Administrative Personnel

NAPC: Number of Administrative Personnel in a Company

TWH (h): Total working hours for a product

- Storage Cost (£):

$$SC = TSAR (\pounds) / (NUSP + NUST). \quad (4-5)$$

Where SC (£): Storage Cost

$TSAR$ (£): Total Storage Area Rent

$NUSP$: Total number of units shipped during previous 12 months

$NUST$: Total number of units that is stored currently

4.3. Obtaining the Generic Cost Model for Calculation of Unit Manufacturing Cost of a Product:

After determining the parameters of the generic cost model and creating the sub-formulas of some cost parameters, the last step will be fulfilled which is obtaining a generic cost model for calculating unit manufacturing cost of a product by combining these four groups that are given in Table 4.2.

The generic cost model has three main levels when these four groups that are presented in Table 4.4 are combined: Direct Process Level Cost, Process Overhead Level Cost and Company Level Cost.

4.3.1. Direct Process Level Cost (£):

In this level, the parameters of variable direct process costs that are given in Table 5.2. and explained in Table 5.1. are summed up. The formula of Direct Process Level Cost is presented in Eq (3-7):

$$DPLC = RMC + EC + SC + PC + OC + FC + GC \quad (4-6)$$

Where $DPLC$ (£): Direct Process Level Cost

RMC (£): Raw Material Cost

EC (£): Electricity Cost (used only for manufacturing process)

SC (£): Storage Cost

PC (£): Packaging Cost

OC (£): Operator Cost per Process

FC (£): Fluid Cost per Process

GC (£): Gas Cost per Process

4.3.2. Process Overhead Level Cost (£):

Process Overhead Level Cost contains fixed and variable overhead cost parameters.

Differently from Direct Process Level Cost and Company Level Cost, utilization rate of the machine, which is calculated through Matlab Software according to active power values of the machine that are obtained from the power meter or estimated based on utilization rate of the machine of previous similar products by utilizing Case Based Reasoning Method, will be considered in this level. The formula of Process Overhead Level Cost is given in Eq (4-7):

$$POLC = CSC + DOM + ST + PAS \quad (4-7)$$

Where *POLC* (£): Process Overhead Level Cost

CSC (£): Computer Software Cost

DOM (£): Depreciation of the Machine

ST (£): Staff Training

PAS (£): Post and Stationery

When utilization rate of the machine is considered in Process Overhead Level Cost, it is seen that the utilization rate of the machine and Process Overhead Level Cost are inversely proportional. It means that when the utilization rate of the machine increases in a manufacturing process, total process overhead cost decreases because the idle time of the machine drops and the efficiency of the machine rises. Therefore, the formula of Process Overhead Level Cost that includes the utilization rate of the machine parameter is shown in Eq (4-8):

$$POLCUR = POLC * (1/URM) \quad (4-8)$$

Where *POLCUR* (£): Process Overhead Level Cost After Considered Utilization Rate of the Machine

POLC (£): Process Overhead Level Cost

URM (£): Utilization Rate of the Machine

4.3.3. Company Level Cost (£): Company Level Cost consists of the sum of the value of fixed and variable company overhead cost parameters. The formula of Company Level Cost is given as Eq (4-9):

$$CLC = PCORB + SC + SW + TS + CS + MI + A + LPF + ET + ROF + I + SEC + APC + MC + T + LH + CS \quad (4-9)$$

Where *CLC* (£): Company Level Cost

PCORB (£): Pension Cost and Other Post-Retirement Benefits

SC (£): Security Charges

SW (£): Staff Welfare

TS (£): Trade Subscriptions

CS (£): Courier Service

MI (£): Medical Insurance

A (£): Accountancy

LPF (£): Legal and Professional Fees

ET (£): Estate Taxes

ROF (£): Rent of the Facility

I (£): Insurance

SEC (£): Standing Electricity Charge

APC (£): Administrative Personnel Cost

MC (£): Maintenance Cost

T (£): Telephone

LH (£): Light and Heat

CS (£): Courier Service

After Direct Process Level Cost, Process Overhead Level Cost and Company Level Cost were explained, the generic cost model for calculating unit manufacturing cost of a product is created and given as Eq (4-10):

$$TUMCP = DPLC + POLCUR + CLC \quad (4-10)$$

Where *TUMCP* (£): Total Unit Manufacturing Cost of a Product

DPLC (£): Direct Process Level Cost

POLCUR (£): Process Overhead Level Cost After Considered Utilization Rate of the Machine

CLC (£): Company Level Cost

CHAPTER 5: SENSOR BASED COST MODELLING AND RELEVANT DATA ACQUISITION

5.1. Introduction:

Obtaining reliable data from a sensor on a machine during a manufacturing process can be a very important step during a cost model generation. Therefore, in order to achieve reliable results from the cost model during testing the model, three experiments were done in three different establishments by assembling a sensor on three different machines. A power meter was utilized as a sensor.

Before the explanation of power meter experiments of these three establishments, the industrial partners of these project that are Company 1 and Company 2 will be introduced in the following:

- Company 1:

Company 1 is a small size company that has specialized in designing and producing industrial filters for 30 years. In order to solve filtration challenges, they have contributed to the market through producing a large variety of industrial filters in any shape or design such as basket strainers, duplex filters, vacuum filters, cone filters, cylinder filters and screen filters in perforated plate, woven wire mesh and expanded metal. Beside the filters, they produce Dixon Ring and have developed a machine for high-speed manufacturing. More recently, they have started to utilize additive manufacturing technology for the industrial filters manufacturing process in order to decrease the cycle time and increase the product quality. In their facility, besides 3D Printing, the dominant manufacturing processes are pressing, welding and machining such as lathing, polishing and drilling. Hence, there are a large variety of machines are located in Company 1. These are seam, spot, plasma and laser welders, milling machines, punch and hydraulic presses, hydraulic guillotines, milling machines, pillar drills, band saws, lathes, manual presses, manual rollers, pneumatic sheet rollers and selective laser sintering machines for additive manufacturing. When the product is obtained after manufacturing processes, they provide the quality control according to customer requirements. All of the products are visually examined and measured by the manager of the department and some of them are put into pressure tests as well as Positive Material Identification (PMI) tests. Lastly, the future plan of Company 1 includes the improvement and distribution of Dixon Rings, Filter Housings and Additive Manufactured parts. Additionally, they aim to expand within current market sectors through an increase in global brand awareness.

- Company 2:

Company 2 is a small size company that can produce up to 5,000 tonnes of strip mill products, which are panel flat de-coiled sheet, slit coil to tolerances of $\pm 0.10\text{mm}$, precision cut blanks to tolerance of $\pm 0.25\text{mm}$ and supply of parent coils, original or rewound to fit customer requirements, each month. The company supplies products that they manufacture to multifarious industries such as automotive, domestic appliances, sheet metal and laser cutting, storage, ducting and ventilation, construction, street furniture and green energy. In order to fulfil the requirements of this wide range of industries, Company 2 utilizes different equipment during the manufacturing process of strip mill products. These pieces of equipment are de-coiling lines (capable of de-coiling in excess of 6 metres), multi strand blanking line, Amada ATF Auto shear, slitting line and guillotines. Their future plan is to add a laser cutting process in their facility in order to extend their customer profile.

The first and second experiments were done by assembling the power meter on a laser sintering machine of Company 1 and cutting machine of Company 2. Subsequently, the last experiment was done through assembling the power meter on the milling machine in the laboratory of Liverpool John Moores University. The explanation of some terms, which will be used during explaining the experiments, are given in the following:

- a) Active (Real) Power: It is the actual power that is used or spread essentially in a circuit.
- b) Reactive Power: It is a “phantom power” that is not expanded in a circuit. However, it misguides the operator that it is spread in the circuit because it is known that reactive loads, for instance, inductors and capacitors, do not spread any power. However, they decrease the voltage and draw current.
- c) Apparent Power: It is the combination of active power and reactive power.

5.2. Experiment with Laser Sintering Machine in Company 1

The aim of the experiment of Company 1 is to observe the laser sintering process, measuring the cycle time of the process and calculating the electrical energy consumption of the laser sintering machine during the time of the experiment in the facility of Company 1. As a result of this, electrical energy cost of the machine, which is one of the parameters of the cost model, can be calculated.

5.2.1. Calculation of Electrical Energy Cost of the Laser Sintering Process

The power meter recorded all data that are the value of voltage and current of the laser sintering machine in three phases. Average interval of each unit of data is one second and in total, 10,062 data were gathered in 6 days 23 hours. In consequence of the gathered data, minimum,

maximum and average value of active, reactive and apparent power of the machine that are calculated by the power meter are given in Table 5.1:

Table 5.1: Minimum, Maximum and Average Value of Power Types of Laser Sintering Machine

Power Type	Minimum	Maximum	Average
Active Power (kW)	-0.000	2.085	0.851
Reactive Power (kvar)	0.000	7.257	1.020
Apparent Power (kVA)	0.000	7.296	1.328

As can be seen in Table 4.1., the average real power that was used by the laser sintering machine during the processes is 0.851 kW. However, the utilized reactive power, which is the stored power in the machine before utilization, is 1.020 kvar. Therefore, the average value of the combination of active and reactive power is 1.328 kVA.

Time based calculation of the total electrical energy cost formula that is Eq (4-1) is given in the following:

$$EC = TAE \text{ (kW)} * EP \text{ (£/kWh)} * CT \text{ (h)}. \quad (4-1)$$

$$Total \ Active \ Energy \ (TAE) = 0.851 \text{ kW}$$

$$Total \ Cycle \ Time \ (CT) = 167.97 \text{ h}$$

$$Electricity \ Price \ (EP) = 0.127 \text{ £/kWh}$$

Hence, according to the experiment results, which was presented by the power meter after the experiment, total experiment time is 167.97 hours. When the consumption of active power is considered, total energy cost is **£18.153**.

5.2.2. Analysis of Laser Sintering Process via Active Power Graph:

In order to fulfil one of the aims of the experiment that is observing the cycle time, the analysis of the graph of Active Power against time is given in the following as Figure 5.1.:

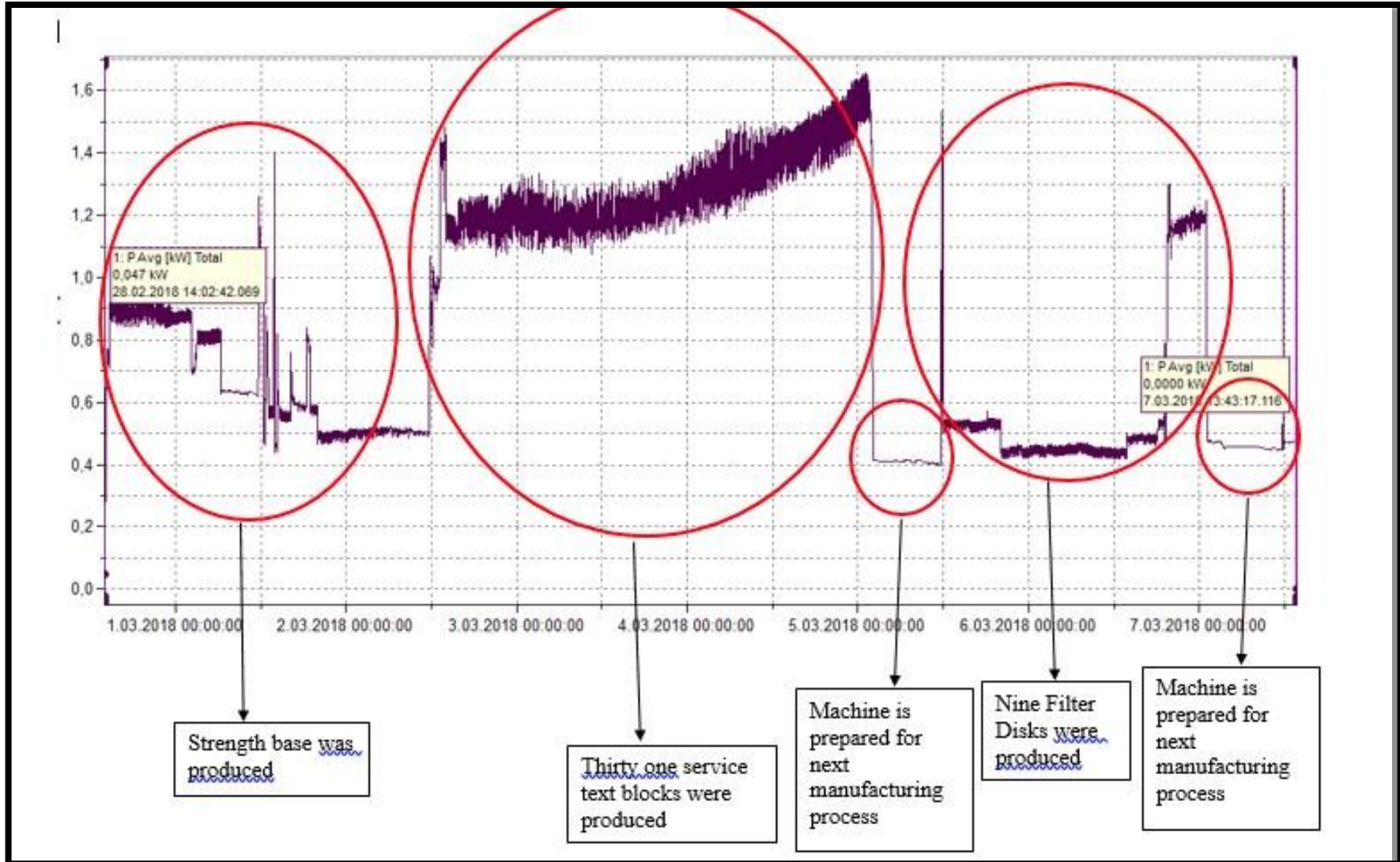


Figure 5.1.: Active Power Graph of the Experiment of Company 1

Figure 5.1. shows active power levels of the laser sintering machine during the time of the experiment. According to power levels and meetings with the operator of the laser sintering process, the type of products that were produced via the laser sintering machine were identified and shown on the graph. Therefore, the analysis of laser sintering process during the experiment time was completed.

5.3. Experiment with Cutting Machine in Company 2:

This section presents the analysis of the experiment that was done with the power meter on the cutting machine in the facility of the Company 2. It is intended that the results of the analysis will show total power consumption and electricity cost as well as the analysis of the cutting process during the experiment period.

5.3.1. Calculation of Electrical Energy Cost of the Laser Sintering Process:

Before calculating the cost of the electrical energy of the cutting process, the cutting process of Company 2 will be explained: The cutting machine has three motors inside of it. The first motor pushes the steel coils to the cutting machine, second one cuts the steels that were uncoiled by the previous motor and lastly, third motor recoils the steels that have been cut.

In order to measure the total power consumption of these three motors, the power meter recorded all data that are the value of voltage and current of the cutting machine in three phases. The average interval of each data unit is one second and in total, 10,081 units of data were gathered in 7 days. As a result of the gathered data, calculated minimum, maximum and average value of active, reactive and apparent power of the machine are given in Table 5.2.:

Table 5.2.: Minimum, Maximum and Average Value of Power Types of Cutting Machine

Power Type	Minimum	Maximum	Average
Active Power (kW)	-0.011	119.054	4.510
Reactive Power (kvar)	0.000	108.496	7.363
Apparent Power (kVA)	0.000	160.703	8.634

In Table 5.2., the average active power that was utilized by the cutting machine during the process is 4.510 kW. However, the utilized reactive power, which is the stored power in the machine before utilization, is 7.363 kvar. Therefore, the average value of the combination of active and reactive power is 8.634 kVA.

Total electrical energy cost of the cutting machine was calculated according to Eq (4-1) and it is presented below:

Total Active Energy = 4.510 kW

Total Cycle Time = 168.475 h

Electricity Price = 0.001084 £/kWh

Based on the power meter results that was given by the power meter after the experiment total experiment time is 168.475 hours. Therefore, when the value of active power is considered in the calculation of total electrical energy cost for the experiment period, total energy cost is £0.82364.

5.3.2. Analysis of Cutting Process through Active Power Graph:

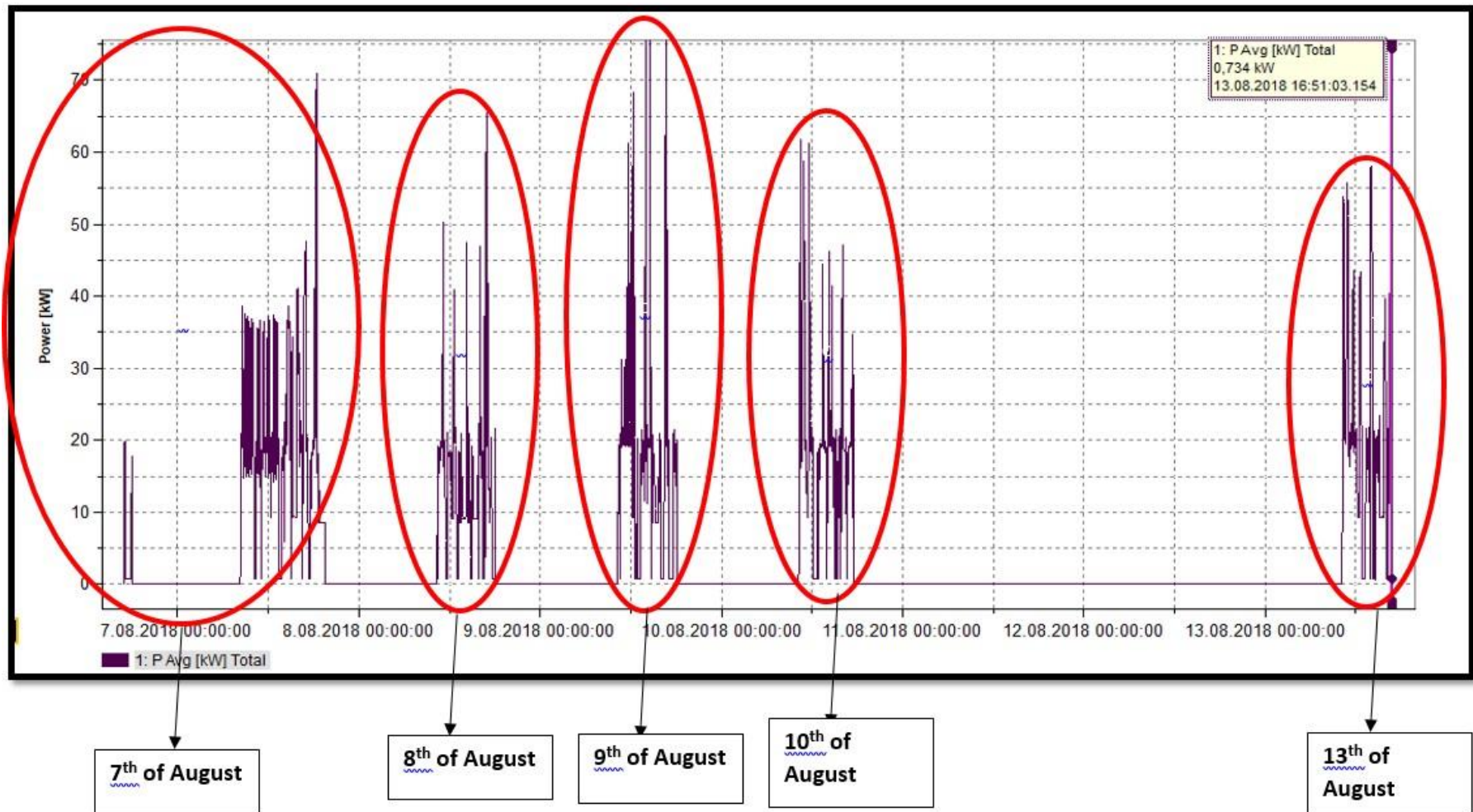


Figure 5.2.: Active Power Graph of the Experiment of Company 2

Figure 5.2. shows active power levels of cutting machine of Company 2 between 7th of August and 13th of August. Based on meetings with the cutting machine operator, patterns of active power graph were separated according to experiment days. In each day, steel coils were cut and recoiled by the cutting machine. However, the measurement of recoiled steels varies based on the requirements of the customers. Therefore, the slitters, which were used during the cutting process, differ according to the measurements of the coils that the customers required. The types of slitters that were utilized during the experiment period were cold reduced slitter, galvanised slitter and ZINTEC slitter. Slitter types and the measurements of the coils that were cut and recoiled are given in the following:

w: width of the coil

L: length of the coil

G: thickness of the coil

a) 7th of August

- Cold Reduced Slitter (w=89.90 m, L=393.00 m, G=1.95)
- Cold Reduced Slitter (w=113.60 m, L=391.00 m, G=1.95)
- Galvanised Slitter (w=941.00 m, L=0.00 m, G=0.90)
- Galvanised Slitter (w=290.00 m, L=1238.00 m, G=0.90)
- Galvanised Slitter (w=941.00 m, L=0.00 m, G=0.90)
- Galvanised Slitter (w=290.00 m, L=1223.00 m, G=0.90)
- Galvanised Slitter (w=118.00 m, L=1829.00 m, G=0.90)

b) 8th of August

- Galvanised Slitter (w=398.70 m, L=172.00 m, G=1.90)
- Galvanised Slitter (w=343.90 m, L=201.00 m, G=0.95)
- ZINTEC Slitter (w=237.00 m, L=153.00 m, G=1.42)
- Galvanised Slitter (w=330.00m, L=259.00 m, G=1.49)
- Galvanised Slitter (w=350.00m, L=264.00 m, G=1.49)

c) 9th of August

- Galvanised Slitter (w=93.00m, L=733.00 m, G=0.78)
- Galvanised Slitter (w=65.00m, L=730.00 m, G=0.78)
- Galvanised Slitter (w=41.00m, L=734.00 m, G=0.78)
- Galvanised Slitter (w=92.00m, L=1007.00 m, G=0.99)
- Galvanised Slitter (w=100.00m, L=1008.00 m, G=0.99)

d) **10th of August**

- Galvanised Slitter (w=137.00m, L=1099.00 m, G=0.55)
- Galvanised Slitter (w=137.00m, L=1099.00 m, G=0.55)
- Galvanised Slitter (w=137.00m, L=930.00 m, G=0.65)
- Galvanised Slitter (w=525.30m, L=174.00 m, G=1.42)
- Galvanised Slitter (w=78.00m, L=1385.00 m, G=0.40)
- Galvanised Slitter (w=170.30m, L=176.00 m, G=1.42)
- Galvanised Slitter (w=650.00m, L=280.00 m, G=1.42)
- Galvanised Slitter (w=525.30m, L=281.00 m, G=1.42)
- ZINTEC Slitter (w=1250.00 m, L=0.00 m, G=0.85)
- ZINTEC Slitter (w=206.50 m, L=717.00 m, G=0.85)

e) **11th – 12th of August**

Weekend Break

f) **13th of August**

- Galvanised Slitter (w=78.00m, L=1385.00 m, G=0.40)
- Galvanised Slitter (w=78.00m, L=1349.00 m, G=0.40)
- Galvanised Slitter (w=621.00m, L=0.00 m, G=0.90)
- Galvanised Slitter (w=941.00m, L=0.00 m, G=0.50)
- Galvanised Slitter (w=290.00m, L=1480.00 m, G=0.50)
- Galvanised Slitter (w=621.00m, L=0.00 m, G=0.90)
- Galvanised Slitter (w=621.00m, L=757.00 m, G=0.90)

- Galvanised Slitter (w=621.00m, L=756.00 m, G=0.90)

5.4. Experiment with Milling Machine in Liverpool John Moores University:

The last experiment that was done for the project is observing the milling machine process through assembling the power meter on the milling machine in the laboratory of Liverpool John Moores University and calculating the cost of total electricity energy of the milling process, which was utilized during the experiment period, based on the power meter data. The power meter recorded the data of the milling machine for 21 days and 30,241 units of data were gathered in total.

5.4.1. Calculation of Electrical Energy Cost of the Milling Process:

Minimum, average and maximum values of active, reactive and apparent power that were obtained from the power meter at the end of the experiment are given below:

Table 5.3.: Minimum, Maximum and Average Value of Power Types of Milling Machine

Power Type	Minimum	Maximum	Average
Active Power (kW)	-5.111	8,129	0,508
Reactive Power (kvar)	0.725	6.897	1.597
Apparent Power (kVA)	0.864	9.652	1.676

According to Table 5.3., the average value of active power is 0.508 kW and the average value of reactive and apparent value is 1.597 kvar and 1.676 kVA. Based on average value of active power, total electricity energy cost of the milling process will be calculated according to the formula that was named as Eq(4-1). The cost parameters of milling machine electrical energy cost calculation are presented in the following:

- *Total Active Power* = 0.508 kW
- *Electricity Price* = 0.127 £/kWh (industry rate)
- *Total Cycle Time* = 504 h

Considering the value of the cost parameters that were given above, total electricity energy cost of the milling machine during the experiment period is £32.516.

5.4.2. The Analysis of Milling Process through Active Power Graph:

The active power graph of the milling machine that shows active power levels against experiment time is given in Figure 5.3.:

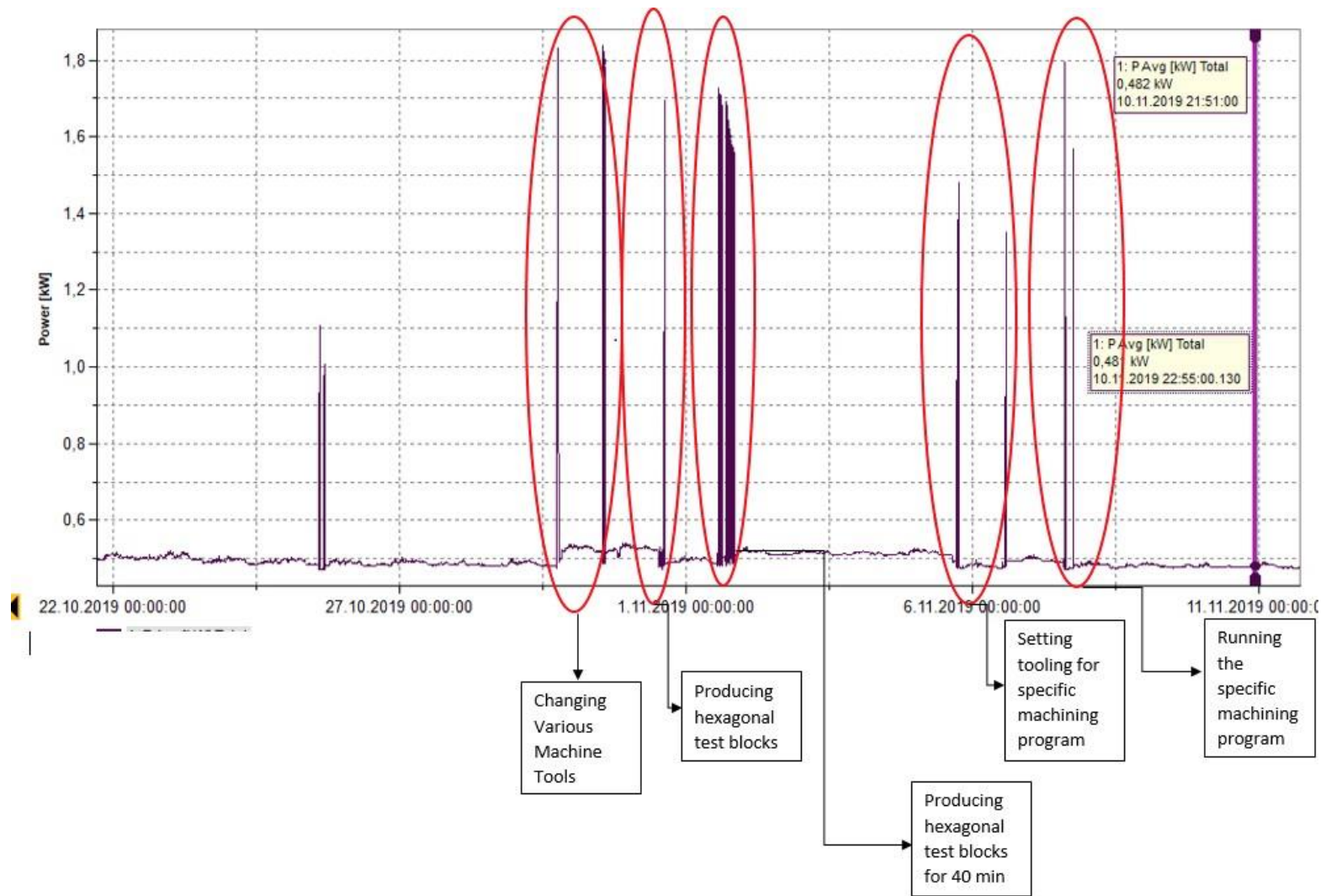


Figure 5.3.: Active Power Graph of the Experiment in the Laboratory of Liverpool John Moores University

Figure 5.3 shows active power patterns of the milling machine during the experiment period. According to the power levels and meetings with the operator, the activities of the alteration of the power levels were presented.

CHAPTER 6: ASSESSMENT OF THE MANUFACTURING PROCESS EFFICIENCY

In this chapter, in order to obtain more reliable results from the cost model and track the manufacturing process more efficiently, a manufacturing process efficiency parameter, which is the utilization rate of the machine, will be calculated and added to the cost model. The parameter is aimed to show the rate of the actual usage of the machine for manufacturing a product during an observation of the process through a sensor.

Utilization rate of the machine is planned to be calculated in 3 steps:

- Classification of the machine status, which are “Stop”, “Idle” or “Working”, according to its power level values that were obtained from the power meter.
- Calculation of times of the target status that is “Working” and total time that includes total time of all kinds of status.
- Validation of the system.

6.1. Classification of a Machine Status:

The aim of this step is classifying the status of milling machine in the Liverpool John Moores University Laboratory according to its power level values. However, before using data that are obtained from the experiment on the milling machine, two kinds of Matlab codes are written in order to apply them to Company 1 and Company 2 experiment data for gaining error-free running codes for the milling machine experiment.

6.1.1. Company 1 Laser Sintering Machine Status Classification:

As is mentioned in previous chapters, the manufacturing process of the laser sintering machine was observed and 1901 units of data were gathered in 6 days 23 hours. The 52 data of the database is given as an example in following with their machine status in order to classify the power levels through Matlab software:

Table 6.1: Power Data of Laser Sintering Machine Experiment of Company 1

Date and Time	Power1(Original)	Power 2	Power 3	Power 4	Power 5	Status
07/03/2018 13:38	18.68	19.27	19.52	18.54	15.40583492	idle
07/03/2018 13:38	19.27	19.52	18.54	15.41	16.54	idle
07/03/2018 13:38	19.52	18.54	15.41	16.54	16.17	idle
07/03/2018 13:38	18.54	15.41	16.54	16.17	17.20	idle
07/03/2018 13:38	15.41	16.54	16.17	17.20	18.44	idle
07/03/2018 13:38	16.54	16.17	17.20	18.44	17.90	idle
07/03/2018 13:38	16.17	17.20	18.44	17.90	17.71	idle
07/03/2018 13:38	17.20	18.44	17.90	17.71	17.95	idle
07/03/2018 13:38	18.44	17.90	17.71	17.95	18.14	idle
:	:	:	:	:	:	:
07/03/2018 13:38	22.90	23.92	25.38	29.64	31.09	working
07/03/2018 13:38	23.92	25.38	29.64	31.09	33.57	working
07/03/2018 13:38	25.38	29.64	31.09	33.57	31.90	working
07/03/2018 13:38	29.64	31.09	33.57	31.90	27.09	working
07/03/2018 13:38	31.09	33.57	31.90	27.09	22.16	working
07/03/2018 13:38	33.57	31.90	27.09	22.16	19.95	working
07/03/2018 13:38	31.90	27.09	22.16	19.95	20.72	working
07/03/2018 13:38	27.09	22.16	19.95	20.72	16.16	working
07/03/2018 13:38	22.16	19.95	20.72	16.16	18.11	working
07/03/2018 13:38	19.95	20.72	16.16	18.11	18.27	idle
07/03/2018 13:38	20.72	16.16	18.11	18.27	18.33	working
07/03/2018 13:38	16.16	18.11	18.27	18.33	18.48	idle
07/03/2018 13:38	18.11	18.27	18.33	18.48	14.66	idle
07/03/2018 13:38	18.27	18.33	18.48	14.66	10.13	idle
07/03/2018 13:38	18.33	18.48	14.66	10.13	1.78	idle
07/03/2018 13:38	18.48	14.66	10.13	1.78	0.56	idle
07/03/2018 13:38	14.66	10.13	1.78	0.56	0.44	idle
07/03/2018 13:38	10.13	1.78	0.56	0.44	0.51	idle
07/03/2018 13:38	1.78	0.56	0.44	0.51	0.57	idle
07/03/2018 13:38	0.56	0.44	0.51	0.57	0.58	stop
07/03/2018 13:38	0.44	0.51	0.57	0.58	0.65	stop
07/03/2018 13:38	0.51	0.57	0.58	0.65	0.54	stop
07/03/2018 13:38	0.57	0.58	0.65	0.54	0.41	stop
07/03/2018 13:38	0.58	0.65	0.54	0.41	0.57	stop
07/03/2018 13:38	0.65	0.54	0.41	0.57	0.55	stop

As can be seen in Table 6.1 that was created by taking the power levels of the machine from the power meter and assigned the machine status according to active power-time graph given in Chapter 5., there are three kinds of status of the laser sintering machine: Stop, Idle and Working.

- “Stop”: It is the status that indicates the machine is off.
- “Idle”: It is the status that shows the machine on but it is not producing a product.
- “Working”: It is the status that displays the machine is on and producing a product.

Three thresholds were determined for the three status according to the original power values of the data set. The thresholds are given in below:

P: Original Power Value

Stop: $0 < P < 1$

Idle: $1 < P < 20$

Working: $20 < P < 40$

In addition to creating the thresholds, the data shifting method was applied as a filter in order to eliminate noises from the data set of Company 1. This method includes shifting the column of original power values four times and obtaining 5 columns as input of a machine learning method.

After the preparation of the database of Company 1, support vector machine and multi-class support vector machine is applied as a machine learning method in order to classify laser sintering machine status.

6.1.1.1. Application of Binary Support Vector Machine to Company 1 Database:

First of all, support vector machine method is applied to the database part of which is given in Table 6.1. in Matlab for determining machine status. The application of Support Vector Machine method in Matlab has three steps:

- a) Training Step: In this step, a part of the database is selected as training data and a model is created in order to teach the thresholds, which are determined before, to the code. The teaching process begins with selecting inputs and output of the model. In this situation, inputs are Power 1 (Original), Power 2, Power 3, Power 4 and Power 5 columns and the output is the Status column. Subsequently, the model is generated according to

inputs and output. The model includes the Matlab function of Support Vector Machine.

- b) Optimization Step: The model is optimized through optimization Matlab function in order to obtain the most reliable result.
- c) Prediction Level: In this level, another part of the database is selected for testing. However, this time, only power columns are selected as inputs because the model is expected to predict the output. Subsequently, the generated model predicts output (status) of the testing data according to thresholds that are taught to the code before. The reliability of the model can be tested through accuracy of the predicted status.

The Matlab code and the explanation of it step by step can be seen in Appendix 1.

The explanations of Matlab code lines can be divided into three categories according to the steps of the Support Vector Machine method that are mentioned in Appendix 1:

- a) Training Level:
 - 1,283 units of data from an Excel file that includes 1,901 units of data are selected as training data and located in sheet 2 of the file that is named as “Shifting”. Subsequently, the data that is between a71 and f1310 is selected.
 - The number of training data sets is defined. The data set starts with n_0 and ends with n .
 - The input X that includes power level values as numbers from column 1 to 5 for training the model is defined.
 - The output Y that includes status as a text is defined. “Stop”, “Idle” and “Working” status is converted into “0”, “1” and “2” in further lines. The status column is changed from text to numbers because the prediction level can be done more easily with a model in which inputs and output are the same kind. Besides, converting the status column from text to numbers can ease the calculation of total time and working time of the machine. This process requires three loops because this method is beneficial to binary classifications. If more than two items need to be classified, it requires more than one “for” loop.
 - L1, L2, L3 arrays are created for each status and they are defined false in the beginning.
 - A variable that is called Y1 is defined in order to give zero to all status.
 - The first loop is created and it means that if the status is “Stop”, Y1 is zero. If it is

“Idle” or “Working”, Y1 is 1.

- A model (Mdl) that includes binary SVM function “fitrsvm” and does SVM classification through using input X and output Y1 is generated. This model only classifies 0 (“Stop”) from 1 (“Idle” and “Working”)
- A variable that is named as Y2 is defined and it is equal to Y1 in the beginning.
- The second loop is generated and it shows that if the status is “Idle”, Y2 is 1. The other status (“Stop” and “Working”) are 2.
- A model (Mdl2) that includes binary SVM function “fitrsvm” and does SVM classification through using input X and output Y2 is generated. This model only classifies 1 (“Idle”) from 2 (“Stop” and “Working”).
- Variable A is created and it can be a number between 1 and 1,900.
- The third and last loop is generated for that if $Y1 = 1$ and $Y2 = 1$, it is 1 (idle). If $Y1 = 1$ and $Y2 = 2$, it is 2 (Working). The remain status is zero (Stop).
- A model (Mdl3) that includes binary SVM function “fitrsvm” and does SVM classification through using input X and output A is created. This model classifies all three status (0, 1 and 2).

b) Optimization Level:

- The accuracy of the last model is tested with the Matlab function “resubLoss”. The third model is considered for the prediction level because only this model could convert all three status from texts to numbers.
- The hyperparameters of the third model (Mdl3) are optimized and the model is renamed as Mdl4. The fourth model is utilized for prediction level. Hyperparameter is a parameter of the model that is determined before training level starts contrary to other parameters that are set in the training level. Hyper parameters do not affect the quality of the model. However, it is critical to optimize them because they have influence over the speed and quality of the machine learning process.

c) Prediction Level:

- 200 units of data are selected from the same Excel file as a test data and the data between a2 and e68 is chosen.
- The number of the testing data set is defined. The data set starts with n_0 and ends with n.
- The input XX that includes 200 power level values from column 1 to 5 is defined.

- The model “yfit” is generated and it covers the prediction function of Matlab that is “predict”. This model does prediction by using the fourth model (Mdl4) and input XX.

Results of the Code:

Optimization Results:

Loss Function before Hyperparameter Optimization: lstd = 0.0973

Hyperparameter Optimization:

Optimization completed.

MaxObjectiveEvaluations of 30 reached.

Total function evaluations: 30

Total elapsed time: 225.7474 seconds.

Total objective function evaluation time: 200.1687

Best observed feasible point:

BoxConstraint	KernelScale	Epsilon
_____	_____	_____
930.37	981.46	0.018009

Observed objective function value = 0.093099

Estimated objective function value = 0.092333

Function evaluation time = 0.25127

Best estimated feasible point (according to models):

BoxConstraint	KernelScale	Epsilon
_____	_____	_____
360.48	187.49	0.023542

Estimated objective function value = 0.092333

Estimated function evaluation time = 0.52415

Explanation:

First of all, in the code the loss function is calculated. If the loss function is zero, it means that the model is 100% efficient and the prediction results are expected to be totally accurate. In the current situation, the loss function is 0.0973. It means that the model is very efficient but it is not 100%. It is aimed to converge zero through the optimization of the hyperparameters of the model.

The next step of the optimization is the optimization of the hyperparameters of the model. The optimization process has 30 iterations of which each includes a combination of hyperparameter values. The optimization was made through Bayesian optimization that is a process that utilizes acquisition function in order to set the following group of hyperparameter values. Total number of iterations is determined 30 as a default by Matlab Software. Moreover, according to results, total optimization time is 225.7474 seconds as well as total objective function, which is the loss function, evaluation time is 200.1687 seconds. There are two kind of values in the optimization process: estimated and observed values. The values are given below:

- a) **Box Constraint:** This parameter controls the number of disadvantages that affect margin based observations and it facilitates restraining overfitting. When the value of the box constraint is increased, it means that the SVM classifier has fewer support vectors. The estimated value of the box constraint is 360.48. On the other hand, the observed value of it is 930.37. It means that the observed optimization result shows the model has fewer support vectors than the estimated result.
- b) **Kernel Scale:** Kernel is the name of a group of machine learning algorithms that does pattern analysis and this group covers the support vector machine method as well. All elements of predictor (input) matrix of X are divided by Kernel Scale value in Matlab. Therefore, if the kernel scale value is high, it means that the positions of support vectors are enlarged more and the model become smoother. Subsequently, the Matlab software chooses a suitable scale factor through utilizing a heuristic process. This process utilizes sub-sampling, therefore predictions can differ from one call to another. Hence, the values of the input matrix of observed optimization results are smaller than the values of the estimated optimization results.

c) **Epsilon:** Epsilon is hyperparameter that can affect the number of support vectors that are utilized for creating the regression function. When the epsilon gets higher, it means that fewer support vectors are selected and this situation leads to more unreliable predictions. It can be said that when epsilon gets smaller, the better results can be expected. However, it cannot be zero because if the value of epsilon is zero, it means overfitting. In this optimization, the estimated value of epsilon is 0.023542 and the observed value of it is 0.018009. This result shows that observed solution of epsilon of the optimization of the model gives better results than estimated results.

As was said before, the objective function is referred to as loss function. Before the optimization, loss function was 0.0973 however after the optimization it dropped to 0.093099. Therefore, it can be seen that the hyperparameter optimization process of the model is successful.

Prediction Results:

Table 6.2: Prediction Results of Binary Support Vector Machine Code

Date and Time	Total Power (Original Value)	Total Power 2	Total Power 3	Total Power 4	Total Power 5	Status (Y)	Yfit (predicted by model)
28/02/2018 14:02	14.41	15.85	17.21	19.02	20	idle	0.8907
28/02/2018 14:02	15.85	17.21	19.02	20	17.67	idle	1.1006
28/02/2018 14:02	17.21	19.02	20	17.67	16.8	idle	1.2807
28/02/2018 14:02	19.02	20	17.67	16.8	17.56	idle	1.3967
28/02/2018 14:02	20	17.67	16.8	17.56	15.6	working	1.5602
28/02/2018 14:02	17.67	16.8	17.56	15.6	15.11	idle	1.3797
28/02/2018 14:02	16.8	17.56	15.6	15.11	17.75	idle	1.1757
28/02/2018 14:02	17.56	15.6	15.11	17.75	18.47	idle	1.2329
.
.
28/02/2018 14:02	14.61	13.99	12.78	13.37	13.91	idle	1.0586
28/02/2018 14:02	13.99	12.78	13.37	13.91	14.45	idle	1.0087
28/02/2018 14:02	12.78	13.37	13.91	14.45	16.58	idle	0.8249
28/02/2018 14:02	13.37	13.91	14.45	16.58	16.14	idle	0.8754
28/02/2018 14:02	13.91	14.45	16.58	16.14	16.9	idle	0.9551
28/02/2018 14:02	14.45	16.58	16.14	16.9	15.54	idle	0.9884
28/02/2018 14:02	16.58	16.14	16.9	15.54	14.35	idle	1.2811
28/02/2018 14:02	16.14	16.9	15.54	14.35	14.1	idle	1.2123
28/02/2018 14:02	16.9	15.54	14.35	14.1	15.19	idle	1.2675
28/02/2018 14:02	15.54	14.35	14.1	15.19	15.74	idle	1.115
28/02/2018 14:02	14.35	14.1	15.19	15.74	13.59	idle	1.0604

Explanation:

In this section, the predicted output results can be analysed. In Table 6.2., the given and predicted status of the laser sintering machine are compared. In the code, before the prediction process, machine status (stop, idle and working) are converted into numbers (0,1 and 2). However, the predicted values are not discrete like given status. They are continuous values. Hence, the number values of given and predicted status of machine status are given in following:

- Given Values of Machine Status:

Stop = 0

Idle = 1

Working = 2

- Predicted Values of Machine Status:

Stop ≤ 0.5

$0.5 < \text{Idle} \leq 1.5$

Working ≥ 1.5

Therefore, it can be seen that the predicted values mostly fit with given values. Because of this, it can be said that the classification of laser sintering machine status according to power values is successful.

6.1.1.2. Application of Multi-Class Support Vector Machine Algorithm to Company 1 for Laser Sintering Machine Status Classification:

In previous section, binary support vector machine method was applied for classification of machine status of the laser sintering machine. However, although the classification was successful, it is seen that the binary SVM method is not appropriate for classifying three kinds of status because three “for” loops are required for the Matlab code. Therefore, the code become more complicated. For this reason, the literature review search was carried out in order to find a more suitable machine learning algorithm for classifying three kind of machine status. When the literature review search was completed, it was observed that one of the most suitable machine learning algorithms is the Multi-Class Support Vector Machine Method.

Therefore, multi-class svm is applied to the data set that is given in Table 6.1. The Matlab code and its step by step explanation is given in Appendix 2.

Although multi-class support vector machine code is similar to binary svm code, there are some differences between related codes:

- The first difference is in converting machine status from text to numbers. During the changing process of “stop, idle and working” to “0,1 and 2”, only one “for” loop is required because multi-class svm can classify more than two items at the same time. In the loop, all three status are converted into numbers and new status are assigned to Y1 variable that is the output of the model.
- Instead of using “fitrsvm” Matlab function that is for binary svm, “fitcecoc” Matlab function is utilized, which is for multiclass models including support vector machines, for creating the model.

Results of the Code:

Optimization Results:

Loss Function before Optimization: lstd = 0.0040

Optimization Result:

Optimization completed.

MaxObjectiveEvaluations of 30 reached.

Total function evaluations: 30

Total elapsed time: 1500.5692 seconds.

Total objective function evaluation time: 1471.9501

Best observed feasible point:

Coding	BoxConstraint	KernelScale
onevsone	14.641	0.94011

Observed objective function value = 0.0048465

Estimated objective function value = 0.0024644

Function evaluation time = 2.722

Best estimated feasible point (according to models):

Coding	BoxConstraint	KernelScale
onevsone	0.040424	0.013209

Estimated objective function value = 0.0024644

Estimated function evaluation time = 15.4723

Explanation of the Optimization Result:

The first indicator of the efficiency of the model is loss function that is shown as lstd. The loss function of multi-class SVM before optimization is 0.0040. On the other hand, the loss function of binary SVM before optimization is 0.0973. Therefore, it can be seen that the model of multi-class SVM has less error about classifying machine status than the model of binary SVM.

The hyperparameter optimization of the model has two kind of results: Estimated and observed point. Total optimization time is 1500.5692 seconds and total loss function evaluation time is 1471.9501 seconds. Subsequently, hyperparameter optimization results are given in the following:

- A one-versus-one coding design is utilized for three classes. This coding design generates $N \times (N - 1) / 2$ classifiers through utilizing all the binary pairwise combinations of the N classes. In this situation, because of the three machines' status, N is three. Therefore, it provides three binary learners in the model in order to classify the machine status according to power levels.
- Estimated box constraint value is 0.040424 and the observed value of it is 14.641. It means that the model that has SVM classifier includes less support vector than its prediction value. The lower number of support vectors means higher misclassification of training algorithm that happens because the separation of the data set is not perfect.

- Another hyperparameter of the model is kernel scale. The estimated kernel scale value is 0.013209. On the other hand, the observed value of the kernel scale is 0.94011. There is a big difference between estimated and observed values. It means that the function of the model is smoother than expected because the distance between the position of its support vectors is bigger than expected.

Lastly, there is a comparison between the loss function of the optimized and non-optimized model. The loss function of the non-optimized model is 0.0040. On the other hand, the loss function of the optimized model is 0.0048465. It can be seen that the loss function of the optimized model is bigger than the loss function of the non-optimized model. Therefore, it can be said that the optimization is not successful and instead of Mdl2 (optimized model), Mdl (non-optimized model) is utilized in the prediction level.

Prediction Result:

The same testing data set with binary SVM code of Company 1 is utilized in multi-classifier SVM code for generating predicted output (yfit) by utilizing testing input (XX). In order to see the efficiency of the prediction level of the code, the comparison of the given output (Y) and predicted output (yfit) is given in Table 6.3. that includes date and time, testing data input (XX) with their given output and predicted output:

Table 6.3: Prediction Results of Multi-Class Support Vector Machine Code

Date and Time	Total Power (Original Value)	Total Power 2	Total Power 3	Total Power 4	Total Power 5	Status (Y)	Yfit (predicted by model)
28/02/2018 14:02	14.41	15.85	17.21	19.02	20.00	idle	1
28/02/2018 14:02	15.85	17.21	19.02	20.00	17.67	idle	1
28/02/2018 14:02	17.21	19.02	20.00	17.67	16.80	idle	1
28/02/2018 14:02	19.02	20.00	17.67	16.80	17.56	idle	1
28/02/2018 14:02	20.00	17.67	16.80	17.56	15.60	working	2
28/02/2018 14:02	17.66	16.79	17.55	15.59	15.11	idle	1
.
.
28/02/2018 14:02	13.91	14.45	16.58	16.14	16.90	idle	1
28/02/2018 14:02	14.45	16.58	16.14	16.90	15.54	idle	1
28/02/2018 14:02	16.58	16.14	16.90	15.54	14.35	idle	1
28/02/2018 14:02	16.14	16.90	15.54	14.35	14.10	idle	1
28/02/2018 14:02	16.90	15.54	14.35	14.10	15.19	idle	1
28/02/2018 14:02	15.54	14.35	14.10	15.19	15.74	idle	1
28/02/2018 14:02	14.35	14.10	15.19	15.74	13.59	idle	1

Explanation:

Initially, the code converted the laser sintering machine status from text (Stop, Idle and Working) to numbers (0,1 and 2) similar to binary SVM code. Subsequently, the code predicted the output according to 66 power values of the machine (testing data) through using the model that includes multi-class SVM Matlab function. The predicted values are discrete values and it can be seen that they fit the given outputs totally. Therefore, it can be said that the machine learning algorithm successfully predicted the status of the laser sintering machine according to its power values.

6.1.2. Application of Decision Tree Algorithm to Company 2 for Cutting Machine Status Classification:

In this section, a Matlab code that includes a classification tree function was created in order to classify the status of the cutting machine of Company 2 successfully. The cutting machine has three status: Stop, Idle and Working. Before explaining the Matlab code and giving its results, a part of the database of Company 2 that includes 6,300 power values of the machine, which were gathered in 7 days with power meter, and their status is shown in the following for classifying the power data in Matlab:

Table 6.4: Power Values of Cutting Machine of Company 2 and Their Status

Date and Time	Power 1 (Original Power Data)	Power 2	Power 3	Power 4	Power 5	Status
08/08/2018 09:30	30030.40	30031.23	29910.27	29936.21	30004.45	working
08/08/2018 09:31	30031.23	29910.27	29936.21	30004.45	29976.05	working
08/08/2018 09:32	29910.27	29936.21	30004.45	29976.05	29964.70	working
08/08/2018 09:33	29936.21	30004.45	29976.05	29964.70	30047.91	working
08/08/2018 09:34	30004.45	29976.05	29964.70	30047.91	30150.03	working
08/08/2018 09:35	29976.05	29964.70	30047.91	30150.03	24834.13	working
08/08/2018 09:36	29964.70	30047.91	30150.03	24834.13	11624.76	working
08/08/2018 09:37	30047.91	30150.03	24834.13	11624.76	1972.44	working
08/08/2018 09:38	30150.03	24834.13	11624.76	1972.44	1827.75	working
08/08/2018 09:39	24834.13	11624.76	1972.44	1827.75	1409.36	working
08/08/2018 09:40	11624.76	1972.44	1827.75	1409.36	1243.42	working
08/08/2018 09:41	1972.44	1827.75	1409.36	1243.42	967.83	idle
08/08/2018 09:42	1827.75	1409.36	1243.42	967.83	624.71	idle
08/08/2018 09:43	1409.36	1243.42	967.83	624.71	310.61	idle
08/08/2018 09:44	1243.42	967.83	624.71	310.61	169.83	idle
08/08/2018 09:45	967.83	624.71	310.61	169.83	260.18	idle
.
.

08/08/2018 10:24	0.03	0.00	0.01	0.02	0.01	idle
08/08/2018 10:25	0.00	0.01	0.02	0.01	0.01	stop
08/08/2018 10:26	0.01	0.02	0.01	0.01	0.02	idle
08/08/2018 10:27	0.02	0.01	0.01	0.02	0.01	idle
08/08/2018 10:28	0.01	0.01	0.02	0.01	0.00	idle
08/08/2018 10:29	0.01	0.02	0.01	0.00	0.00	idle
08/08/2018 10:30	0.02	0.01	0.00	0.00	0.00	idle
08/08/2018 10:31	0.01	0.00	0.00	0.00	0.00	idle
08/08/2018 10:32	0.00	0.00	0.00	0.00	0.00	stop
08/08/2018 10:33	0.00	0.00	0.00	0.00	0.00	stop
08/08/2018 10:34	0.00	0.00	0.00	0.00	0.00	stop
08/08/2018 10:35	0.00	0.00	0.00	0.00	0.01	stop
08/08/2018 10:36	0.00	0.00	0.00	0.01	0.01	stop
08/08/2018 10:37	0.00	0.00	0.01	0.01	0.00	stop
08/08/2018 10:38	0.00	0.01	0.01	0.00	0.00	stop

As can be seen in Table 6.4., the first columns of the table are the original power data of the cutting machine. In order to eliminating the noises in the dataset, a filter is applied to it through shifting the column of the original power data and obtaining four more power columns. The status column is created according to the thresholds that are determined based on the original power data. These thresholds are given below:

P: Power Value

Stop: $P=0$

Idle: $0 < P \leq 10000$

Working: $10000 < P \leq 36000$

After creating the data set and giving the status according to the thresholds, the Matlab code that has classification tree algorithm is created in order to classify machine status (stop, idle or working) successfully. The code and its line-by-line explanation is presented in Appendix 3.

Explanation of the Code that is given in Appendix 3:

a) Training Level:

- The Excel file, which contains 6,300 units of power data of the cutting machine of Company 2 and their given status, is identified. Subsequently, Sheet 1 of the file is read and the data that are between a794 and f1213 are defined as training data for training the decision tree model.

- The number of the training data set is defined. The training data set starts with n0 and it ends with n.
- Input of the decision tree model, which includes power values of the machine from column 1 to 5, is defined and is named as X.
- The output of decision tree model, which covers “Status” column of the dataset, is defined and is called Y.
- L1, L2, L3 arrays are created and defined false for the beginning.
- Afterwards, the status, which is “stop”, “idle” and “working”, is converted to numbers that is 0,1 and 2 through “for” loop. In the loop, if the status is “Idle”, L1 becomes True and the status is converted to 1. If it is “Working”, L2 becomes True and the status is changed to 2. Lastly, if the status is “Stop”, L3 becomes True and the status is converted to 0. The code recognises the status via their character numbers.
- The model that has classification tree function is generated. The input of the model is X and the output of it is Y1. The aim is training the model to classify machine status successfully according to the thresholds that are mentioned above.

b) Optimization Level:

- Initially, a classification tree graph of the model is created in order to see the thresholds that the model determined. If the thresholds in the graph fit with the thresholds that are determined before writing the code, it means that the training process of the model is completed successfully.
- The second optimization step is k-fold cross-validation of the model in order to test the performance of the classification tree model. Cross-validation process predicts loss of the model through these three steps: Firstly, the data set is randomly divided into 10 equal parts as training and test data. Subsequently, the model is trained 10 times. For each training process, a division is selected as test data from part of the data set and the other 9 parts of dataset are left for training. After cross-validation is completed, the loss function of the cross-validated model is calculated. If the result is zero, there is no misclassification.

c) Prediction Level:

- In the prediction level, the data between a1214 and f6301 in sheet 1 of the Excel file that is identified in the first line of the code is read. Afterwards, the number of the testing

dataset is defined.

- Input of the prediction level, which includes 5,086 power level values as numbers from column 1 to 5, is defined and named as XX.
- The code predicts the status of the testing data through utilizing the created model and input XX. The model is expected to classify testing data by using 0,1 and 2 machine status. The predicted output is called “yfit”.

Results of the Decision Tree Code:

a) **Optimization Result:**

Classification Tree:

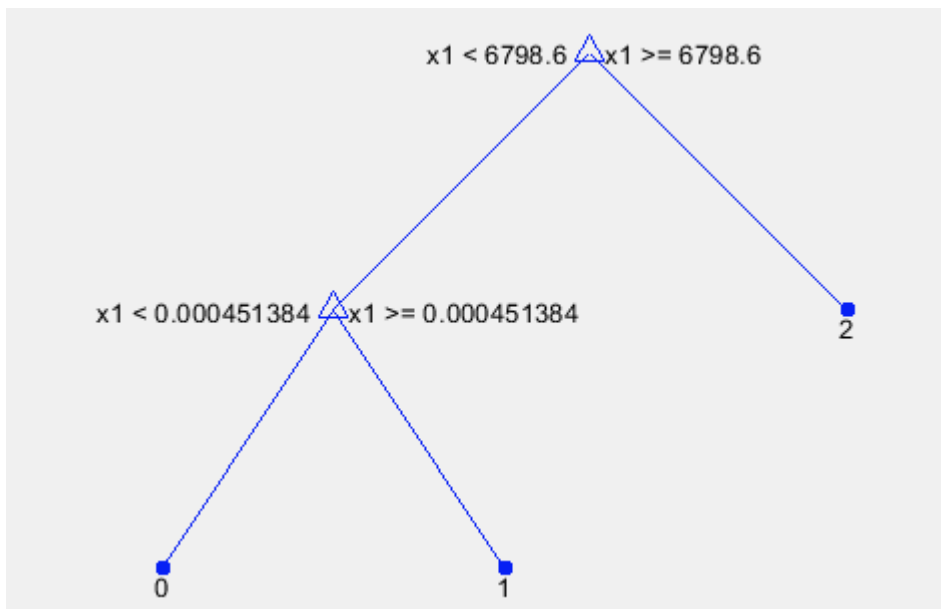


Figure 6.1: Classification Tree of Status of Cutting Machine

Cross-Validation Result

classreg.learning.partition.ClassificationPartitionedModel

CrossValidatedModel: 'Tree'

PredictorNames: {'x1' 'x2' 'x3' 'x4' 'x5'}

ResponseName: 'Y'

NumObservations: 418

KFold: 10

Partition: [1×1 cvpartition]

ClassNames: [0 1 2]

ScoreTransform: 'none'

Properties, Methods

$$L= 8.5763e^{-04}$$

Explanation of Optimization Results:

Firstly, the classification tree of the model is created in order to see whether the thresholds of the data set that the model determined are similar to the thresholds that were decided during the generation of the dataset. According to Figure 6.1, the thresholds of the model are as follows:

x1: Power Value

If $x1 \geq 6798.6$ then it is 2 (working) else if $x1 \geq 0.000451384$ then it is 1 (idle) else it is 0 (stop).

On the other hand, the thresholds that are created before the generation of the code are given below:

Stop (0): $P=0$

Idle (1): $0 < P \leq 10000$

Working (2): $10000 < P \leq 36000$

As can be seen in the comparison of two kind of thresholds, there is a small misclassification about “idle” status. In decided thresholds, “idle” status is smaller than 6798.6 and in the thresholds that are determined by the model, it is smaller than 10000. This misclassification can cause small errors in the process of the classification of machine status of the cutting machine by the model.

According to cross-validation results, the type of the cross- validated model is “tree”. The predictor names that show inputs’ names are x1, x2, x3, x4 and x5. It is reasonable because the input of the model has 5 columns of power data and each predictor name represents each column. Moreover, Response Name means output name and, as it can be seen in the code, the determined output of the model is Y. On the other hand, NumObservations symbolises the number of training data set that is 418. The value of KFold is 10. It means that the data set is randomly separated into 10 equal parts as training and testing data and the model was trained 10 times. The next term of the result of cross-validation is Partition. Cvpartition represents the class of partition and it is utilized in order to generate a cross-validated model. Besides, the result of cross- validation of the model gives class names that are 0, 1 and 2. The class names

fit with machine status that were determined and converted to numbers. The last term is ScoreTransform that is responsible for converting estimated classification scores. In this situation, ScoreTransform is “none” which means that ScoreTransform is x and there is no transformation of classification scores. After the cross validation of the model, the loss function of the cross-validation process is 8.5763e-04. The loss function value is very close to zero, it means that there is almost no misclassification during the cross-validation of the model.

a) Prediction Results:

A part of the prediction result of decision tree code of the classification of the cutting machine of Company 2 is given in Table 6.5 that includes date and time of recording the power value, power values from column 2 to 6, determined status (Y) and predicted status (yfit):

Table 6.5: Prediction Results of Decision Tree Code of Company 2

Date and Time	Power 1 (Original Data)	Power 2	Power 3	Power 4	Power 5	Determined Status (Y)	Predicted Status (yfit)
07/08/2018 16:01	0.000627	0.002083	0.000587	0.000426	0	idle	1
07/08/2018 16:02	0.002083	0.000587	0.000426	0	0	idle	1
07/08/2018 16:03	0.000587	0.000426	0	0	0.000177	idle	1
07/08/2018 16:04	0.000426	0	0	0.000177	0.000294	idle	0
07/08/2018 16:05	0	0	0.000177	0.000294	0.000614	stop	0
07/08/2018 16:06	0	0.000177	0.000294	0.000614	0.000371	stop	0
07/08/2018 16:07	0.000177	0.000294	0.000614	0.000371	0	idle	0
07/08/2018 16:08	0.000294	0.000614	0.000371	0	0	idle	0
07/08/2018 16:09	0.000614	0.000371	0	0	0.000173	idle	1
07/08/2018 16:10	0.000371	0	0	0.000173	0	idle	0
07/08/2018 16:11	0	0	0.000173	0	0	stop	0
07/08/2018 16:12	0	0.000173	0	0	0	stop	0
07/08/2018 16:13	0.000173	0	0	0	0	idle	0
07/08/2018 16:14	0	0	0	0	0	stop	0
07/08/2018 16:15	0	0	0	0	0.001093	stop	0
07/08/2018 16:16	0	0	0	0.001093	0.00177	stop	0
07/08/2018 16:17	0	0	0.001093	0.00177	0.001428	stop	0
07/08/2018 16:18	0	0.001093	0.00177	0.001428	0.000598	stop	0
07/08/2018 16:19	0.001093	0.00177	0.001428	0.000598	0.000797	idle	1
07/08/2018 16:20	0.00177	0.001428	0.000598	0.000797	0.000592	idle	1
07/08/2018 16:21	0.001428	0.000598	0.000797	0.000592	0.000726	idle	1
07/08/2018 16:22	0.000598	0.000797	0.000592	0.000726	0.001496	idle	1
07/08/2018 16:23	0.000797	0.000592	0.000726	0.001496	0.001647	idle	1
07/08/2018 16:24	0.000592	0.000726	0.001496	0.001647	0.000869	idle	1
07/08/2018 16:25	0.000726	0.001496	0.001647	0.000869	0.00027	idle	1
07/08/2018 16:26	0.001496	0.001647	0.000869	0.00027	0.000288	idle	1
07/08/2018 16:27	0.001647	0.000869	0.00027	0.000288	0	idle	1
07/08/2018 16:28	0.000869	0.00027	0.000288	0	0	idle	1

07/08/2018 16:29	0.00027	0.000288	0	0	0	idle	0
07/08/2018 16:30	0.000288	0	0	0	0	idle	0
07/08/2018 16:31	0	0	0	0	0	stop	0
07/08/2018 16:32	0	0	0	0	0	stop	0
07/08/2018 16:33	0	0	0	0	0.000131	stop	0
07/08/2018 16:34	0	0	0	0.000131	0	stop	0
07/08/2018 16:35	0	0	0.000131	0	0	stop	0
07/08/2018 16:36	0	0.000131	0	0	0	stop	0
07/08/2018 16:37	0.000131	0	0	0	0	idle	0
07/08/2018 16:38	0	0	0	0	0	stop	0
07/08/2018 16:39	0	0	0	0	0.000666	stop	0
07/08/2018 16:40	0	0	0	0.000666	0.001623	stop	0
07/08/2018 16:41	0	0	0.000666	0.001623	0.003268	stop	0
07/08/2018 16:42	0	0.000666	0.001623	0.003268	0.002773	stop	0
07/08/2018 16:43	0.000666	0.001623	0.003268	0.002773	0	idle	1
07/08/2018 16:44	0.001623	0.003268	0.002773	0	0.001135	idle	1
07/08/2018 16:45	0.003268	0.002773	0	0.001135	0.000498	idle	1
07/08/2018 16:46	0.002773	0	0.001135	0.000498	0.000284	idle	1
07/08/2018 16:47	0	0.001135	0.000498	0.000284	0.00052	stop	0
07/08/2018 16:48	0.001135	0.000498	0.000284	0.00052	0.00053	idle	1
07/08/2018 16:49	0.000498	0.000284	0.00052	0.00053	0.000192	idle	1
07/08/2018 16:50	0.000284	0.00052	0.00053	0.000192	0.000143	idle	0
07/08/2018 16:51	0.00052	0.00053	0.000192	0.000143	0	idle	1
07/08/2018 16:52	0.00053	0.000192	0.000143	0	0	idle	1
07/08/2018 16:53	0.000192	0.000143	0	0	0.000545	idle	0
07/08/2018 16:54	0.000143	0	0	0.000545	0.000961	idle	0
07/08/2018 16:55	0	0	0.000545	0.000961	0.000153	stop	0

In Table 6.5., the model classified the original power values of the cutting machine to 0 (stop), 1 (idle) and 2(working). The predicted output of the model is compared to determined output that was decided before the generation of the model. It can be seen that there are some mistakes in the classification of “idle”(1) status of the machine as was predicted while analysing the classification tree graph. The classification mistakes are shown with yellow colour.

6.1.3. Application of Multi-Class Support Vector Machine for Milling Machine Status Classification

The aim of this section is classifying the status of the milling machine in the Liverpool John Moores University Laboratory according to its power values. As a result of the experiment, the power meter had been logging on the milling machine for three weeks and 30244 units of power value data were collected. In order to complete the classification of machine status successfully, it was decided to utilise multiclass model for support vector machine (svm) in Matlab.

Subsequently, every week was analysed separately and in the first week, the code was trained and predicted the status of the power level of the first week. In the second and third weeks, the code only made predictions based on the trained model of svm. The machine has two status: Idle (the machine is on however it does not do milling.) and working (the machine does milling.). The thresholds are given below:

P: Power Value

Idle: $P \leq 600$

Working: $600 < P \leq 2000$

A part of the data set of the milling machine is given in Table 6.6. as an example in order to start the classification through svm method:

Table 6.6: A Part of the Data set of the Milling Machine

<u>Date and Time</u>	<u>Power 1 (Original Data)</u>	<u>Power 2</u>	<u>Power 3</u>	<u>Power 4</u>	<u>Power 5</u>	<u>Status</u>
01/11/2019 15:58	506.05	506.00	506.01	506.27	505.91	idle
01/11/2019 15:59	506.00	506.01	506.27	505.91	506.35	idle
01/11/2019 16:00	506.01	506.27	505.91	506.35	507.78	idle
01/11/2019 16:01	506.27	505.91	506.35	507.78	506.92	idle
.
.
01/11/2019 16:31	482.24	651.51	1665.92	1693.77	1673.93	idle
01/11/2019 16:32	651.51	1665.92	1693.77	1673.93	733.86	working
01/11/2019 16:33	1665.92	1693.77	1673.93	733.86	504.97	working
01/11/2019 16:34	1693.77	1673.93	733.86	504.97	813.20	working
01/11/2019 16:35	1673.93	733.86	504.97	813.20	1678.40	working
01/11/2019 16:36	733.86	504.97	813.20	1678.40	1690.12	working
01/11/2019 16:37	504.97	813.20	1678.40	1690.12	1659.19	idle
01/11/2019 16:38	813.20	1678.40	1690.12	1659.19	1681.51	working
01/11/2019 16:39	1678.40	1690.12	1659.19	1681.51	1695.35	working
01/11/2019 16:40	1690.12	1659.19	1681.51	1695.35	1658.00	working
01/11/2019 16:41	1659.19	1681.51	1695.35	1658.00	1680.25	working
01/11/2019 16:42	1681.51	1695.35	1658.00	1680.25	1598.95	working
01/11/2019 16:43	1695.35	1658.00	1680.25	1598.95	1620.84	working
01/11/2019 16:44	1658.00	1680.25	1598.95	1620.84	1683.47	working
01/11/2019 16:45	1680.25	1598.95	1620.84	1683.47	1173.69	working
01/11/2019 16:46	1598.95	1620.84	1683.47	1173.69	490.21	working

01/11/2019 16:47	1620.84	1683.47	1173.69	490.21	489.46	working
01/11/2019 16:48	1683.47	1173.69	490.21	489.46	531.44	working
01/11/2019 16:49	1173.69	490.21	489.46	531.44	488.86	working
01/11/2019 16:50	490.21	489.46	531.44	488.86	487.95	idle
01/11/2019 16:51	489.46	531.44	488.86	487.95	488.48	idle
01/11/2019 16:52	531.44	488.86	487.95	488.48	488.02	idle
01/11/2019 16:53	488.86	487.95	488.48	488.02	488.84	idle
01/11/2019 16:54	487.95	488.48	488.02	488.84	488.86	idle
01/11/2019 16:55	488.48	488.02	488.84	488.86	487.85	idle
01/11/2019 16:56	488.02	488.84	488.86	487.85	487.57	idle
01/11/2019 16:57	488.84	488.86	487.85	487.57	488.54	idle

A filter is applied to the data set of the milling machine by shifting the column of the original power values of the machine (Power 1) four times in order to make the data set cleaner. The status column is determined according to original power values. The matlab codes of week 1, week 2 and week 3 of the experiment on the milling machine are shown in Appendix 4:

Explanation of the Code that is Presented in Appendix 4:

Week 1:

- An Excel file that includes 30,244 training and testing power values of the milling machine is defined and second sheet of the file that is called “Shifting” is read and the data between a5575 and f6000 is selected for training the model.
- The number of the training data set that starts with the parameter n0 and ends with the parameter n is defined.
- The input of the model that covers power level values of the machine from column 1 to 5 is identified and it is called X.
- The output of the model that has the status of the milling machine (idle or working) based on its power values is defined as text.
- An array, which is named as L1, is defined and is given False before converting the status from text to numbers.
- A variable Y1 is defined through giving zero to each status of the power values of the milling machine.
- The “for” loop is created in order to convert machine status from text to numbers. In the loop, if the status is “Working”, L1 becomes True and Y1 becomes 2. If it is not “Working”, Y1 becomes 2. Therefore, Y1 becomes the new output of the model that includes status as numbers. (1 or 2).

- The last part of the training level is creating the model that has multi-class support vector machine function. The model is generated based on the input X and the output Y1.
- In the optimization level, the loss function of the model “Mdl” is calculated in order to see the efficiency of the model.
- The second step of the optimization level is the optimization of the hyperparameters of the model “Mdl”. The optimized model is called “Mdl2”.

Week 2

- Sheet 2 of the file that is defined in the first line of the code, which is named as “Shifting”, is defined and the data between a11634 and f14430 are selected as testing data set of week 2.
- The testing data set of week 2 is identified and the beginning of the dataset is n0 and the end of it is n.
- The input of the prediction level, which is called XX, is defined. The input covers 2,795 power values of the milling machine from column 1 to 5.
- The code predicts the output of the testing data through using the model “Mdl2”. The predicted output of the week 2 is yfit2.

Week 3

- Sheet 2 of the file that is defined in the first line of the code, which is named as “Shifting”, is defined and the data between a21484 and f23949 are selected as testing data set of week 3.
- The testing data set of week 3 is identified and the beginning of the dataset is n0 and the end of it is n.
- The input of the prediction level, which is called XX, is defined. The input covers 2,464 power values of the milling machine from column 1 to 5.
- The code predicts the output of the testing data through using the model “Mdl2”. The predicted output of the week 3 is yfit3.

Results of Multi- Support Vector Machine Code of Milling Machine:

Optimization Result

The Loss Function before the Optimization of the Hyperparameters of the Model “Mdl”:

lStd = 0.9434

The Hyperparameters Optimization Results:

Optimization completed.

MaxObjectiveEvaluations of 30 reached.

Total function evaluations: 30

Total elapsed time: 214.6179 seconds.

Total objective function evaluation time: 179.5665

Best observed feasible point:

<u>Coding</u>	<u>BoxConstraint</u>	<u>KernelScale</u>
onevsall	15.511	994.76

Observed objective function value = 0

Estimated objective function value = -0.0020574 Function evaluation time = 1.7179

Best estimated feasible point (according to models):

<u>Coding</u>	<u>BoxConstraint</u>	<u>KernelScale</u>
onevsall	0.0077052	584.47

Estimated objective function value = -0.0020574

Estimated function evaluation time = 0.12954

Explanation of the Optimization Results:

The loss function of the model is 0.9434 before the optimization of the hyperparameters of the model. If it were zero, it would mean that the model classifies the power values with zero error.

In the optimization results, total time of the model hyperparameters optimization is 214.6179 seconds as well as total loss function evaluation time is 179.5665 seconds during 30 iterations. The hyperparameter optimization of the model has two kind of results: Estimated and observed point. According to these points, the explanation of the hyperparameters is given in the following:

Coding: One-versus-all coding design is utilized during the optimization. In this kind of coding design, it includes 2 different classifiers because there are 2 different classes that are 1 (working) and 2 (idle). The first classifier is trained while assigning all the samples in the first

class as positive and the others are negative.

Box Constraint: Estimated box constraint value is 0.0077052. On the other hand, observed box constraint value is 15.511. This result shows that there are fewer support vectors of the model in the observed situation than in the estimated situation. When the number of support vectors decrease, the misclassification of the data set in multi-class SVM algorithm increases.

Kernel Scale: The estimated kernel scale value is 584.47 and the observed value of it is 994.76. It can be seen that the distance of the support vectors of the model in the observed optimization results is bigger than the model in the estimated results. Therefore, the function of the model is smoother based on observed optimization results than the function based on estimated optimization results.

As a result of the hyperparameters optimization, there are two kind of objective function value, which is loss function value, results: the estimated and observed objective function values. The estimated objective (loss) function value is -0.0020574 and the observed value of it is 0. It means that in the expected situation, because of the negative loss function value, the data set is misclassified. However, in the observed situation, the loss function is 0. According to the observed function, the model can classify the dataset with zero error.

Besides, the loss function value is decreased compared to the value before the optimization. It means that the hyperparameter optimization of the model achieved its goal and made the loss function zero.

Prediction Results:

a) **Prediction Results of Week 1:**

In order to examine the accuracy of the predicted output of week 1, Table 6.7, which covers power values of the milling machine, given output Y and predicted output yfit, is created and is given in the following:

Table 6.7: Prediction Results of Week 1 of Milling Machine Experiment

Date and Time	Power 1 (Original Data)	Power 2	Power3	Power4	Power5	Given Status (Y)	Estimate d Output (yfit)
23/10/2019 05:47	516.54	517.00	515.87	515.88	515.80	idle	1
23/10/2019 05:48	517.00	515.87	515.88	515.80	516.56	idle	1
23/10/2019 05:49	515.87	515.88	515.80	516.56	516.76	idle	1
.
.

23/10/2019 09:47	497.93	500.87	499.95	500.15	500.00	idle	1
23/10/2019 09:48	500.87	499.95	500.15	500.00	502.07	idle	1
23/10/2019 09:49	499.95	500.15	500.00	502.07	501.82	idle	1
23/10/2019 09:50	500.15	500.00	502.07	501.82	501.47	idle	1
23/10/2019 09:51	500.00	502.07	501.82	501.47	500.83	idle	1

As can be seen in Table 6.7., the predicted status of the milling machine that is called yfit and given status of the machine, which are decided based on original power data and are named as Y, fit with each other. It means that the classification of the machine status of week 1 of the milling machine experiment is successful.

b) Prediction Level Results of Week 2:

The comparison of given status and predicted status of week 2 of the milling machine experiment is presented in Table 6.8 through displaying a part of the results of it:

Table 6.8: A Part of the Prediction Results of Week 2 of Milling Machine Experiment

Date and Time	Power 1 (Original Data)	Power 2	Power3	Power4	Power5	Given Status (Y)	Estimated Output (yfit2)
30/10/2019 12:42	520.23	519.50	519.37	519.37	520.37	Idle	1
30/10/2019 12:43	519.50	519.37	519.37	520.37	519.89	Idle	1
30/10/2019 12:44	519.37	519.37	520.37	519.89	519.82	Idle	1
30/10/2019 12:45	519.37	520.37	519.89	519.82	520.89	Idle	1
.
.
30/10/2019 13:06	506.81	1406.11	1838.49	1809.81	1826.25	Idle	1
30/10/2019 13:07	1406.11	1838.49	1809.81	1826.25	1806.85	Working	2
30/10/2019 13:08	1838.49	1809.81	1826.25	1806.85	1830.52	Working	2
30/10/2019 13:09	1809.81	1826.25	1806.85	1830.52	1798.46	Working	2
.
.
30/10/2019 13:25	1802.74	1190.01	492.76	499.58	488.60	Working	2
30/10/2019 13:26	1190.01	492.76	499.58	488.60	499.16	Working	2
30/10/2019 13:27	492.76	499.58	488.60	499.16	1091.63	Idle	1
30/10/2019 13:28	499.58	488.60	499.16	1091.63	1822.64	Idle	1
30/10/2019 13:29	488.60	499.16	1091.63	1822.64	1793.69	Idle	1
30/10/2019 13:30	499.16	1091.63	1822.64	1793.69	1801.35	Idle	1
30/10/2019 13:31	1091.63	1822.64	1793.69	1801.35	1810.93	Working	2
30/10/2019 13:32	1822.64	1793.69	1801.35	1810.93	1804.13	Working	2
30/10/2019 13:33	1793.69	1801.35	1810.93	1804.13	1782.16	Working	2
30/10/2019 13:34	1801.35	1810.93	1804.13	1782.16	1582.67	Working	2
30/10/2019 13:35	1810.93	1804.13	1782.16	1582.67	1802.16	Working	2

•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
30/10/2019 13:50	1632.43	1228.99	492.64	489.43	488.99	Working	2
30/10/2019 13:51	1228.99	492.64	489.43	488.99	488.48	Working	2
30/10/2019 13:52	492.64	489.43	488.99	488.48	488.37	Idle	1
30/10/2019 13:53	489.43	488.99	488.48	488.37	489.37	Idle	1
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
31/10/2019 00:12	536.29	536.88	536.61	534.98	535.48	Idle	1

A part of the data set of the second week of the milling machine experiment is given in Table 5.8., the given and estimated status of the machine are presented with the data set as well. The Matlab model that includes multi-class SVM function estimated the status correctly because, as it can be seen in Table 6.8., the predicted and estimated status are same.

c) Prediction Level Results of Week 3:

A part of the data set of week 3 of the milling machine experiment is presented in Table 6.9 with its given and estimated status (Y and yfit3):

Table 6.9: A Part of the Prediction Results of Week 3 of Milling Machine Experiment

Date and Time	Power 1 (Original Data)	Power 2	Power3	Power4	Power5	Given Status (Y)	Estimated Output (yfit3)
05/11/2019 16:57	493.32	493.28	493.10	493.34	491.83	Idle	1
05/11/2019 16:58	493.28	493.10	493.34	491.83	492.61	Idle	1
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
05/11/2019 17:05	492.16	491.62	895.40	495.58	491.49	Idle	1
05/11/2019 17:06	491.62	895.40	495.58	491.49	492.62	Idle	1
05/11/2019 17:07	895.40	495.58	491.49	492.62	491.08	Idle	2
05/11/2019 17:08	495.58	491.49	492.62	491.08	491.52	Idle	1
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
05/11/2019 17:15	513.22	841.51	936.79	920.75	477.15	Idle	1
05/11/2019 17:16	841.51	936.79	920.75	477.15	520.97	Idle	2
05/11/2019 17:17	936.79	920.75	477.15	520.97	964.60	Idle	2
05/11/2019 17:18	920.75	477.15	520.97	964.60	509.78	Idle	2
05/11/2019 17:19	477.15	520.97	964.60	509.78	485.09	Idle	1
05/11/2019 17:20	520.97	964.60	509.78	485.09	484.63	Idle	1
05/11/2019 17:21	964.60	509.78	485.09	484.63	477.43	Idle	2

05/11/2019 17:22	509.78	485.09	484.63	477.43	477.12	Idle	1
.
05/11/2019 17:29	477.79	823.62	1387.06	551.33	478.51	Idle	1
05/11/2019 17:30	823.62	1387.06	551.33	478.51	479.10	Idle	2
05/11/2019 17:31	1387.06	551.33	478.51	479.10	1363.14	Working	2
05/11/2019 17:32	551.33	478.51	479.10	1363.14	745.42	Idle	1
05/11/2019 17:33	478.51	479.10	1363.14	745.42	835.59	Idle	1
05/11/2019 17:34	479.10	1363.14	745.42	835.59	1042.70	Idle	1
05/11/2019 17:35	1363.14	745.42	835.59	1042.70	1345.59	Working	2
05/11/2019 17:36	745.42	835.59	1042.70	1345.59	480.20	Idle	2
05/11/2019 17:37	835.59	1042.70	1345.59	480.20	478.19	Idle	2
05/11/2019 17:38	1042.70	1345.59	480.20	478.19	480.22	Working	2
05/11/2019 17:39	1345.59	480.20	478.19	480.22	479.37	Working	2
05/11/2019 17:40	480.20	478.19	480.22	479.37	479.88	Idle	1
.
05/11/2019 18:10	484.72	1484.17	827.66	479.48	478.13	Idle	1
05/11/2019 18:11	1484.17	827.66	479.48	478.13	479.77	Working	2
05/11/2019 18:12	827.66	479.48	478.13	479.77	478.99	Idle	2
05/11/2019 18:13	479.48	478.13	479.77	478.99	477.30	Idle	1
.
06/11/2019 01:08	487.91	488.49	487.67	487.86	488.21	Idle	1

In week 3 prediction results of the milling machine, the given status of the rows that are coloured yellow in Table 5.9 are “Idle”. However, the code predicted them as “Working”.

Therefore, it can be seen that the model could not estimate the status of the week 3 experiment of the milling machine with 100% accuracy.

6.1.4. Application of Decision Tree Algorithm for Classification of Status of Milling Machine:

After the results of the application of multi-class support vector machine to the data set of the milling machine power values, it was decided to implement the decision tree algorithm to same data set in order to classify the status of the milling machine in the Liverpool John Moores University Laboratory. A part of the data set of the week 1 experiment is given in Table 6.10.

Table 6.10: A Part of the Data set of Milling Machine Power Values for Decision Tree

Date and Time	Power1 (Original Data)	Moving Average (5)	Status
21/10/2019 17:08	498.72	498.39	idle
⋮	⋮	⋮	⋮
21/10/2019 18:28	498.88	498.53	idle
21/10/2019 18:29	498.39	498.75	idle
21/10/2019 18:30	498.30	498.72	idle
21/10/2019 18:31	498.17	499.04	idle

As can be seen in Table 6.10., the moving average method is applied as a filter in order to eliminate noises in the data set. The moving average method is implemented via taking average of power values of each five value. The Matlab code of the classification of milling machine status that includes classification tree function is presented in Appendix 5 with its line by line explanation.

Explanation of the Code that is Given in Appendix 5:

Week 1

- An Excel file that has training data of week 1 and testing data of week1, week2 and week3 is defined. Subsequently, sheet of the Excel file that is called “Moving Average” is read and a part of week1 data that is between a11500 and c13000 is chosen for training the code.
- The number of training data set that begins with n0 and ends with n is defined.
- Input of the classification tree model that is named as X and includes power level values as numbers from column 1 to 2 is defined.
- Output of the classification tree model, which is called Y and covers “idle” and “working” status as a text, is identified.
- L1 array is defined and given false for these two status in the beginning.
- Variable Y1 is generated and is given zero for “idle” and “working”.
- A loop is created in order to convert the “idle” and “working” status from text to numbers. In the loop, if the status is “working”, L1 is true and Y1 is 2; Else L1 is false and Y1 is 1. Y1 becomes the new output of the model.
- A model, which generates classification tree and does classification through using input X

and output Y1, is created. This model can classify status 1 and 2. This line is end of the training level.

- In optimization level, initially, the classification tree that is generated by the model is analysed. Subsequently, the model is cross-validated in order to check whether it fits the classification rules that are determined before. Lastly, the loss function of the cross-validated model is calculated.
- In the testing level, firstly, sheet 3 of the file that is called “Decision Tree week 1” is read and the data between b2 and c10492 is selected for testing the model.
- The number of the testing data set is defined.
- The input XX that includes 10,491 power level values as numbers from column 1 to 2 is created for testing the model.
- Lastly, the code predicts the status of the testing data for week 1 through using the “predict” function of Matlab and the model Mdl and named the output set as “yfit”.

Week 2

- An Excel file that has training data of week 1 and testing data of week1, week2 and week3 is defined. Subsequently, sheet of the Excel file that is called “Decision Tree week 2” is read and week2 data that is between b2 and c10081 is chosen for testing the model.
- The number of the testing data set that begins with n0 and ends with n is defined.
- The input XX that includes 10,080 power level values as numbers from column 1 to 2 is created for testing the model.
- Lastly, the code predicts status of the testing data for week 2 through using the “predict” function of Matlab and the model Mdl and named the output set as “yfit2”.

Week 3

- An Excel file that has training data of week 1 and testing data of week1, week2 and week3 is defined. Subsequently, sheet of the Excel file that is called “Decision Tree week 3” is read and week3 data that is between b2 and c9671 is chosen for testing the model.
- The number of the testing data set that begins with n0 and ends with n is defined.
- The input XX that includes 9,670 power level values as numbers from column 1 to 2 is created for testing the model.
- Lastly, the code predicts status of the testing data for week 3 through using “predict” function of Matlab and the model Mdl and named the output set as “yfit3”.

Results of Decision Tree Code of Milling Machine:

Optimization Results

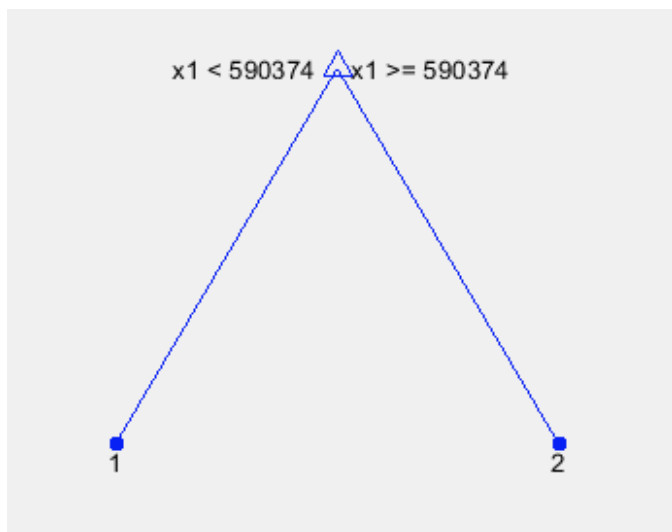


Figure 6.2: Classification Tree of Status of Milling Machine

- CrossValidatedModel: 'Tree'
PredictorNames: {'x1' 'x2'}
ResponseName: 'Y'
NumObservations: 1499 KFold: 10
Partition: [1×1 cvpartition]
ClassNames: [1 2]
ScoreTransform: 'none'
- L = 0.02068

Explanation of the Optimization Results of the Code:

- Initially, the classification tree of the model is generated and presented in Figure 6.2 in order to analyse the thresholds of the model in the code. According to the classification tree, original power values are shown as x1 and if x1 is smaller than 590.374, the status is 1 (idle). If x1 is equal or bigger than 590.374, the status is 2 (working). On the other hand, the threshold that was determined before creating the code is:

x1 < 600 : idle

x1 >= 600 : working

It can be seen that the threshold of the model is very close to the threshold that was decided before. However, more detailed analysis is done through cross-validation.

- In the results of cross-validation of the model, the type of the cross validated model is “tree”.

Predictor names that means inputs of the model are x1, which refers to original power data and x2 that represents the moving average column of the data set. The Response name that is the output of the model is Y. Moreover, the number of observations (NumObservations), which is the number of training power values, is 1,499. On the other hand, the value of KFold is 10. It means that the data set is randomly separated into 10 equal parts as training and testing data and the model was trained 10 times. Also, the class names of cross-validated model are 1 and 2. 1 is idle status and 2 means working. Lastly, ScoreTransform is “none” which means that ScoreTransform is x and there is no transformation of classification scores. After the cross-validation process of the model, the loss function of this process is calculated and the result is 0.02068. The value of loss function is very close to zero that is the ideal value. It means that the cross-validated model can classify the status of the milling machine with almost zero error.

Prediction Results:

a) Week 1 Result:

A part of the result of the classification of the power values of the week 1 milling machine experiment is given in Table 6.11 that covers date and time, original power values, moving average column, determined status (Y) and predicted status (yfit). Table 6.11 follows:

Table 6.11.: A Part of the Prediction Result of Week 1 Milling Machine Experiment

<u>Date and Time</u>	<u>Power1 (Original Data)</u>	<u>Moving Average (5)</u>	<u>Determined Status (Y)</u>	<u>Predicted Status (yfit)</u>
21/10/2019 17:08	498.72	498.39	idle	1
21/10/2019 17:09	499.42	498.16	idle	1
21/10/2019 17:10	498.25	497.79	idle	1
•	•	•	•	•
•	•	•	•	•
21/10/2019 17:55	500.68	500.14	idle	1
21/10/2019 17:56	499.75	500.02	idle	1
21/10/2019 17:57	500.22	500.13	idle	1

In Table 6.11, all status of milling machine in the week 1 experiment are idle and, as can be seen in the last column of the table, the model predicted the status of the milling machine with 100% accuracy.

b) Week 2 Result:

A part of the week 2 prediction results of the milling machine experiment is given in Table

6.12 with date and time power values, determined and predicted status:

Table 6.12: A Part of the Prediction Result of Week 2 Milling Machine Experiment

Date and Time	Power1 (Original Data)	Moving Average (5)	Determined Status (Y)	Predicted Status (yfit)
31/10/2019 13:58	478.24	492.36	idle	1
31/10/2019 13:59	489.38	491.96	idle	1
.
31/10/2019 14:41	496.86	563.20	idle	1
31/10/2019 14:42	815.93	598.50	idle	2
31/10/2019 14:43	489.08	684.15	idle	1
31/10/2019 14:44	507.94	925.35	idle	1
31/10/2019 14:45	506.21	1144.49	idle	1
31/10/2019 14:46	673.36	1298.35	working	2
31/10/2019 14:47	1244.14	1501.08	working	2
.
31/10/2019 15:00	1693.77	768.78	working	2
31/10/2019 15:01	690.59	527.29	working	2
31/10/2019 15:02	486.73	486.46	idle	1
31/10/2019 15:03	486.48	486.36	idle	1
.
31/10/2019 15:15	486.30	486.18	idle	1

In Table 6.12., when the determined and predicted output of the machine are compared with each other, it is seen that the code predicted output of the week 2 experiment of the milling machine with almost zero error. The misclassification of the prediction results is shown with yellow colour. The code predicted the status as 2 (working), when the status is idle according to determined thresholds.

c) Week 3 Results:

A part of the comparison of the determined and predicted output of the week 3 of the milling machine experiment is presented in Table 6.13:

Table 6.13: A Part of the Prediction Result of Week 3 Milling Machine Experiment

Date and Time	Power1 (Original Data)	Moving Average (5)	Determined Status (Y)	Predicted Status (yfit)
05/11/2019 15:08	509.54	509.37	idle	1
.
05/11/2019 17:06	491.62	573.34	idle	1
05/11/2019 17:07	895.40	573.23	working	2
05/11/2019 17:08	495.58	492.46	idle	1
05/11/2019 17:09	491.49	491.65	idle	1
05/11/2019 17:10	492.62	492.32	idle	1
.
05/11/2019 17:15	513.22	737.88	idle	1
05/11/2019 17:16	841.51	739.43	working	2
05/11/2019 17:17	936.79	764.05	working	2
05/11/2019 17:18	920.75	678.65	working	2
05/11/2019 17:19	477.15	591.52	idle	1
05/11/2019 17:20	520.97	593.01	idle	1
05/11/2019 17:21	964.60	584.31	working	2
05/11/2019 17:22	509.78	486.81	idle	1
.
05/11/2019 17:26	477.12	545.03	idle	1

As can be seen in Table 6.13, the model in the Matlab code predicted the status of the milling machine of the last week of the experiment with zero misclassification.

Before the next section that includes Matlab codes of the calculation of the total “working” status time, total time and utilization rate of the machine, there will be discussion about the reason for the utilization of the moving average filter method in the data set of the decision tree algorithm and the shifting method in the data set of the multi-class support vector machine algorithm.

An analysis was done in Matlab with the classification learner algorithm by using the milling machine experiment data set after some misclassifications in prediction results of Company 2. Initially, the moving average filter method was applied to the milling machine data set.

Subsequently, decision tree and binary svm algorithms were applied to dataset. The same confusion matrix is obtained as a result of the two machine learning method. The confusion matrix is given in Figure 6.3.:



Figure 6.3: The Confusion Matrix of the Binary SVM and Decision Tree Algorithm

As can be seen in Figure 6.3., the model predicted 29896 “idle” status and 346 “working” status. The model only did one misclassification which is that it predicted one “idle” status as “working”.

Afterwards, the shifting method was applied to the milling machine experiment data set as a filter and the decision tree algorithm was applied to the data set. The Confusion matrix of this process is shown as Figure 6.4.:



Figure 6.4: The Confusion Matrix of the Decision Tree Algorithm of “Shifting” Filter Method

Figure 6.4. indicates that the model estimated 29896 “idle” status and 344 “working” status. However, it predicted one “idle” status as “working” and 2 “working” status as “idle”. Therefore, it can be seen that the decision tree algorithm gives a better solution through using the data set that uses the moving average method as a filter.

6.2. Calculation of the Utilization Rate of the Machine:

After prediction of the status of the laser sintering, cutting and milling machines, the total time of status 2 (“working”) and the total time of the experiment will be calculated through Matlab Software codes in this section. Subsequently, based on total experiment time and total time of the status 2, the utilization rate of the machine will be calculated in the same codes. The formula of the utilization rate of the machine is given in the following before starting to analyse the code of these three machines:

$$URM = (TTS2/TTE) * 100$$

(6-1)

Where *URM*: Utilization Rate of the Machine

TTS2 (h): Total Time of Status 2

TTE (h): Total Time of the Experiment

6.2.1. Calculation of the Utilization Rate of the Laser Sintering Machine:

The aim of the experiment of the laser sintering machine is to calculate the utilization rate of the machine. Therefore, initially, the classification of the machine status was done and the Matlab model that has the multi-class svm algorithm classified the power values into three status: Stop (0), Idle (1) and Working (2). In this section, the utilization rate of the machine will be calculated according to the predicted status of the model through Matlab Software.

A part of the data set of the model that covers date and time, power values and predicted status is given below:

Table 6.14: A Part of the Dataset of Laser Sintering Machine for Time Calculation

Date and Time	Total Power (Original Value)	Predicted Status (yfit)
28/02/2018 14:02	14.41	1
28/02/2018 14:02	15.85	1
28/02/2018 14:02	17.21	1
28/02/2018 14:02	19.02	1
28/02/2018 14:02	20.00	2
28/02/2018 14:02	17.67	1
:	:	:
28/02/2018 14:02	15.15	1
28/02/2018 14:02	23.20	2
:	:	:
28/02/2018 14:02	20.04	2
28/02/2018 14:02	20.11	2
28/02/2018 14:02	18.36	1
28/02/2018 14:02	20.26	2
28/02/2018 14:02	22.33	2
:	:	:
28/02/2018 14:02	20.88	2
28/02/2018 14:02	19.09	1
28/02/2018 14:02	18.83	1
28/02/2018 14:02	16.63	1
28/02/2018 14:02	18.28	1
28/02/2018 14:02	19.06	1
28/02/2018 14:02	20.01	2
28/02/2018 14:02	19.65	1
28/02/2018 14:02	19.29	1
28/02/2018 14:02	18.09	1
28/02/2018 14:02	17.35	1
28/02/2018 14:02	18.77	1
28/02/2018 14:02	18.15	1
28/02/2018 14:02	22.39	2
28/02/2018 14:02	21.65	2
28/02/2018 14:02	21.42	2
28/02/2018 14:02	21.65	2
28/02/2018 14:02	17.23	1
28/02/2018 14:02	18.92	1

28/02/2018 14:02	17.95	1
28/02/2018 14:02	18.83	1
28/02/2018 14:02	18.41	1
28/02/2018 14:02	21.93	2
:	:	:
28/02/2018 14:02	30.23	2
28/02/2018 14:02	23.03	2
28/02/2018 14:02	19.05	1
28/02/2018 14:02	17.92	1
28/02/2018 14:02	19.21	1
28/02/2018 14:02	19.57	1
28/02/2018 14:02	20.60	2
28/02/2018 14:02	19.68	1
28/02/2018 14:02	18.50	1
28/02/2018 14:02	19.89	1
28/02/2018 14:02	18.97	1
28/02/2018 14:02	16.08	1
28/02/2018 14:02	19.26	1
28/02/2018 14:02	20.57	2
:	:	:
28/02/2018 14:02	20.31	2
28/02/2018 14:02	19.29	1
:	:	:
28/02/2018 14:02	15.90	1
28/02/2018 14:02	20.04	2
28/02/2018 14:02	20.25	2
28/02/2018 14:02	19.50	1
28/02/2018 14:02	18.52	1
28/02/2018 14:02	19.03	1
28/02/2018 14:02	19.00	1
28/02/2018 14:02	17.50	1
28/02/2018 14:02	22.08	2
:	:	:
28/02/2018 14:02	23.66	2
28/02/2018 14:02	15.73	1
:	:	:

28/02/2018 14:02	19.05	1
28/02/2018 14:02	24.21	2
28/02/2018 14:02	26.97	2
28/02/2018 14:02	24.00	2
28/02/2018 14:02	23.86	2
28/02/2018 14:02	22.04	2
28/02/2018 14:02	17.33	1
28/02/2018 14:02	18.22	1
28/02/2018 14:02	20.40	2
28/02/2018 14:02	23.20	2
28/02/2018 14:02	21.67	2
28/02/2018 14:02	19.33	1
28/02/2018 14:02	18.89	1
28/02/2018 14:02	16.10	1
28/02/2018 14:02	18.53	1
28/02/2018 14:02	17.38	1
28/02/2018 14:02	20.21	2
28/02/2018 14:02	19.04	1
28/02/2018 14:02	15.40	1
28/02/2018 14:02	19.19	1
28/02/2018 14:02	20.50	2
28/02/2018 14:02	22.87	2
28/02/2018 14:02	24.21	2
28/02/2018 14:02	23.63	2
28/02/2018 14:02	23.19	2
28/02/2018 14:02	17.16	1
:	:	:
28/02/2018 14:02	19.52	1
28/02/2018 14:02	22.40	2
28/02/2018 14:02	21.76	2

Table 6.14. indicates the part of the data set of the experiment of the Laser Sintering Machine of Company 1 that was explained in Section 1.2. The table includes the predicted status of the machine that was estimated via the multi-class support vector machine algorithm in Matlab Software.

The line-by-line explanation of Matlab code that has the calculation of the utilization rate of the laser sintering machine of Company 1 is given in Appendix 6.

Explanation of the Code of the Calculation of the Utilization Rate:

- The Excel file that includes the data set is defined and named as “FileName”.
- The name of the sheet that has the data set in the Excel file is given to the code and it is “moving average”.
- The range of the data set is determined and it is between A2 and E900.
- A variable that is called “ReadVariableNames” is defined and is given false in the beginning.
- Target status is determined as Status 2, that is “Working” status.
- Time column that includes date and time and status column of the excel file is introduced to the code. Time column is the first and status column is the fourth column.
- After the variables, which represent the Excel file, are defined, the Excel file is read by the code according to these variables.
- The size of time and status column should be equal in order to run the code. Therefore, if time column and status column do not have an equal number of data, exit the system.
- After Excel file variables, data set variables are defined. Initially, TSize variable that is equal to the size of the time vector is introduced. Subsequently, lengths of the vectors are defined with TLength variable. It has 1 column and 1 row and has the same number of data as the time column. Next, TempTv variable stores the time vector that has 1 column and the number of rows that are equal to TLength and default time is set as Tv that is an unknown date/time value. Lastly, TempSv variable stores the status vector that that has 1 column and the number of rows that are equal to TLength and default status is set as Sv, that is zero.
- It is important that time gaps between the data should be detected and should not be added to utilization rate calculation. Therefore, the maximum gap that is allowed between data is 1 hour and it is defined with “MaxTimeGapAllowed. If it is more than 1 hour, it is considered as missing time.
- Before creating the loops of the code, loop variables are presented. Firstly, “TotalValidTime” variable holds total experiment time (excluding missing time). Subsequently, “TotalTargetTime” variable holds total time of Status 2. It does not include missing time. On the other hand, total missing time is recorded by “MissedRecordsTime” variable. The beginning of Status 2 in the dataset is represented by “RangeStartIndex”. Lastly, the number of Status 2 in the data set is held by the “TargetCounter” variable.

- In the first loop, if the machine status is target status, which is status 2 (working), 1 is added to “TargetCounter” variable and “RangeStartIndex” becomes 1 because it is the beginning of Status 2.
- In the next loop, the period between two data in the data set is identified and is showed with “tb” variable.
- Normally, dates in the time column are in “day/month/year, hour:minute:second” format. However, the time column of the database is divided into ‘years’, ‘months’, ‘days’, ‘time’ of tb (the period between two data) with split function in order to calculate working, total experiment and missing time in the third loop.
- In the subsequent loop, if time, day, month or year is bigger than the gap allowed, MissedRecordsTime is increased as tb and if the status vector equals the target status, total target time is increased as the period between the last data that is working and the current one and RangeStartIndex variable is equal to the number of the next data. If there is no time gap, total valid time is increased as tb.
- The next loop is about next data that is mentioned in the previous bullet point. If i^{th} status is target status, 1 is added to TargetCounter.
- If $(i-1)^{\text{th}}$ status is target status instead of i^{th} one, RangeStartIndex is $(i - 1)$. If the previous $(i - 1)$ status is target status but it is not the first target status in the database, (RangeStartIndex is not equal to $(i - 1)$) and total target time is increased as the time period between the data that is target status and current one $(i-1)$.
- In the last loop, if last status is target status, total target time equals the time period between i^{th} status and status, that number equals RangeStartIndex.
- After the loops are completed, TotalTargetTime represents total time that the machine is actually producing a product and TotalValidTime refers to total idle and working time of the machine, converted from calendar duration to time duration. Subsequently, the formula of the utilization rate of the machine is written.

The Result of the Code:

The result of the Matlab code of Company 1 is given below:

```
'No. Records Searched:' [899]
'No. Records of ' [2] [424]
'Total Valid Time:' [0h 0m 8.001s]
'Total Target Time:' [0h 0m 5s]
```

'Missing Records For:' [6d 23h 35m 19s]

'Utilization Rate:' [62.4922]

The Explanation of the Code:

Initially, the code gives the number of data in the data set of the laser sintering machine that is 899. The data set includes date and time, power values and Status 1 (idle) and Status 2 (working) that were predicted by the multi-class SVM method. The number of Status 2 among the dataset is 424. Subsequently, there are time calculations of the Matlab code: Total Valid Time that covers idle and working time of the machine is 8.001 seconds and, among total valid time, the machine works for 5 seconds. It can be seen that total valid and working time is too low considering the experiment time. However, based on missing records results, the power meter did not record time for almost 7 days. On the other hand, according to the utilization rate result, the code and the calculation of the utilization rate of the laser sintering machine is accurate and it is 62.4922%.

6.2.2. Calculation of the Utilization Rate of the Cutting Machine:

After calculating the utilization rate of the laser sintering machine of Company 1, the utilization rate of the cutting machine of Company 2 will be calculated according to the status prediction results of Section 1.1.3 through Matlab A part of the dataset of the Matlab Code for calculating the utilization rate of the machine is given in Table 6.15:

Table 6.15: A Part of the Data set of Cutting Machine for Time Calculation

Date and Time	Power 1 (Original Value)	Power 2	Power 3	Power 4	Power 5	Predicted Status (yfit)
10/08/2018 04:10	0.0000	0.0008	0.0006	0.0004	0.0004	0
10/08/2018 04:11	0.0008	0.0006	0.0004	0.0004	0.0001	1
10/08/2018 04:12	0.0006	0.0004	0.0004	0.0001	0.0004	1
10/08/2018 04:13	0.0004	0.0004	0.0001	0.0004	0.0006	0
10/08/2018 04:14	0.0004	0.0001	0.0004	0.0006	0.0000	0
10/08/2018 04:15	0.0001	0.0004	0.0006	0.0000	0.0007	0
10/08/2018 04:16	0.0004	0.0006	0.0000	0.0007	0.0006	0
10/08/2018 04:17	0.0006	0.0000	0.0007	0.0006	0.0003	1
10/08/2018 04:18	0.0000	0.0007	0.0006	0.0003	0.0004	0
10/08/2018 04:19	0.0007	0.0006	0.0003	0.0004	0.0002	1
10/08/2018 04:20	0.0006	0.0003	0.0004	0.0002	0.0003	1
10/08/2018 04:21	0.0003	0.0004	0.0002	0.0003	0.0004	0
.
.
10/08/2018 04:40	0.0003	0.0009	0.0021	0.0021	0.0015	0
10/08/2018 04:41	0.0009	0.0021	0.0021	0.0015	0.0014	1
.
.
10/08/2018 05:00	0.0012	0.0005	0.0004	0.0005	0.0004	1
10/08/2018 05:01	0.0005	0.0004	0.0005	0.0004	0.0005	0
10/08/2018 05:02	0.0004	0.0005	0.0004	0.0005	0.0010	0
10/08/2018 05:03	0.0005	0.0004	0.0005	0.0010	0.0016	1
10/08/2018 05:04	0.0004	0.0005	0.0010	0.0016	0.0011	0
10/08/2018 05:05	0.0005	0.0010	0.0016	0.0011	0.0008	1
.
.
10/08/2018 05:11	0.0005	0.0004	0.0006	0.0020	0.0023	1
10/08/2018 05:12	0.0004	0.0006	0.0020	0.0023	0.0018	0
10/08/2018 05:13	0.0006	0.0020	0.0023	0.0018	0.0010	1
10/08/2018 05:14	0.0020	0.0023	0.0018	0.0010	0.0008	1
10/08/2018 05:15	0.0023	0.0018	0.0010	0.0008	0.0004	1
10/08/2018 05:16	0.0018	0.0010	0.0008	0.0004	0.0005	1
10/08/2018 05:17	0.0010	0.0008	0.0004	0.0005	0.0010	1
10/08/2018 05:18	0.0008	0.0004	0.0005	0.0010	0.0025	1
10/08/2018 05:19	0.0004	0.0005	0.0010	0.0025	0.0034	0
10/08/2018 05:20	0.0005	0.0010	0.0025	0.0034	0.0024	1
.
.

10/08/2018 05:30	0.0016	11510.96	11488.45	11418.97	11342.28	1
10/08/2018 05:31	11510.96	11488.45	11418.97	11342.28	11401.36	2
•	•	•	•	•	•	•
•	•	•	•	•	•	•
10/08/2018 10:30	8924.05	3465.33	605.41	631.75	585.74	2
10/08/2018 10:31	3465.33	605.41	631.75	585.74	491.46	1
•	•	•	•	•	•	•
•	•	•	•	•	•	•
10/08/2018 11:26	0.0014	0.0000	0.0000	0.0000	0.0000	1
10/08/2018 11:27	0.0000	0.0000	0.0000	0.0000	0.0000	0
•	•	•	•	•	•	•
•	•	•	•	•	•	•
10/08/2018 11:40	0.0000	0.0130	0.0000	0.0000	0.0000	0
10/08/2018 11:41	0.0130	0.0000	0.0000	0.0000	0.0000	1
10/08/2018 11:42	0.0000	0.0000	0.0000	0.0000	0.0000	0
•	•	•	•	•	•	•
•	•	•	•	•	•	•
10/08/2018 12:00	0.0000	0.0007	0.0006	0.0008	0.0011	0
10/08/2018 12:01	0.0007	0.0006	0.0008	0.0011	0.0012	1
10/08/2018 12:02	0.0006	0.0008	0.0011	0.0012	0.0006	1
10/08/2018 12:03	0.0008	0.0011	0.0012	0.0006	0.0000	1
10/08/2018 12:04	0.0011	0.0012	0.0006	0.0000	0.0000	1

As can be seen in Table 6.15, the data set includes date and time, power values and predicted status (yfit) that were estimated through the decision tree algorithm in Matlab. After the dataset was prepared, the code that is presented and explained in Section 6.2. was applied and the results are given in the following:

'No. Records Searched:' [5088]

'No. Records of ' [2] [1199]

'Total Valid Time:' [84h 46m 59.877s]

'Total Target Time:' [19h 58m 59.83s]

'Missing Records For:' [0]

'Utilization Rate:' [23.5698]

The Explanation of the Results:

In the results of the Matlab code, the number of data that includes date and time, power values of the machine and predicted status is 5,088 and 1,199 of them are Status 2 (working). Total time that covers Status 0 (stop), Status 1 (idle) and Status 2 (working) is 84 hours 46 minutes 59.877 seconds. On the other hand, total time of Status 2, that indicates total time that the machine is actually working, is 19 hours 58 minutes 59.83 seconds. There are no missing records in the data set and , based on these results, the utilization rate of the cutting machine during the experiment is 23.5698%.

6.2.3. Calculation of the Utilization Rate of the Milling Machine:

In Section 2.1. and 2.2., the utilization rates of the laser sintering and cutting machines were calculated successfully and it is verified that the Matlab code, which is given in Section 2.1., runs effectively. Therefore, in this section, the Matlab code will be applied to milling machine experiment that was carried out in the Liverpool John Moores University laboratory for three weeks. The utilization rate of the machine will be calculated for each week according to power values that were predicted in Section 6.1.5. A part of the data set of the week 1 of the milling machine experiment for calculating the utilization rate of the machine is presented as an example in Table 6.16:

Table 6.16: A Part of the Data set of Milling Machine for Time Calculation

Date and Time	Original Power Values	Predicted Status (yfit)
25/10/2019 16:20	491.49	1
25/10/2019 16:21	489.05	1
25/10/2019 16:22	502.23	1
25/10/2019 16:23	983.49	2
25/10/2019 16:24	585.57	1
25/10/2019 16:25	837.35	2
25/10/2019 16:26	607.01	2
25/10/2019 16:27	487.07	1
25/10/2019 16:28	472.51	1
25/10/2019 16:29	472.82	1
25/10/2019 16:30	744.53	2
25/10/2019 16:31	975.31	2
25/10/2019 16:32	950.86	2
25/10/2019 16:33	863.55	2
25/10/2019 16:34	481.10	1
25/10/2019 16:35	474.54	1
25/10/2019 16:36	874.80	2
25/10/2019 16:37	1009.60	2
25/10/2019 16:38	976.28	2
25/10/2019 16:39	512.02	1
25/10/2019 16:40	474.39	1
25/10/2019 16:41	479.54	1
25/10/2019 16:42	937.13	2
25/10/2019 16:43	987.27	2
25/10/2019 16:44	770.91	2
25/10/2019 16:45	822.36	2

25/10/2019 16:46	495.47	1
25/10/2019 16:47	483.96	1
25/10/2019 16:48	484.43	1
25/10/2019 16:49	491.62	1
25/10/2019 16:50	484.64	1
25/10/2019 16:51	484.01	1
25/10/2019 16:52	484.20	1

As can be seen in Table 5.16, the data set of each week of the milling machine experiment contains date and time, original power values and predicted status that were predicted through the decision tree algorithm in Section 6.1.5.

After the Matlab code was applied to the dataset of week 1, week 2 and week 3 of the milling machine experiment, total experiment time, total working time of the milling machine and the utilization rate of the machine are calculated and given below for each week:

a) The Results of Week 1 of the Experiment

'No. Records Searched:' [10491]

'No. Records of ' [2] [24]

'Total Valid Time:' [174h 49m 58.06s]

Total Target Time:' [0h 23m 59.673s]

'Missing Records For:' [0]

'Utilization Rate:' [0.2287]

Explanation: The number of data in the data set of the week 1 experiment of the milling machine is 10,491 and, among these data, 24 of them are Status 2 (working). Therefore, total experiment time is 174 hours 49 minutes and 58.06 seconds. On the other hand, total working hours of the milling machine is 23 minutes 59.673 seconds. There are no missing records and, based on these results, the utilization rate of the milling machine for week 1 is 0.2287%.

b) The Results of Week 2 of the Experiment

'No. Records Searched:' [10080]

'No. Records of ' [2] [279]

'Total Valid Time:' [167h 59m 0.067s]

'Total Target Time:' [4h 39m 0.283s]

'Missing Records For:' [0]

'Utilization Rate:' [2.7682]

Explanation: The number of data in the data set of the week 2 experiment of the milling machine is 10,080 and, among these data, 279 of them are Status 2 (working). Therefore, total experiment time is 167 hours 59 minutes and 0.067 seconds. On the other hand, total working hours of the milling machine is 4 hours 39 minutes 0.283 seconds. There are no missing records and, based on these results, the utilization rate of the milling machine for week 2 is 2.7682%.

c) The Results of Week 3 of the Experiment

'No. Records Searched:' [9670]

'No. Records of ' [2] [279]

'Total Valid Time:' [161h 9m 0.065s]

'Total Target Time:' [4h 39m 0.283s]

'Missing Records For:' [0]

'Utilization Rate:' [2.8856]

Explanation: The number of data in the data set of the week 3 experiment of the milling machine is 9,670 and, among these data, 279 of them are Status 2 (working). Therefore, total experiment time is 161 hours 9 minutes and 0.065 seconds. On the other hand, total working hours of the milling machine is 4 hours 39 minutes 0.283 seconds. There are no missing records and, based on these results, the utilization rate of the milling machine for week 3 is 2.8856%.

6.4. Validation of the System of Calculation of the Utilization Rate of the Machine

After classification of the status of laser sintering, cutting and milling machine (stop, idle or working) according to thresholds that were determined based on power values of the machines and calculation of total experiment time, total working time and the utilization rate of each machine according to predicted status, the system that includes machine status classification and utilization rate calculation will be verified through applying the Case Based Reasoning Method.

CHAPTER 7: FRAMEWORK of the KNOWLEDGE SUPPORT SYSTEM

This chapter presents the framework development and the explanations of the knowledge support system that is generated for the calculation of unit manufacturing cost of a product and manufacturing process optimization. The features of the database are given in the following:

- The capability to indicate the utilization rate of each machine in a product manufacturing process. The utilisation rate is calculated based on the time and power data that is transferred from power meter to knowledge support system via the Internet. through the Matlab code explained in Chapter 6.
- The calculation of the cost of a product in a manufacturing process can be calculated through a generic cost model that includes direct and overhead costs of the process with the consideration of production time and power data collected from the machines. The utilization rate of the machine is added to the cost model during the calculation of the process overhead costs.
- Cased Based Reasoning method is applied in order to estimate the utilization rate of a machine in the manufacturing process of a product based on similar previous cases. This estimation is for calculating the cost of a process before starting the manufacturing process.
- The facility that allows an operator to check the steps of the manufacturing process and the drawings of the product that were produced via the related process. Also the operator can record the manufacturing process in the database as a case and it can see the previous cases of the manufacturing process.

This chapter consists of two sub-sections: Designing Level of the system and Creation Level of the system.:

7.1. Designing Level of the Knowledge Support System:

This level includes the flow diagram of the base of the knowledge support system database. The flow diagram was built in order to design the knowledge support system is which for optimization of a manufacturing process and calculation or estimation of unit manufacturing cost of a product through a sensor-based cost model that gives information about efficiency of a manufacturing process. The flow diagram in Figure 7.1 follows:

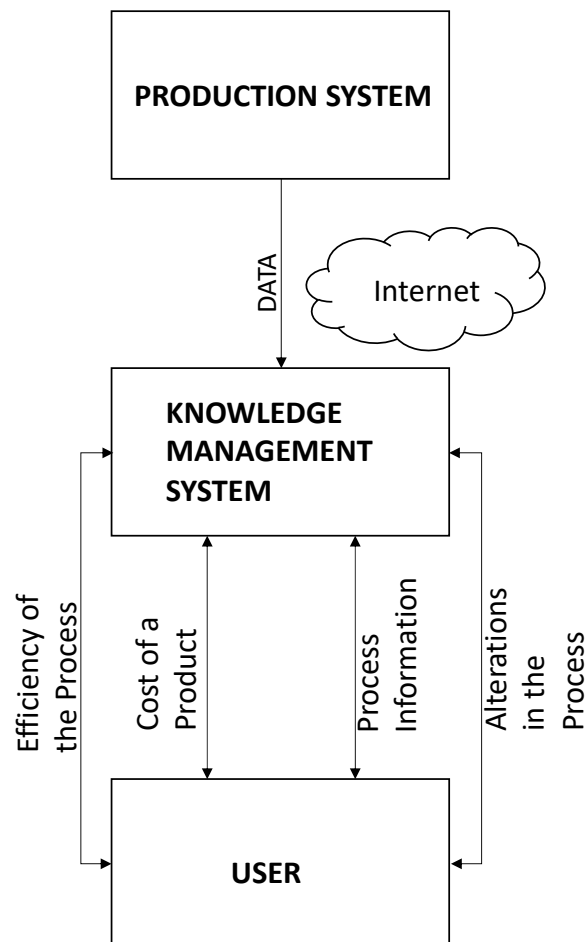
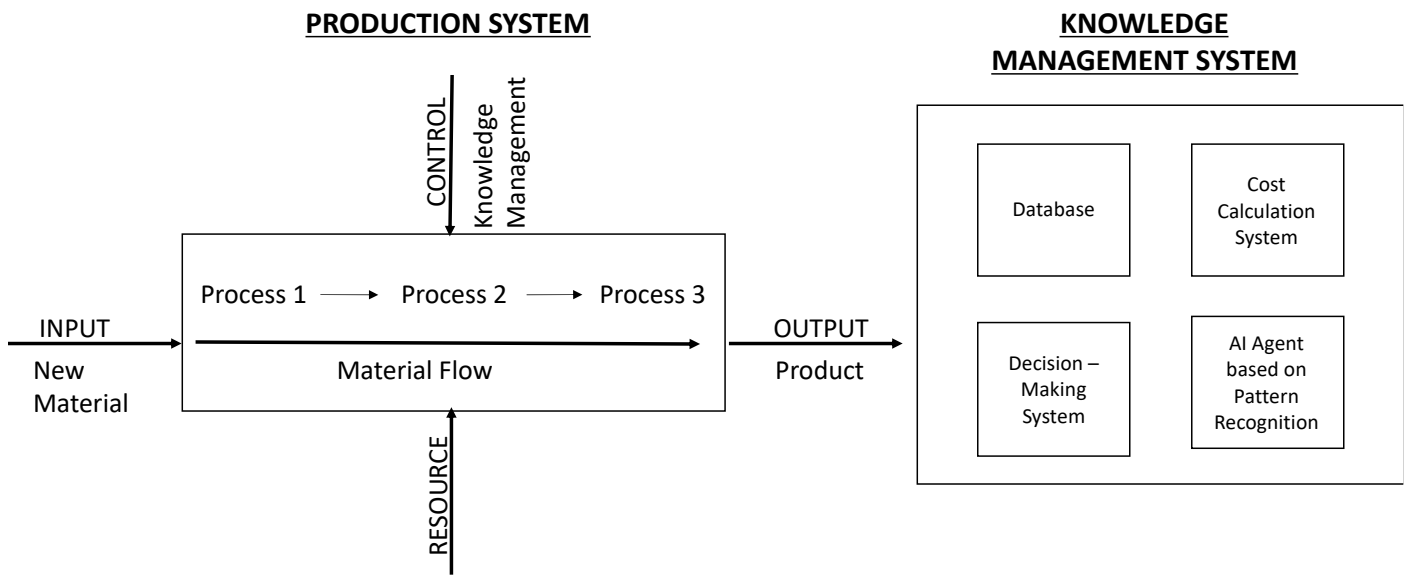


Figure 7.1: Diagram of the Structure of the Knowledge Support System

As can be seen in the diagram of the structure of knowledge support system, the flow is built on production system structure. In the creation level of the knowledge support system, the elements of a production system, which are manufacturing processes of a product, input of the system that can be the material of the product, resource of the system and control mechanism of the user of the database or the operator of the manufacturing system that can be knowledge

management of the production system, are considered. The result of the collaboration of the elements of the production system is the output is the new product. In this project, a knowledge management system framework that can be one of the control mechanisms of a production system is aimed to be established. There will be four elements inside the knowledge management system. These elements are the database that allows to the user operate the knowledge management system, cost calculation system that is for calculating or estimating unit manufacturing cost of a product, decision-making system that includes case-based reasoning system and lastly, AI agent based on pattern recognition that provides calculating utilization rate of the machine of the production system through machine learning methods by analysing the power values of the machine. Another aim of the project covers the combination of the production system and knowledge management system through internet. As a result of this combination, the user can obtain 4 outputs from the knowledge management system database: Efficiency of the production process, unit manufacturing cost of a product, process information and alterations in the process.

7.2. Creation Level of the Knowledge Support System:

After planning the knowledge support system that is explained in the designing level, the creation process of the database is begun. Before explaining the system step by step, the flowchart of whole database system and the explanation of the flowchart are given in the figure below as a summary:

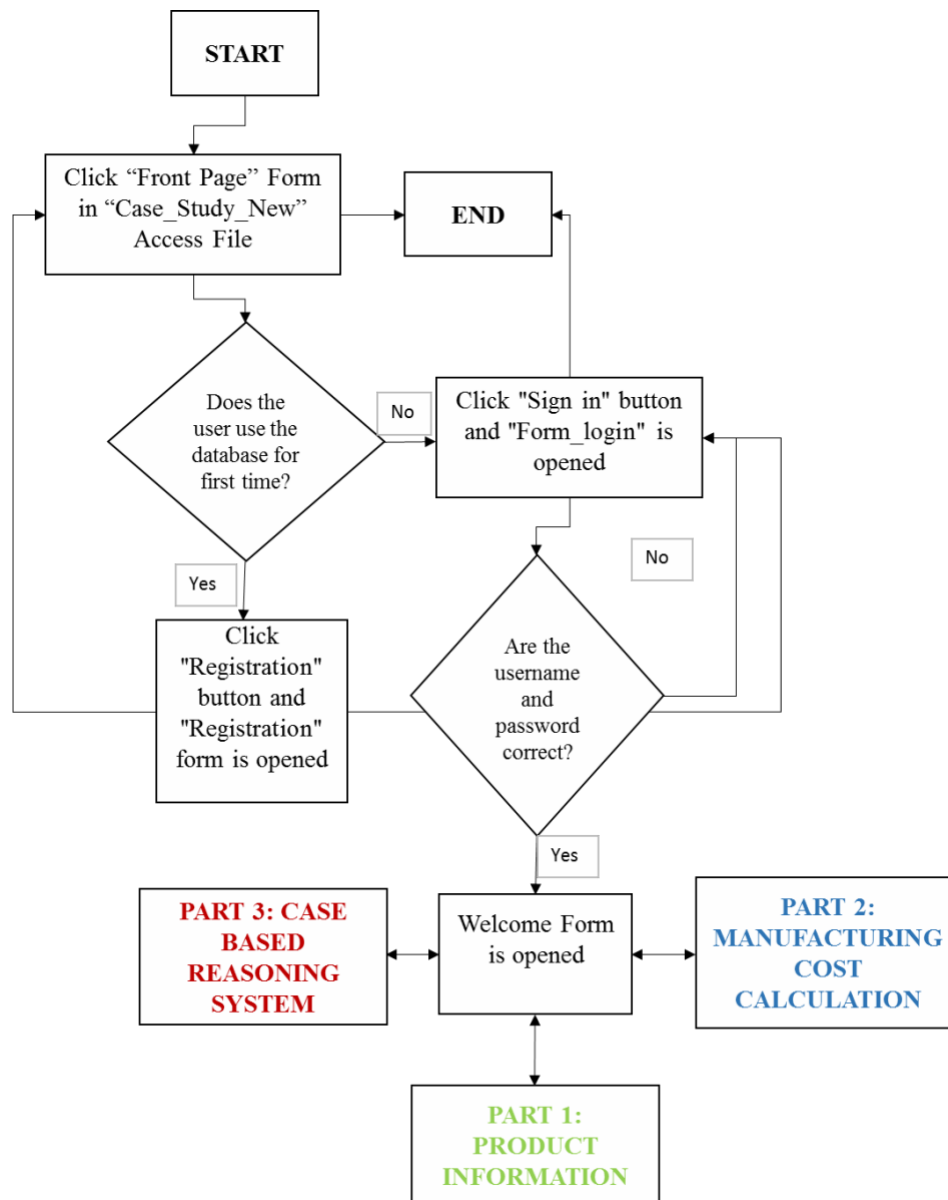


Figure 7.2.: The Flowchart of the Summary of the Whole Database System

7.2.1. Explanation of the Flowchart of the Knowledge Support System Database:

After the flowchart of the knowledge support system database is given above, it is explained in three parts:

- Beginning of the Knowledge Support System Database:

The beginning of the flowchart that includes front page, log on form, registration form and welcome form of the database is given in Figure 6.3:

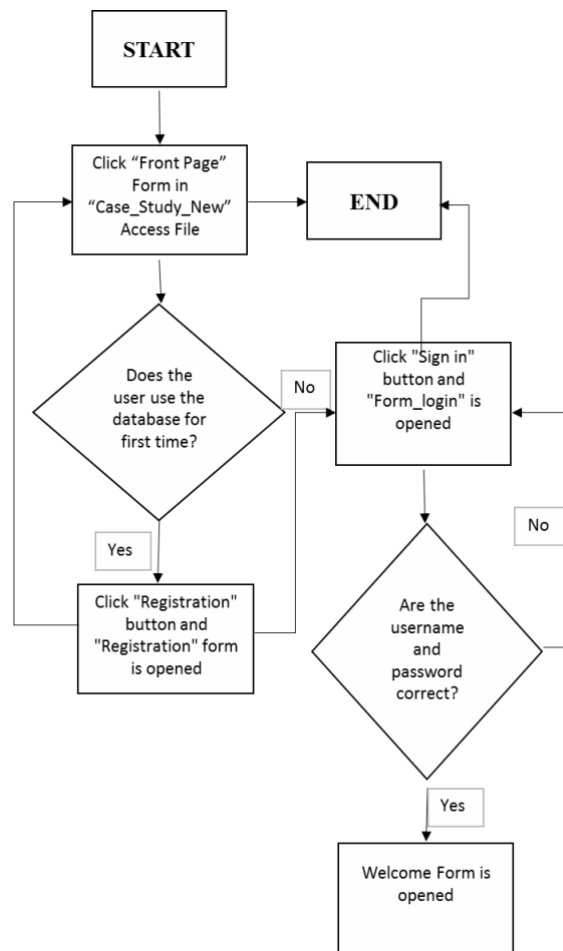


Figure 7.3.: Beginning of the Database

As can be seen in Figure 7.3, the main database, which is called “Case_Study_New”, starts with the front page of the database that contains login or registration options. If the user uses the database for the first time, it chooses the registration form in which it put his details and it chooses a username and a password. Subsequently, it turns back to the front page and it logs into the system with his username and password. On the other hand, if it is not the user’s first time, it directly clicks the “log in” button and it enters the system with its username and password. After the user has entered the system, it sees the “Welcome” form. In “Welcome” form, there are several options of the database for the user. These options are product information that includes product types and information about their manufacturing processes, cost calculation that covers the cost model for calculating unit manufacturing cost calculation of the product types and the case based reasoning system that is the system for estimating the utilization rate of the machine (for manufacture) of a product based on most similar previous products.

- Part 1 of the Flowchart of the Knowledge Support System Database: Product Information:

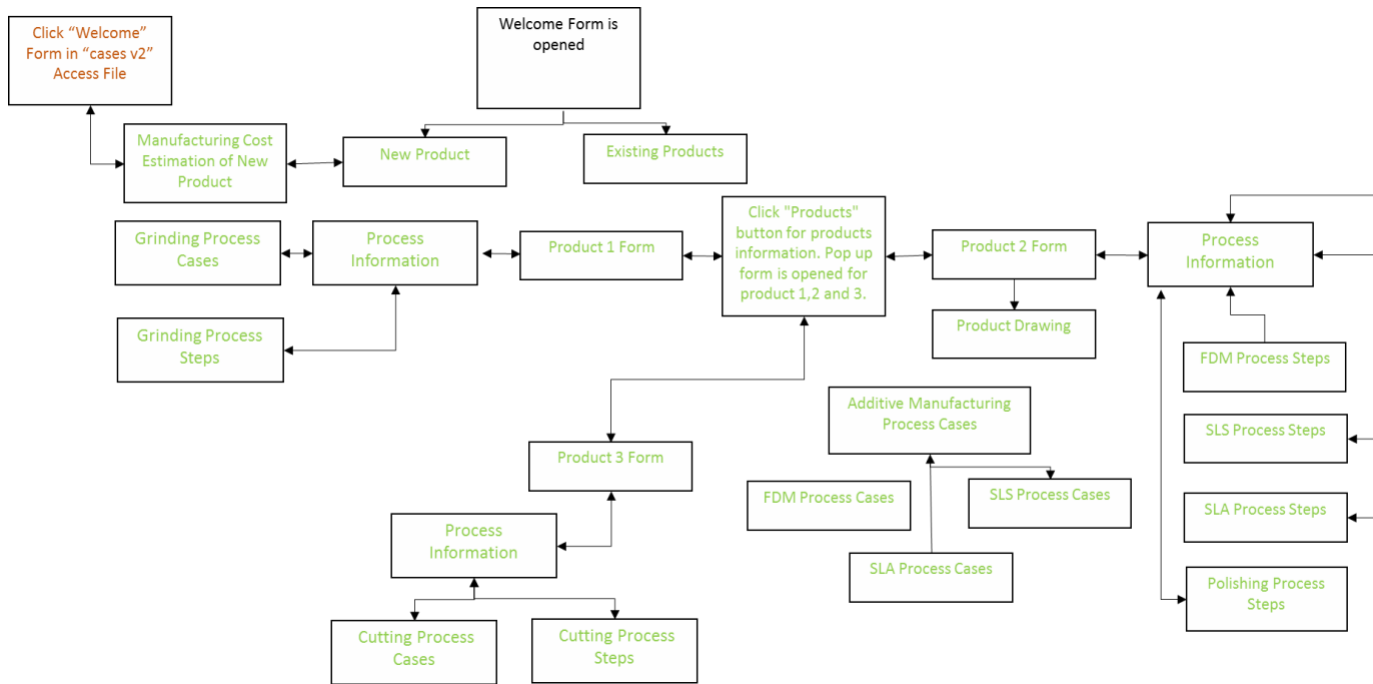


Figure 7.4.: Part 1 of the Flowchart of the Database

In “Welcome” form, there are two options: New Product and Existing Products. In “New Product” option, the user can record the design of the new product type of the facility and it can estimate the manufacturing cost of the new product type. The cost estimation button is linked to the case based reasoning system of the database in order to ensure that the user estimates the utilization rate of the machine for the new product type, for estimating the process overhead cost of the product. On the other hand, when the users click the “Product” options after the “Welcome” form, a pop-up form, which leads the user to three types of product that are put to the database as an example as existing product types, is opened. Each type of product has its individual form in the database. When the user chooses Product 1 that is manufactured via a grinding process, Product 1 form is opened and the user can check the process steps through the checklist of the grinding process steps and it can investigate the inputs and outputs of previous cases (products) that were produced through the grinding process; it can also record the input and output parameters of the current product. If the user clicks Product 2 option, it can see that Product 2 has additive manufacturing and polishing processes in its manufacturing system. Therefore, the user can check Fused Deposition Modelling (FDM), Selective Laser Sintering (SLS) and Stereolithography (SLA) process steps as additive manufacturing process steps options. Also, it can examine the input and outputs of the previous products that were produced through FDM, SLS or SLA processes and can record new products. Lastly, it can see some drawings of the products that are Product 2 type. In Product 3 option that is manufactured

via a cutting process, the user can check the cutting process steps and it can analyse input and outputs of the previous products and can record input and outputs of the new one.

- Part 2 of the Flowchart of the Knowledge Support System Database: Manufacturing Cost Calculation:

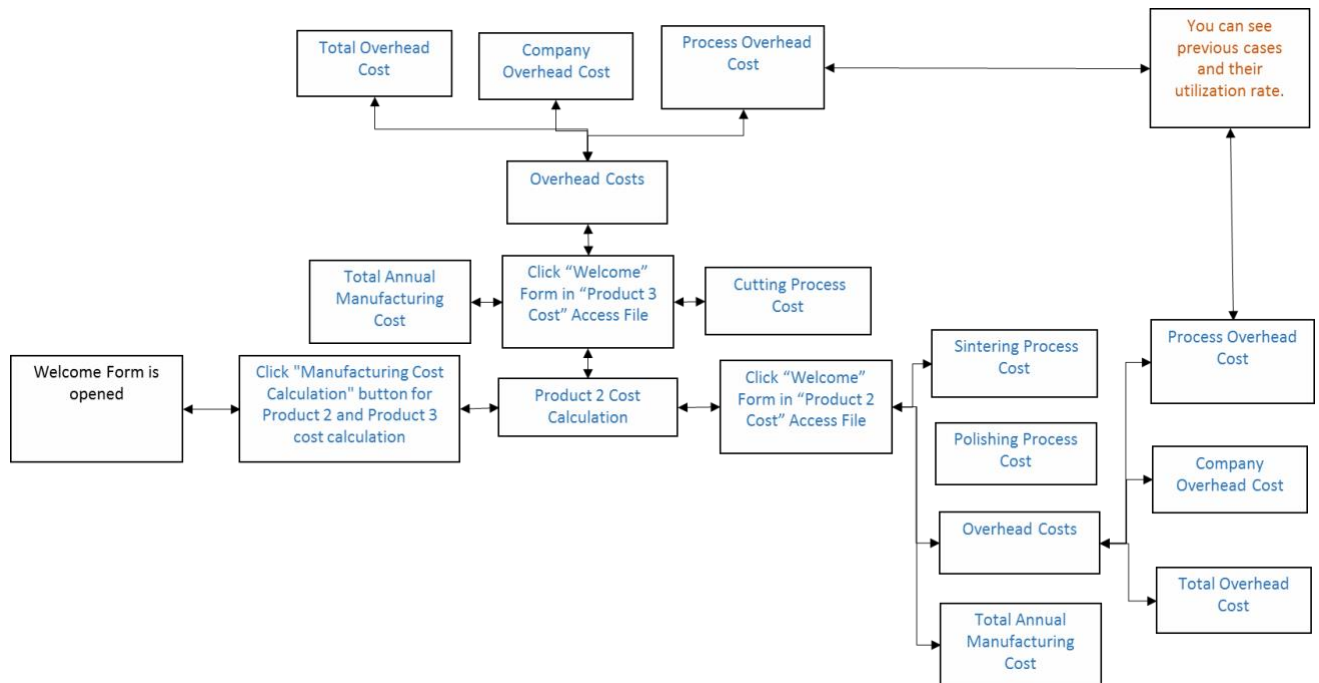


Figure 7.5.: Part 2 of the Flowchart of the Database

The second option of the “Welcome” form of the main database is the unit manufacturing cost calculation of Product 2 and Product 3. Therefore, when the user clicks the “manufacturing cost calculation” button, it can choose three options: Generic Cost Calculation that is an empty database that can be used for calculating or estimating unit manufacturing cost of a new product, Product 2 cost calculation that has selective laser sintering and polishing process in its manufacturing system and Product 3 cost calculation that includes the cutting process in its manufacturing system. The system of unit manufacturing cost calculation of Product 2 and Product 3 was created in two different databases that have the same structure with the generic cost calculation database.

Hence, the buttons of Product 2 and Product 3 are connected to these databases with a hyperlink. The database of unit manufacturing cost calculation of Product 2 and Product 3 has the same structure. The database starts with the “welcome” form that has direct process cost, overhead costs and total unit manufacturing cost options. Therefore, in Product 2, direct process cost contains direct laser sintering and direct polishing costs. On the other hand, Product 3 has direct cutting process costs. The next element of the cost model is overhead costs. In Product 2 and Product 3, overhead costs covers three selections: Total Process Overhead Cost that includes indirect cost parameters that can affect the manufacturing process indirectly, Total Company Overhead Cost that covers indirect cost parameters that do not affect the manufacturing process indirectly and lastly, Total Overhead Costs that is the sum of Process Overhead Cost and Company Overhead Cost. The last option of the “welcome” form of unit manufacturing cost calculation of the database is “Total Unit Manufacturing Cost”. This option includes the sum of Direct Process Cost and Total Overhead Cost.

- Part 3 of the Flowchart of the Knowledge Support System Database: Case Based Reasoning System:

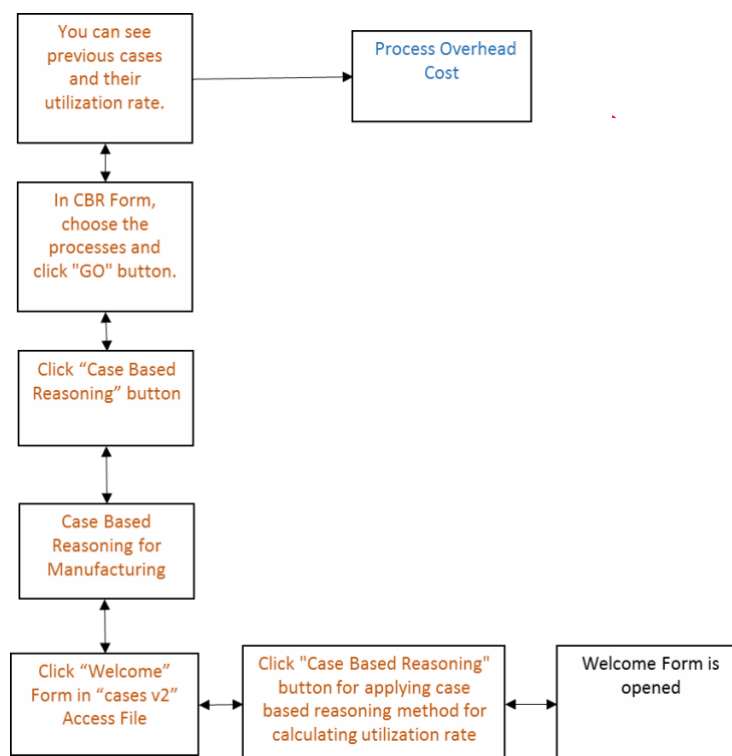


Figure 7.6.: Part 3 of the Flowchart of the Database

The last option of the “Welcome” form of the main database is estimation of the utilization rate of the machine in a manufacturing system through utilizing the case based reasoning system method. The case based reasoning system has 63 previous cases (products) for searching the most similar previous product to the current product, based on manufacturing processes of the

previous products, in order to estimate the utilization rate of the machine used for the current product in a different database that is called “cases v2”. After the user clicks “Case Based Reasoning” button, the database leads him to “cases v2” database. In this database, when the user clicks the “Case Based Reasoning for Manufacturing” option in the “Welcome” form, the button leads him to the “CBR” form, which has all the manufacturing processes of the previous products; this is opened and the user put “one” next to the processes that the current product has and it presses the “GO” button. The value of the rest of the processes remain zero. Subsequently, a form is opened that has product information, similarity value and the utilization rate of the most similar previous products that are chosen by the case based reasoning system of the database. Lastly, the user can select the most similar product that has highest similarity value and it can use its utilization rate in manufacturing cost estimation of the current product. It can use the utilization rate value in the Process Overhead Cost part of Part 3 of the flowchart of the whole database.

- The Relationship between the Parts of the Flowchart of the Database:

As can be seen in the flowchart of the knowledge support system database, there are some links between the parts of the database. These links are given in the following:

- The first link is between Part 2 and Part 3. It means that the cost estimation button of the new product type is linked to the case based reasoning system sub-database. Therefore, the user can estimate the utilization rate of the machine used for the new product and it can use this value in the process overhead cost estimation.
- The second link is between Part 1 that includes product types information and Part 2, which covers unit manufacturing cost calculation, of the database. The information form of Product 2 and Product 3 has the cost calculation button that is connected to unit manufacturing cost calculation sub-databases of Product 2 and Product 3 with a hyperlink.
 - The third and last link is between Part 2 that has manufacturing cost calculation and Part 3, which includes the case based reasoning system, of the database. The cost model of the manufacturing cost calculation can be used for manufacturing cost estimation before starting the manufacturing process. Therefore, in the process overhead cost form of Product 2 and Product 3 cost calculation sub-databases, there is a button, which is linked to the case based reasoning system sub-database, next to total process overhead cost. Therefore, the user can estimate the utilization rate of the machine based on the most similar previous product, subsequently, it can write the utilization rate value to the process overhead cost form and it can see the value of the total process overhead cost that is calculated through considering the estimated utilization rate value.

7.2.2. Detailed Explanation of Knowledge Support System Database:

After the flowchart of the knowledge support system database was explained, there will be a detailed analysis of the database part by part with some figures of the forms of the database.

- Beginning of the Knowledge Support System Database:

The database begins with “Front Page” form that has Registration and Sign-in options. The form is given in Appendix 7. In Appendix 7., if the user utilizes the database first time, it should go to registration option in order to record his details to the system and create a username and password. Subsequently, it can log in to the database with his username and password by clicking “Log in” button. If it is not his first time, it can go directly to log in form. The “Registration” form is shown in Appendix 8.

When the user enters “Registration” form, it should record his name, surname, e-mail, phone number and the company name in which it works. Later, it should decide a username and a password. After it decided his username and password, it should go back to Front Page and select Log in option to enter the Welcome form. Log in form is presented in Appendix 9.

When the user enters his username and password in the Log in form, if the username and password is correct, the database leads him to the “Welcome” form that has database options. The form is given in Appendix 10. In “Welcome” form there are several options: “Existing Products” option that has the information of existing product types in the facility, “New Product” option that is for recording the information of new product type of the facility and cost estimation of the new product type, “Manufacturing Cost Calculation” is for calculating unit manufacturing cost of the existing product types and cost estimation of new product types and lastly, “Case Based Reasoning for Utilization Rate of a Machine” option that can be utilized for estimating the utilization rate of the machine for the new product type, for estimating the process overhead cost of the product.

- Part 1 of the Knowledge Support System Database:

In this section, Part 1 of the knowledge support system database will be explained. It starts with the Welcome form that is shown in Appendix 10. Subsequently, the user can select “Existing Products” and “New Product” options. “Existing Products” option represents the information of existing product types of a facility. On the other hand, “New Product” option covers new product type of a facility that is in design level and has not been produced yet. If the user clicks “New Product” option, it comes across to the form that is given in Appendix 11. “New Product Form” contains “Design of New Product” option that has the design of new product type of the facility and is currently under development. The second option is “Manufacturing Cost

Estimation of New Product” that includes the cost estimation system of new product type and this button is linked to the Case Based Reasoning System is which Part 3 of the knowledge support system database. When the user estimates the utilization rate of the machine for the new product type based on similar previous products through the case based reasoning method, it can go to the manufacturing cost sub-database of the knowledge support system, which is part 3 of the database system, and estimate the manufacturing cost of the new product type before it starts the manufacturing process.

Another option of the “Welcome” form is “Existing Products”. In this selection, there are current product types of a facility. Therefore, when the user selects this option, a pop-up form is opened that has product types of the facility. The pop-up form is presented in Appendix 12. In Appendix 12, three types of product are given as an example for testing the database. Product 1 is manufactured through a Grinding process, Product 2, which is one of the product types of Company 1 that is one of the industrial partners of this project, is produced by applying additive manufacturing and polishing processes and , lastly, Product 3, which is one of the product types of Company 2 that is other industrial partner of this project, is manufactured by utilizing a Cutting process. The information structures of these three product types are the same. Therefore, the information structure of Product 2 will be explained as an example. When the user selects Product 2, “Product 2” form that is shown as Appendix 13 is opened. “Product 2” form includes product drawing of Product 2, process information of Product 2 that covers additive manufacturing and polishing process steps and input and output parameters of the previous cases of additive manufacturing processes and the last option is “Manufacturing Cost Calculation” that is linked to the manufacturing cost calculation sub-database that is Part 3 of the knowledge support system database. Therefore, the user can calculate the unit manufacturing cost of Product 2 after it has completed the manufacturing processes. Product Drawing form is presented in Appendix 14. As can be seen in Appendix 14, the enquiry/order form of two products of Company 1, which are a perforated plate basket and a cone filter and were produced through an additive manufacturing process, were added as examples of product drawing. The enquiry/order forms include product drawings of the products as well. Another option of “Product 2” form is process information of Product 2. The “Process Information” option leads the user to another form that is presented as Appendix 15. In Appendix 15, the form contains additive manufacturing and polishing process steps and additive manufacturing cases that have input and output parameters of previous additive manufacturing cases. The first option is “Additive Manufacturing Process Steps” that leads the user to another pop-up form that is given as Appendix 16. In Appendix 16, the database presents the user the process steps of three type of additive manufacturing processes: Fused Deposition Modelling (FDM),

Selective Laser Sintering (SLS) and Stereolithography (SLA). The process steps form of these three processes and polishing process has same structure. Therefore, the process steps form of FDM is given as Appendix 17 as an example. As can be seen in Appendix 17, a checklist is prepared that contains FDM process steps. It is designed so that the operator can do a quality control through ticking each process step after it has completed it.

The third option of the “Processes of Product 2” Form is “Additive Manufacturing Cases (Inputs and Outputs)”. When the user selects this option, a pop-up form that is shown as Appendix 18 is opened. The “Additive Manufacturing Cases” form includes the input and output parameters of three types of additive manufacturing cases: Fused Deposition Modelling (FDM) Cases, Selective Laser Sintering (SLS) Cases and Stereolithography (SLA) Cases. The form of FDM Cases is given as Appendix 19 as an example. In Appendix 19., the form has case (product) name, the type of manufacturing process, material type, software type and the drawing of the previous case (product). Company 1 only has SLS process in their facility. Therefore, the previous cases in Figure 7.18 were taken from the Liverpool John Moores University laboratory.

- Part 2 of the Knowledge Support System Database:

In this section, part 2 of the knowledge support system database that covers unit manufacturing cost calculation will be explained. When the user clicks the “Manufacturing Cost Calculation” button of the “Welcome” form of the main database, the form, which is given as Appendix 20., is opened. The “Manufacturing Cost Calculation” Form has three options: Generic Database that is a sub- database that the user can fill in and calculate unit manufacturing cost of a new product type, Product 2 and Product 3 Cost that has the same structure with Generic Database. The cost parameters of the forms of Product 2 and Product 3 Cost databases were filled with real values of Company 1 and Company 2. Therefore, unit manufacturing cost of Product 2 and Product 3 were calculated in order to test the Generic Database. The structure of the Generic Database will be explained as an example. The user starts from the “ Welcome” form of the Generic Database. The form is shown as Appendix 21. The “Welcome” form of the Generic Cost Database has several selections: “Overhead Costs” that covers manufacturing process overhead costs of the product and company overhead costs, “Direct Process Level Cost” that includes direct fixed and variable costs of the manufacturing process of the product and “Total Unit Manufacturing Cost” option that is the sum of total direct process costs and total overhead costs. When the “Overhead Cost” option is selected, a pop-up form that is presented as Appendix 22, is opened. In Appendix 22, overhead cost types and total overhead cost is presented. The Process Overhead Cost form is given as Appendix 23. The Process Overhead

Cost Form includes cost parameters that affect the manufacturing process indirectly. These cost parameters are depreciation of the machine used in the process, computer (software) cost, post and stationery, which can be used before, during or after the manufacturing process and staff training for the manufacturing process. The sum of these cost parameters gives the total process overhead cost. However, the percentage of the utilization rate of the machine should be considered because if the utilization rate of the machine increases in a process, the total overhead cost should decrease. The formula between total overhead cost and utilization rate of a machine is given in the following:

$$POLCUR = POLC * (1/URM) \quad (4-8)$$

When a user writes the percentage of utilization rate in the box next to the “Utilization Rate” parameter, the form calculates “Total Process Overhead Cost After Considered Utilization Rate” automatically. The user can calculate the utilization rate through power signals and time that come from a power meter on the machine. In this situation, the user classifies power signals according to machine status (stop, idle or working). Subsequently, based on machine status and time, the user can calculate total machine-on time (the machine can be “idle” or “working”), total “working” time and total utilization rate of the machine through the Matlab code that was created in this project. On the other hand, the user can estimate the utilization rate of the machine based on similar previous products via the Case Based Reasoning Method. In order to achieve this estimation, it clicks the “Estimate from Case Based Reasoning” button and the button leads him to Part 4 of the database system in which the Case Based Reasoning sub-database.

The next option is calculating Company Overhead Cost by clicking the “Company Overhead Cost” button in Appendix 22. Company Overhead Cost Form is shown as Appendix 24. The form “Company Overhead Cost” covers fixed and variable overhead costs of a facility that are not related to a manufacturing process. Total Company Overhead Cost is the sum of the parameters of “Company Overhead Cost.”

After calculating process and company overhead cost, total overhead cost of a facility is calculated via the form that is shown as Appendix 25. The next option of the “Welcome” form of the Generic Cost Database is “Direct Process Cost” of which the form is shown as Appendix 26. In Appendix 26., the cost parameters of the “Direct Process Cost” form are variable costs that directly affect the manufacturing process. Overhead costs are not included in this form.

The last option is calculating unit manufacturing cost of the product after calculating total overhead cost and total process cost. The form of Total Unit Manufacturing Cost is shown in

Figure Appendix 27. The values in the forms of this section are given to the cost parameters in order to test the system.

- Part 3 of the Knowledge Support System Database:

The last part of the knowledge support system database is the Case Based Reasoning System in order to estimate the utilization rate of the machine of a new product based on similar previous products in a facility.

The Case Based Reasoning System includes the following steps:

- Initially, the Case Based Reasoning Method is applied to the products that were produced through the machining process. Therefore, the information of 65 products that have the machining processes in their manufacturing system was taken from Company 1 as the case base of the Case Based Reasoning System. The information of each product contains its manufacturing process steps, total number of working and idle status and total cycle time and total working time of the machines that were utilized during the manufacturing processes. A table is generated in Microsoft Access by utilizing the information of the products.
- The table contains 65 products, 13 different machining processes, total cycle, working and idle time of the machines and total working and idle status of the machines used for the manufacture of the products. The processes are given below:
 - Measure Existing Product
 - Purchase Wedge Wire
 - Buy in Screw Fitting Section
 - Buy Filter Bags
 - Buy in Plug Section
 - Punching
 - Meshing
 - Cutting
 - Welding
 - Grinding
 - Rolling (Forming)
 - Checking
 - Packing

According to the table., if the related product has one or more processes in its manufacturing system, 1 is given to the processes. Otherwise, the processes are zero. On the other hand, the utilization rate of the machines for each product is calculated based on total working and cycle time of the machines through utilizing the formula that is given in Section 2.

- In the third step, after the case table is prepared that includes the calculation of the utilization rate of the product, the Case Based Reasoning Form, which consists of the machining processes, is generated in Microsoft Access. If the manufacturing system of the current product that is planned to be manufactured has one or more than one of these processes, the user puts 1 next to the related processes. Other processes remain as zero.
- The Case Based Reasoning form includes the processes of the current product (case) and the case table represents previous cases. It is designed so that the user can see the utilization rate of the machines of the previous product that has the most similar manufacturing processes to the current product. Therefore, it can use the value of the utilization rate in the manufacturing cost estimation of the current product. As a result, the access database should calculate a similarity parameter that includes the values of the machining processes of the Case Based Reasoning Form and the values of the machining processes of the case table. The similarity formula is:

$$(similarity)_i = 1 - [[(current\ product)_i - (previous\ product)_i] / Variable\ Range]^2 \quad (Li, 1996) \quad (7-2)$$

Where $(current\ product)_i$: Total value of the manufacturing processes of i^{th} current product. This value will be taken from the manufacturing processes in Case Based Reasoning Form.

$(previous\ product)_i$: Total value of the manufacturing processes of i^{th} previous product. This value will be taken from the manufacturing processes in the case table.

Variable Range: The gap between the values of a manufacturing process. The value of a manufacturing process in the case table or Case Based Reasoning Form is either 0 or 1. Therefore, variable range is always 1.

- A query is created in Microsoft Access and the fields (parameters) of the case table are transferred into this query. Moreover, similarity for each product in the case table is calculated according to the similarity formula through taking the value of the manufacturing processes from the Case Based Reasoning Form and the case table.
- In the last step of the Case Based Reasoning System, the user puts 1 next to the manufacturing processes that are in the manufacturing system of the current product and the access database gives similar previous products from the case table with their similarity value and utilization rate. The user can choose the most similar previous product among other products according to its similarity value in order to utilize the utilization rate of machines of the product for manufacturing cost estimation.

The Case Based Reasoning System starts with the “Welcome” form of the sub-system. The form is given in Appendix 28. In the “Welcome” form, there are three options, Case Based

Reasoning based on Design of previous products which is under development, Case Based Reasoning based on Manufacturing Process of previous products and Case Based Reasoning based on Unit Manufacturing Cost of previous products which is under development. When the user selects Case Based Reasoning based on Manufacturing Process of previous products, a manufacturing process form that is shown as Appendix 29 is opened. In Appendix 29, 13 machining processes in a facility are presented to the user. It puts 1 next to the manufacturing processes that are planned to apply to the new product. Other processes remain zero. As an example, punching, grinding and packing processes are considered. When the user clicks the “GO” button, a form appears that shows previous products that have at least one of the selected processes. The form is opened that is shown as Appendix 30. The “Previous Cases” form covers id, Enquiry ID and name of the product, machining processes, total working and machine-on time, similarity value and utilization rate of the previous product. The value of the manufacturing process is 1 if the manufacturing system of the related previous product has this manufacturing process. There are two terms in the form: Total Similarity and Utilization Rate.

- Utilization Rate: The value of the utilization rate gives the percentage of the utilization rate of the machine used for the manufacture of the previous product during the manufacturing process. The user should examine all the records of the form (all the previous products that the Case Based Reasoning system presents) and it should select the previous product that has the highest similarity value. Subsequently, it can use the utilization rate value for estimating the unit manufacturing cost of the new product. The formula of the utilization rate is given below:

$$\text{Utilization Rate of the Machine} = (\text{Total Working Time} / \text{Total Machine-on-time}) * 100 \quad (6-1)$$

CHAPTER 8: CONCLUSION and FUTURE WORKS

8.2. Discussion and Conclusion:

As a summary, this project offers a sensor based cost model which is a knowledge support system that can be used for manufacturing cost calculation, receiving information from the manufacturing processes of the product and ensuring Industry 4.0 approach in a facility via the sensors on the machines that can provide data from sensors via the internet for the cost model.

The conclusion of each chapter of the thesis is given in following:

- Chapter 2 – Literature Review: In conclusion, significant research was carried out on the literature in order to generate the background for the project and presented in Chapter 2. Chapter 2 gives examples about production process planning and control methods from the literature and explains the methods of cost estimation of the production process as well as knowledge management systems, which include sub-topics such as artificial intelligence, machine learning, data mining and the Case Based Reasoning Method. Therefore, during searching the literature, it became apparent that an example of a cost model, that calculates the utilization rate of a machine during the manufacturing process as an efficiency indicator and is sensor-based at the same time, does not exist. In order to fill this gap in the literature, the bottom-up cost estimation technique will be utilized during generation of the sensor-based cost model and Support Vector Machine, Multi-Class Support Vector Machine and Decision Tree methods will be used as machine learning methods for providing data to the cost model in this project. The limitations of other machine learning methods for this project are given below:
 - Hidden Markov Model may not tolerate the large amount of data of the power meter that was assembled on the milling machine of Liverpool John Moores University.
 - Bayesian Networks may not be useful for continuous power data of this project.
 - In the Artificial Neural Network Algorithm, in order to obtain an effective result from this method, a large amount of input data should be gathered and it requires high complexity in the problem (Comert & Kocamaz, 2017). In this project, a method that gives an effective result in the shortest time is more suitable considering time management.
 - Multiple Linear Regression Analysis is not precise enough to analyse a power-time graph of a machine and finding their cycle time.
 - On the other hand, as a data mining operation, classification will be applied to power

data of cutting, laser sintering and milling machine in order to assign machine status to power values via Matlab Software.

- Chapter 4 - Generic Cost Model Development: In Chapter 4, the creation process of the generic cost model of the knowledge support system database that is for manufacturing cost optimization and calculating unit manufacturing cost of a product was explained by presenting the cost parameters of the model and sub-formulas of it.

The generic cost model of the knowledge support system database contains two novelties of this project:

- The utilization rate of the machine was added to the Process Overhead Level Cost of the generic cost model as a cost parameter. Hence, a parameter that gives a strong idea about the efficiency of a manufacturing process is considered during calculating unit manufacturing cost of a product. Besides, this parameter converts the cost model from traditional to dynamic cost model. It means that the cost model is able to indicate the increase or decrease of the efficiency of the manufacturing process.
 - The cost of the electricity energy consumption of a manufacturing process is considered and added to the Direct Process Level Cost of the generic cost model through calculating the electricity energy consumption of a machine in a manufacturing process via the data of the power meter that is assembled on the machine.
- Chapter 5 - Sensor Based Cost Modelling and Relevant Data Acquisition: This chapter presented three experiment, which were done for calculating electrical energy cost of the machine of the manufacturing process and observing the process by analysing the power levels of the machine, were presented. The cost of the electrical energy of the machine was calculated during the data that was come from the power meter on the machine and it shows the cost of the electricity energy consumption of the machine during the manufacturing process only. In order to calculate the electricity energy cost of the experiments, the value of active power data of the machine of the manufacturing process and total experiment time were obtained from the power meter. On the other hand, via the active power – time graph of the machine that can be taken from the power meter, total working time of the machine, which indicates the time that the machine actually works, can be calculated. Therefore, the total utilization rate of the machine can be calculated with the total experiment time and total working time of the machine during the experiment.

- Benefits of the Experiments for Sensor Based Cost Model:

The aim of these three experiments is obtaining power values of the machine during the manufacturing process and calculating active energy and electricity cost of the machine. This electricity cost will be the cost of electricity that was used during the manufacturing process only. Electricity consumption of the facility apart from the manufacturing process is not included. Therefore, total electricity cost of the manufacturing process that was calculated during these experiments will be used as a direct variable cost in the total direct process cost of the generic cost model.

In order to calculate the electricity energy cost of the experiments, the value of active power data of the machine of the manufacturing process and total experiment time are obtained from the power meter. On the other hand, via the active power – time graph of the machine that can be taken from the power meter, total working time of the machine, which indicates the time that the machine actually works, can be calculated. Therefore, the total utilization rate of the machine can be calculated with the total experiment time and total working time of the machine during the experiment.

- Chapter 6 - Assessment of The Manufacturing Process Efficiency: In this chapter, utilization rate of the machines, which were utilized during the experiments in Chapter 5, were calculated through the utilization rate formula that was given in this chapter. In order to calculate the utilization rate of the machine effectively, the power values of these three machines were classified according to machine status through using decision tree and support vector machine algorithm. Subsequently, through analysing the classified power values and their status (stop, idle or working), total experiment time and total working time of these three machines were calculated by Matlab code. These two variables were used in the utilization rate formula.

The utilization rate of the machine can be a strong indicator about the efficiency of the manufacturing process and it can also be a useful parameter for calculating manufacturing process overhead cost accurately. Therefore, it was decided to use the utilization rate of the machine parameter in the manufacturing cost calculation model.

- Chapter 7 - Framework of The Knowledge Support System: Chapter 7 of the thesis explained the structure of the knowledge support system and the parts of the knowledge support system database. This knowledge support system database is designed to calculate unit manufacturing cost of a product in small or medium size companies and

facilitate optimization of a manufacturing processes.

As a general conclusion, the knowledge support system of this project can help SMEs in several ways:

- ✓ It can calculate unit manufacturing cost of a product through a novel generic sensor-based cost model for SMEs in order to add flexibility to the knowledge support system. The novelty of the cost model comes from the one of its cost parameters that is an indicator of the efficiency of the manufacturing process of a product. This cost parameter is the utilization rate of the machine of the manufacturing process.
- ✓ The knowledge support system database can give information about steps of the manufacturing process of the current product and previous products (cases) that were produced via the same manufacturing process.
- ✓ The knowledge support system can estimate the manufacturing cost of the new product as well, before the beginning of the manufacturing process. In order to estimate the utilization rate of the machine of the new product, the Case Based Reasoning (CBR) System, which chooses the previous product that has the most similar manufacturing system to the new product among the previous cases of the facility and gives the utilization rate of the machine of the previous product, was established inside the database.
- ✓ A base for Industry 4.0 approach can be established in a manufacturing system through the sensor-based cost model.
- ✓ The sensors that will be applied on machines in a facility show the efficiency of production and the power level of the machine, which is one of the outcomes of the sensors, indicates potential quality problems and productivity problems of the manufacturing process.

After presenting potential benefits of the knowledge support system that is offered in this thesis, the novelties of the project are given in the following:

- The utilization rate of a machine during a manufacturing process will be calculated through the data, which are utilized inside of a suitable machine learning technique for the calculations of the utilization rate, from the sensor on the machine. Receiving the value of the utilization rate of the machine has two benefits:
 - ✓ The efficiency of a manufacturing process can be seen and measured effectively through the value of the utilization rate of the machine.
 - ✓ Overhead costs of the manufacturing process can be calculated more effectively

by using the value of the utilization rate of the machine.

- The cost of the electricity energy consumption of a manufacturing process is considered and added to the Direct Process Level Cost of the generic cost model through calculating the electricity energy consumption of a machine of a manufacturing process via the data of the power meter that is assembled on the machine.

8.2. Future Works:

The knowledge support system database requires some improvements that can be done as future works in order to create a fully automatic knowledge support system for SMEs. These improvements are presented in the following:

- In the future, the connection between Matlab software and Microsoft Access database will be created so that the user can enter the power values of the machine that come from a sensor in Matlab Software, run the Matlab code and obtain the value of the utilization rate of the machine of the product. Subsequently, the user can calculate the unit manufacturing cost of the existing product through utilizing the knowledge support system database. Therefore, the self-learning capability and flexibility of the system can be improved.
- The knowledge support system will be tested in industry in order to obtain the feedback of different small or medium companies for carrying the knowledge support system forward.

References

- Abellan-Nebot, J. V., & Subirón, F. R. (2010). A review of machining monitoring systems based on artificial intelligence process models. *Int J Adv Manuf Technol*, 47, 237–257.
- Aggarwal, C. C. (2015). An Introduction to Data Classification. In C. C. Aggarwal (Ed.), *Data Classification: Algorithms and Applications* (pp. 2-31). Boca Raton: CRC Press.
- Ahuja, Y., & Yadav, S. K. (2012). Multiclass Classification and Support Vector Machine. *Global Journal of Computer Science and Technology Interdisciplinary*, 12(11), 15-19.
- Alavi, M., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, 25(1), 107-136.
- Angelis, L., & Stamelos, I. (2000). A Simulation Tool for Efficient Analogy Based Cost Estimation. *Empirical Software Engineering*, 5, 35–68.
- Askarany, D., Yazdifar, H., & Askary, S. (2010). Supply chain management, activity-based costing and organisational factors. *International Journal of Production Economics*, 127(2), 238-248.
- Auria, L., & Moro, R. A. (2008). Support Vector Machines (SVM) as a technique for solvency analysis. Berlin: ISSN.
- Awad, M., & Khanna, R. (2015). Support Vector Machines for Classification: Theories, Concepts, and Applications for Engineers and System Designers. In *Efficient Learning Machines* (pp. 39-66). Apress.
- Baldinger, M., Levy, G., Schönsleben, P., & Wandfluh, M. (2016). Additive manufacturing cost estimation for buy scenarios. *Rapid Prototyping Journal*, 22(6), 871-877.
- Bhatt, G. D. (2001). Knowledge management in organizations: examining the interaction between technologies, techniques and people. *Journal of Knowledge Management*, 5(1), 68-75.
- Blocher, E., Wong, B., & McKittrick, C. T. (2002). Making bottom-up ABC work at Reichhold Inc. (Activity-Based Costing). *Strategic Finance*, 83(10), 51.
- Boral, S., Chaturvedi, K. S., & Naikan, V. (2019). A case-based reasoning system for fault detection and isolation: a case study on complex gearboxes. *Journal of Quality in Maintenance Engineering*, 213-235.
- Brettel, M., Friederichsen, N., Keller, M., & Rosenberg, M. (2014). How Virtualization, Decentralization and Network Building Change the Manufacturing Landscape: An Industry 4.0 Perspective . *International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering*, 8(1), 37-44.
- Camargo, M., Rabenasolo, B., & Jolly-Desodt, A.-M. (2003). Application of the Parametric Cost Estimation in the Textile Supply Chain . *Journal of Textile and Apparel, Technology and Management*, 3(1), 1-12.
- Carolus, J. F., Hanley, N., Olsen, S. B., & Pedersen, S. M. (2018). A bottom-up approach to environmental Cost-Benefit Analysis. *Environment and Development Economics*, 3, 1-25.
- Chan, C. W., & Huang, G. H. (2003). Artificial intelligence for management and control of pollution minimization and mitigation processes. *Engineering Applications of Artificial Intelligence*, 16, 75-90.
- Chou, J.-S., Tai, Y., & Chang, L.-J. (2010). Predicting the development cost of TFT-LCD manufacturing equipment with artificial intelligence models. *International Journal of Production Economics*, 128(1), 339-350.

- Chu, W., Keerthi, S. S., & Ong, C. J. (2002). A GENERAL FORMULATION FOR SUPPORT VECTOR MACHINES. Singapore: IEEE.
- Clements, P. C., McGregor, J. D., & Cohen, S. G. (2005). *The Structured Intuitive Model for Product Line Economics (SIMPLE)*. Virginia.
- Comert, Z., & Kocamaz, A. F. (2017). A Study of Artificial Neural Network Training Algorithms for Classification of Cardiotocography Signals. *Journal of Science and Technology*, 7(2), 93-103.
- Cunnam, L., Sinanovic, E., Ramma, L., Foster, N., Berrie, L., Stevens, W., . . . Vassall, A. (2016). Using Top-down and Bottom-up Costing Approaches in LMICs: The Case for Using Both to Assess the Incremental Costs of New Technologies at Scale. *Health Econ.*, 25, 53–66.
- Dai, W., & Ji, W. (2014). A MapReduce Implementation of C4.5 Decision Tree Algorithm. *International Journal of Database Theory and Application*, 7(1), 49-60.
- Datta, P. P., & Roy, R. (2010). Cost modelling techniques for availability type service support contracts: A literature review and empirical study. *CIRP Journal of Manufacturing Science and Technology*, 3(2), 142-157.
- Degirmenci, A. (2014). *Introduction to Hidden Markov Models*. Retrieved July 22, 2020, from https://scholar.harvard.edu/files/adeqirmenci/files/hmm_adeqirmenci_2014.pdf
- Deng, S., & Yeh, T.-H. (2011). Using least squares support vector machines for the airframe structures manufacturing cost estimation. *International Journal of Production Economics*, 131(2), 701-708.
- Dingsøyr, T., Rolland, K.-H., & Jaccheri, M. L. (2004). *The Benefits and Limitations of Knowledge Management in Global Software Development*. Edinburgh: ICSE.
- Domingos, P. (2012). A Few Useful Things to Know about Machine Learning. *Communications of the ACM*, 55(10).
- Duran, O., Maciel, J., & Rodriguez, N. (2012). Comparisons between two types of neural networks for manufacturing cost estimation of piping elements. *Expert Systems with Applications*, 39(9), 7788-7795.
- Edosio, U. (2014). *Knowledge Management Concept*. Bradford.
- EDUCBA. (2019). *Data Mining vs Machine Learning*. Retrieved August 15, 2019, from <https://www.educba.com/data-mining-vs-machine-learning/>
- Fan, S. (2018, May 7). *Shuzhan Fan*. Retrieved January 13, 2020, from <https://shuzhanfan.github.io/2018/05/understanding-mathematics-behind-support-vector-machines/>
- Fournier-Viger, P., Lin, J. C.-W., Kiran, R. U., Koh, Y. S., & Thomas, R. (2017). A Survey of Sequential Pattern Mining. *Data Science and Pattern Recognition Ubiquitous International*, 1(1), 54-77.
- Franc, V., & Hlavac, V. (2002). *Multi-class support vector machine*. Los Alamitos: IEEE Computer Society.
- Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. *International Journal of Production Economics*, 15-26.
- Geetha, M. (2015). Application of Classification Technique in Data Mining for Agricultural Land. *International Journal of Advanced Research in Computer and Communication Engineering*, 4(3), 352-355.

- Gholami, R., & Fakhari, N. (2017). Chapter 27 - Support Vector Machine: Principles, Parameters, and Applications. In *Handbook of Neural Computation* (pp. 515-535). Academic Press.
- Girard, J., & Girard, J. (2015). Defining knowledge management: Toward an applied compendium . *Online Journal of Applied Knowledge Management*, 3(1), 1-20.
- Haddadi, M. H., & Seyednezhad, M. J. (2015). Comparative Study of Traditional and Activity-Based Costing in Forging Companies of Iran Tractor. *International Journal of Management Sciences and Business Research*, 4(3), 1-10.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Waltham: Elsevier.
- Hi, W., Aas, E., Roman, E., & Smith, A. (2015). Comparison of different costing methods. Milan: ISPOR.
- Holmes, W., & Huckvale, M. (1994). WHY HAVE HMMS BEEN SO SUCCESSFUL FOR AUTOMATIC SPEECH RECOGNITION AND HOW MIGHT THEY BE IMPROVED? *Speech Hearing and Language - Work in Progress*, 8, 1-11.
- HUANG, S., CAI, N., PACHECO, P. P., NARANDES, S., WANG, Y., & XU, W. (2017). Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. *Cancer Genomics Proteomics*, 15(1), 41-51.
- Ishwaran, H., & Rao, J. (2009). Decision tree: introduction. In M. Kattan (Ed.), *Encyclopedia of Medical Decision Making* (pp. 323-328). California: Sage Inc.
- Jaccard, M., Murphy, R., & Rivers, N. (2004). Energy–environment policy modeling of endogenous technological change with personal vehicles: combining top-down and bottom-up methods. *Ecological Economics*, 51(1-2), 31-46.
- Javid, M., Hadian, M., Ghaderi, H., Ghaffari, S., & Salehi, M. (2016). Application of the Activity-Based Costing Method for Unit-Cost Calculation in a Hospital . *Global Journal of Health Science*, 8(1), 165-172.
- Jeon, J. (2015). The Strengths and Limitations of the Statistical Modeling of Complex Social Phenomenon: Focusing on SEM, Path Analysis, or Multiple Regression Models . *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering* , 9(5), 1634-1642.
- Joshi, A. V. (2019). Decision Trees. In *Machine Learning and Artificial Intelligence* (pp. 53-63). Switzerland: Springer, Cham.
- Khosravani, M. R., & Nasiri, S. (2019). Injection molding manufacturing process: review of case-based reasoning applications. *Journal of Intelligent Manufacturing*, 847–864.
- Kim, G.-H., An, S.-H., & Kang, K.-I. (2004). Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Building and Environment*, 39(10), 1235-1242.
- Kim, J. M., & Park, C.-K. (2006). Top-down Budgeting as a Tool for Central Resource Management. *OECD JOURNAL ON BUDGETING*, 6(1), 88-126.
- Kim, K. J., & Kim, K. (2010). Preliminary Cost Estimation Model Using Case-Based Reasoning and Genetic Algorithms . *Journal of Computing in Civil Engineering*, 24(6), 499-505.
- Kolodner, J. (2014). *Case-Based Reasoning*. San Mateo: Morgan Kaufmann Publishers.

- Kowsari, F. (2013). Changing in Costing Models from Traditional to Performance Focused Activity Based Costing (PFABC) . *European Online Journal of Natural and Social Sciences*, 3(2), 2497-2508.
- Leask, M., Lee, C., Milner, T., Norton, M., & Rathod, D. (2008). *Knowledge management tools and techniques: helping you access the right knowledge at the right time* . London: Improvement and Development Agency .
- Lee, J., Kao, H.-A., & Yang, S. (2014). Service innovation and smart analytics for Industry 4.0 and big data environment . *Procedia* , 16, 3 – 8 .
- Leung, J. (2014). *MUMT621-Assignment 3*. Retrieved February 17, 2017, from http://www.music.mcgill.ca/~jason/mumt621/621Presentation3_HMMs.pdf
- Li, Y. (1996). *Intelligent Selection of Grinding Conditions*. Liverpool: Liverpool John Moores University.
- Liao, S.-h. (2003). Knowledge management technologies and applications—literature review from 1995 to 2002. *Expert Systems with Applications*, 25, 155–164.
- Lihong, Q., & Shengping, L. (2012). An improved genetic algorithm for integrated process planning and scheduling. *Int J Adv Manuf Technol*, 58, 727-740.
- Loyer, J.-L., Henriques, E., Fontul, M., & Wiseall, S. (2016). Comparison of Machine Learning methods applied to the estimation of manufacturing cost of jet engine components. *Int. J. Production Economics*, 178, 109-119.
- Lozano-Pérez, T., & Kaelbling, L. (2002). *MIT Open Courseware*. Retrieved January 15, 2017, from <https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-825-techniques-in-artificial-intelligence-sma-5504-fall-2002/lecture-notes/Lecture1Final.pdf>
- Lu, Y., Yang, Z., Eddy, D., & Krishnamurty, S. (2018). Self-Improving Additive Manufacturing Knowledge Management. *ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. 26–29). Quebec: ASME.
- Luger, G. F. (2005). Artificial Intelligence-A Summary. In *Artificial Intelligence-Structures and Strategies For Complex Problem Solving* (p. 30). Essex: Pearson Education Limited.
- MABROUKEH, N. R., & EZEIFE, C. I. (2010). A Taxonomy of Sequential Pattern Mining Algorithms. *ACM Computing Surveys*, 43(1), 3:1-3:41.
- Marjanović, D., Štorga, M. Š., Bojčetić, N., & Pavković, N. (2018). Application of Knowledge Management System to Injection Mold Design and Manufacturing in Small Enterprises. *DESIGN 2018 15th International Design Conference*, (pp. 1733-1744). Dubrovnik.
- MathWorks. (n.d.). *MATLAB for Machine Learning*. Retrieved April 20, 2020, from <https://www.mathworks.com/help/stats/support-vector-machines-for-binary-classification.html>
- Mevelléc, P. (2009). *Cost Systems Design* (1st ed.). London: Palgrave Macmillan UK.
- Mijwel, M. M. (2018, January 1). *Artificial Neural Networks Advantages and Disadvantages*. Retrieved July 24, 2020, from https://www.researchgate.net/publication/323665827_Artificial_Neural_Networks_Advantages_and_Disadvantages
- Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247-259.

- NASA. (2015). *NASA Cost Estimating Handbook, Version 4.0*. Washington: National Aeronautics and Space Administration (NASA).
- Neha, D., & Vidyavathi, B. (2015). A Survey on Applications of Data Mining using Clustering Techniques. *International Journal of Computer Applications*, 126(2), 7-12.
- Olaniyi, S. A., Adewole, K. S., & Jimoh, R. (2011). Stock Trend Prediction Using Regression Analysis –A Data Mining Approach. *ARNP Journal of Systems and Software*, 1(4), 154-157.
- Pandis, N. (2016). Multiple linear regression analysis. *Am J Orthod Dentofacial Orthop*, 149, 581.
- Pimenov, D., Bustillo, A., & Mikolajczyk, T. (2018). Artificial intelligence for automatic prediction of required surface roughness by monitoring wear on face mill teeth. *Journal of Intelligent Manufacturing*, 29, 1045–1061.
- Priyam, A. A., Gupta, R., Rathee, A., & Srivastava, S. (2013). Comparative Analysis of Decision Tree Classification Algorithms. *International Journal of Current Engineering and Technology*, 3(2), 334-337.
- Rasiah, D. (2011). Why activity based costing (ABC) is still tagging behind the traditional costing in Malaysia? *Journal of Applied Finance & Banking*, 1(1), 83-106.
- Ren, L., Sparks, T., Ruan, J., & Liou, F. (2010). Integrated Process Planning for a Multiaxis Hybrid Manufacturing System. *J. Manuf. Sci. Eng.*, 132(2).
- Rokach, L. (2016). Decision forest: Twenty years of research. *Information Fusion*, 27, 111-125.
- Rokach, L., & Maimon, O. (2005). Decision Trees. In L. Rokach, & O. Maimon (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 165-192). Springer.
- Ruth, J. (2006). *World Bank*. Retrieved February 26, 2020, from http://siteresources.worldbank.org/INTPRS1/Resources/383606-1201883571938/Cost_Definitions_and_Methods.pdf
- Schott, P., Lederer, M., Eigner, I., & Bodendorf, F. (2020). Case-based reasoning for complexity management in Industry 4.0 . *Journal of Manufacturing Technology Management*.
- Schulze, M., Seuring, S., & Ewering, C. (2012). Applying activity-based costing in a supply chain environment. *International Journal of Production Economics*, 135(2), 716-725.
- Sharma, V., & Kumar, R. (2017). Dating of ballpoint pen writing inks via spectroscopic and multiple linear regression analysis: A novel approach. *Microchemical Journal*, 134, 104-113.
- Shetelig, H. (2013). *Shipbuilding Cost Estimation-Parametric Approach* . Trondheim.
- SINGH, N., RAW, R. S., & R.K., C. (2012). DATA MINING WITH REGRESSION TECHNIQUE. *Journal of Information Systems and Communication*, 3(1), 199-202.
- Smart, C., & Culver, G. (2009). An Analytical Framework for Cost Estimating Relationship Development . ICEAA.
- Sozontov, A., Ivanova, M., & Gibadullin, A. (2019). Implementation of artificial intelligence in the electric power industry. Irkutsk: IEEE.
- Strandhagen, J. W., Alfnes, E., Strandhagen, J. O., & Vallandingham, L. R. (2017). The fit of Industry 4.0 applications in manufacturing logistics: a multiple case study. *Advances in Manufacturing*, 344–358.
- Su, Y., Yang, S., Liu, K., Hua, K., & Yao, Q. (2019). Developing A Case-Based Reasoning Model for Safety Accident Pre-Control and Decision Making in the Construction Industry. *International Journal of Environmental Research and Public Health*, 16(1511), 1-20.

- Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2005). Classification: Basic Concepts, Decision Trees, and Model Evaluation. In *Introduction to Data Mining* (pp. 145-203). Pearson.
- Thames, L., & Schaefer, D. (2016). Software-Defined Cloud Manufacturing for Industry 4.0. *Procedia CIRP*, 52, 12- 17.
- Tokucoglu, H., Chen, X., El Rhalibi, A., & Opoz, T. (2019). Sensor Based Cost Modelling for a Knowledge Support System Development. 2019 25th International Conference on Automation and Computing (ICAC) (pp. 1-6). Lancaster: IEEE.
- Toth, C. A. (2006). *A Bottoms-Up Approach to Cost Estimation Using Parametric Inputs*. Athens.
- Twongyirwe, T. M., & Jude, L. (2018). Evaluation of a Knowledge Management Dashboard for Manufacturing SMEs in Resource Constrained Areas. *ICICKM 2018 International Conference on Intellectual Capital Knowledge Management & Organisational Learning* (pp. 397-407). Cape Town: Academic Conferences and Publishing International Limited.
- Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling*, 203, 312-318.
- Vural, V., & Dy, J. G. (2004). A Hierarchical Method for Multi-Class Support Vector Machines. Banff.
- Wang, L. (2013). Machine availability monitoring and machining process planning towards Cloud manufacturing. *CIRP Journal of Manufacturing Science and Technology*, 6, 263–273.
- Wang, X. V., Givehchi, M., & Wang, L. (2017). Manufacturing system on the cloud: a case study on cloud-based process planning. *Procedia CIRP*, 63, 39 – 45 .
- Westreich, D., Lessler, J., & Jonsson Funk, M. (2010). Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. *Journal of Clinical Epidemiology*, 63(8), 826-833.
- Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning tools and Techniques* (3rd ed.). Burlington: Elsevier.
- Wright, A. P., Wright, A. T., McCoy, A. B., & Sittig, D. F. (2015). The use of sequential pattern mining to predict next prescribed medications. *Journal of Biomedical Informatics*, 53, 73-80.
- Wu, J. (2012). *Advances in K-means Clustering: A Data Mining Thinking*. Berlin: Springer.
- Xu, J., Liu, X., Huo, Z., Deng, C., Nie, F., & Huang, H. (2017). Multi-Class Support Vector Machine via Maximizing Multi-Class Margins. Melbourne.
- Zheng, P., Wang, H., Sang, Z., Zhong, R. Y., Liu, Y., Liu, C., . . . Xu, X. (2018). Smart manufacturing systems for Industry 4.0: Conceptual framework, scenarios, and future perspectives. *Frontiers of Mechanical Engineering*, 137–150.
- Zhou, Y., Fenton, N., & Neil, M. (2014). Bayesian network approach to multinomial parameter learning using data and expert judgments. *International Journal of Approximate Reasoning*, 55, 1252-1268.

APPENDIX 1: BINARY SUPPORT VECTOR MACHINE CODE of COMPANY 1 **DATABASE**

```
filename = 'c_datamatlab.xlsx';  
[Anum, Btxt] = xlsread(filename, 'shifting', 'a71:f1310');  
n0 = 1;  
n=1238;  
X = Anum (n0:n,1:5);  
Y = Btxt (n0:n);
```

```
L1= false(n,1);  
L2= false(n,1);  
L3= false(n,1);
```

```
Y1= zeros(n,1);
```

```
for i=1:n  
    L1(i)= strcmp(Y(i), 'idle',4);  
    L2(i)= strcmp(Y(i), 'working',7);  
    L3(i)=  
    strcmp(Y(i), 'stop',4); if L3(i)  
        Y1(i)= 0;  
    else Y1(i)=  
    1;  
end  
end
```

```
Mdl1= fitcsvm(X,Y1);
```

```
Y2 = Y1;
```

```
for i=1:n  
    L1(i)= strcmp(Y(i), 'idle',4);  
    L2(i)= strcmp(Y(i), 'working',7);  
    L3(i)= strcmp(Y(i), 'stop',4);  
    if L1(i)  
        Y2(i)= 1;  
    else  
        Y2(i)= 2;  
    end  
end
```

```
Mdl2 = fitcsvm(X,Y2);
```

```
A = Anum (n0:n);
```

```
for i=1:n
    if (Y1(i) == 1) && (Y2(i) ==
        1) A(i) = 1;
    elseif (Y1(i) == 1) && (Y2(i)
        == 2) A(i) = 2;
    else
        A(i) = 0;
    end
end
```

```
Mdl3 = fitcsvm(X,A);
```

```
lStd = resubLoss(Mdl3);
```

```
Mdl4 = fitcsvm(X,A,'OptimizeHyperparameters','auto',...
'HyperparameterOptimizationOptions',struct('AcquisitionFunctionName',...
'expected-improvement-plus'));
```

```
filename = 'c_datamatlab.xlsx';
```

```
[Anum,Craw] = xlsread(filename,'shifting','a2:e68');
```

```
n0 = 1;
```

```
n = 66;
```

```
XX = Anum (n0:n,1:5);
```

```
yfit = predict(Mdl4,XX);
```

APPENDIX 2: MULTI-CLASS SUPPORT VECTOR MACHINE CODE of COMPANY 1 DATABASE

```
filename = 'c_datamatlab.xlsx';

[Anum, Btxt] = xlsread(filename,'shifting','a71:f1310');

n0 = 1;
n=1238;

X = Anum (n0:n,1:5);

Y = Btxt (n0:n);

L1= false(n,1);
L2= false(n,1);
L3= false(n,1);
for i=1:n
    L1(i)= strcmp(Y(i),'idle',4);
    L2(i)= strcmp(Y(i),'working',7);
    L3(i)=
strcmp(Y(i),'stop',4); if
L1(i)
    Y1(i)= 1;
elseif L2(i)
Y1(i)= 2;
elseif L3(i)
Y1(i)= 0;
end
end

Y1= zeros(n,1);

Mdl = fitcecoc(X,Y1);

IStd = resubLoss(Mdl);

Mdl2 = fitcecoc(X,Y1,'OptimizeHyperparameters','auto',...
'HyperparameterOptimizationOptions',struct('AcquisitionFunctionName',...
'expected-improvement-plus'));

filename = 'c_datamatlab.xlsx';

[Anum,Craw] = xlsread(filename,'shifting','a2:e68');

n0 = 1;
n = 66;

XX = Anum (n0:n,1:5);

yfit = predict(Mdl,XX);
```

APPENDIX 3: DECISION TREE ALGORITHM CODE of COMPANY 2 **DATABASE**

```
filename = 'values for clustering.xlsx';

[Anum, Btxt] = xlsread(filename,'Sheet1','a794:f1213');

n0 = 1;
n=418;

X = Anum (n0:n,1:5);

Y = Btxt (n0:n);

L1= false(n,1);
L2= false(n,1);
L3= false(n,1);
Y1= zeros(n,1);

for i=1:n
    L1(i)= strcmp(Y(i),'idle',4);
    L2(i)= strcmp(Y(i),'working',7);
    L3(i)= strcmp(Y(i),'stop',4); if
L1(i)
    Y1(i)= 1;
elseif L2(i) Y1(i)= 2;
elseif L3(i) Y1(i)= 0;
end
end

Mdl = fitctree(X,Y1);

view(Mdl,'Mode','graph')

cvmodel = crossval(Mdl);

L = kfoldLoss(cvmodel);

[Anum,Craw] = xlsread(filename,'Sheet1','a1214:f6301');

n1 = 1;
n2 = 5086;

XX = Anum (n1:n2,1:5);

yfit = predict(Mdl,XX);
```

APPENDIX 4: MULTI-CLASS SUPPORT VECTOR MACHINE ALGORITHM **CODE for MILLING MACHINE STATUS CLASSIFICATION**

Week 1

```
filename = 'milling machine power data v4.xlsx';

[Anum, Btxt] = xlsread(filename,'Shifting','a5575:f6000');

n0 = 1;
n=424;

X = Anum (n0:n,1:5);

Y = Btxt (n0:n);

L1= false(n,1);

Y1= zeros(n,1);
for i=1:n
    L1(i)=
    strcmp(Y(i),'working',7); if
    L1(i)
        Y1(i)= 2;
    else Y1(i)=
    1;
end
end

Mdl = fitcecoc(X,Y1);

IStd = resubLoss(Mdl);

Mdl2 = fitcecoc(X,Y1,'OptimizeHyperparameters','auto',...
'HyperparameterOptimizationOptions',struct('AcquisitionFunctionName',...
'expected-improvement-plus'));

[Anum,Craw] = xlsread(filename,'Shifting','a2200:e2445');
n0 = 1;
n = 244;

XX = Anum (n0:n1,1:5);

yfit = predict(Mdl,XX);
```


Week 2

```
[Anum,Craw] = xlsread(filename,'Shifting','a11634:e14430');
```

```
n0 = 1;
```

```
n = 2795;
```

```
XX = Anum (n0:n,1:5);
```

```
yfit2 = predict(Mdl2,XX);
```

Week 3

```
[Anum,Craw] = xlsread(filename,'Shifting','a21484:e23949');
```

```
n0 = 1;
```

```
n = 2464;
```

```
XX = Anum (n0:n,1:5);
```

```
yfit3 = predict(Mdl2,XX);
```

APPENDIX 5: DECISION TREE ALGORITHM CODE for MILLING MACHINE STATUS CLASSIFICATION

Week 1

```
filename = 'milling machine power data v4.xlsx';

[Anum, Btxt] = xlsread(filename,'Moving Average','a11500:c13000');

n0 = 1;
n=1499;

X = Anum (n0:n,1:2);
Y = Btxt (n0:n);

L1= false(n,1);

Y1= zeros(n,1);

for i=1:n
    L1(i)=
    strcmp(Y(i),'working',7); if
    L1(i)
        Y1(i)= 2;
    else
        Y1(i)= 1;
    end
end

Mdl = fitctree(X,Y1);

view(Mdl,'Mode','graph')

cvmodel = crossval(Mdl);

L = kfoldLoss(cvmodel);

[Anum,Craw] = xlsread(filename,'Decision Tree week 1','b2:c10492');

n0 = 1;
n = 10491;

XX = Anum (n0:n,1:2);

yfit = predict(Mdl,XX);
```

Week 2

```
filename = 'milling machine power data v4.xlsx';  
[Anum,Craw] = xlsread(filename, 'Decision Tree week 2', 'b2:c10081');  
  
n0 = 1;  
n = 10080;  
  
XX = Anum (n0:n,1:2);  
yfit2 = predict(Mdl,XX);
```

Week 3

```
filename = 'milling machine power data v4.xlsx';  
[Anum,Craw] = xlsread(filename, 'Decision Tree week 3', 'b2:c9671');  
  
n0 = 1;  
n = 9670;  
  
XX = Anum (n0:n,1:2);  
yfit3 = predict(Mdl,XX);
```

**APPENDIX 6: The MATLAB CODE of the CALCULATION of the UTILIZATION
RATE of the MACHINE**

```
FileName = 'c_datamatlab.xlsx';  
Sheet = 'moving average';  
Range = 'A2:D900';  
ReadVariableNames = false;  
TargetStatus = 2;  
TimeColumn = 'Var1';  
StatusColumn = 'Var4';
```

```
T=readtable(FileName,  
e,... 'Sheet', Sheet,...  
'Range', Range,...  
'ReadVariableNames', ReadVariableNames);
```

```
if (size(T.(TimeColumn)) ~=  
size(T.(StatusColumn))) return;  
end
```

```
TSize = size(T.(TimeColumn));  
TLength = TSize(1);  
TempTv = T.(TimeColumn)(1:TLength);  
Tv = NaT(1, TLength);  
TempSv = T.(StatusColumn)(1:TLength);  
Sv = zeros(0, TLength); (Default status is set as zero)
```

```
MaxTimeGapAllowed = duration(01,00,00);  
MaxDayGapAllowed = 0;  
MaxMonthGapAllowed = 0;  
MaxYearGapAllowed = 0;
```

```
TotalValidTime = 0;  
TotalTargetTime = 0;  
MissedRecordsTime = 0;  
RangeStartIndex = 0;  
TargetCounter = 0;
```

```
if isequal(Sv(1), TargetStatus)  
TargetCounter =  
TargetCounter+1;  
RangeStartIndex =  
1;  
end
```

```

for i = 2:TLength
    tb = between(Tv(i - 1), Tv(i));
    [y,m,d,t] = split(tb, {'years','months','days','time'});

    if ( t > MaxTimeGapAllowed || y > MaxYearGapAllowed ...
        || m > MaxMonthGapAllowed || d > MaxDayGapAllowed)
        MissedRecordsTime = MissedRecordsTime + tb;
        if isequal(Sv(i - 1), TargetStatus)
            TotalTargetTime = TotalTargetTime + ...
                between(Tv(RangeStartIndex), Tv(i - 1));
        end RangeStartIndex = i;
    else
        TotalValidTime = TotalValidTime + tb;
    end

    if isequal(Sv(i), TargetStatus)
        TargetCounter = TargetCounter+1;
        if ~(isequal(Sv(i - 1), TargetStatus))
            RangeStartIndex = i - 1;
        end
    else
        if isequal(Sv(i - 1), TargetStatus)
            TotalTargetTime = TotalTargetTime + ...between(Tv(RangeStartIndex), Tv(i - 1));
        end
    end

    if isequal(Sv(TLength),
        TargetStatus) TotalTargetTime =
        TotalTargetTime + ...
            between(Tv(RangeStartIndex), Tv(i));
    end

    t1 = time(TotalTargetTime);
    t2 = time(TotalValidTime);

    utilizationrate = (t1/t2)*100;

    format short
    utilizationrate;

    disp({'No. Records Searched:', TLength});
    disp({'No. Records of ', TargetStatus, TargetCounter});
    disp({'Total Valid Time:',
        TotalValidTime}); disp({'Total Target
    Time:', TotalTargetTime}); disp({'Missing
    Records For:', MissedRecordsTime});

```

APPENDIX 7: “FRONT PAGE” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE



APPENDIX 8: “REGISTRATION” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE

Registration

Name

Surname

User E-mail

User Phone Number

Company Name

Username

Password

APPENDIX 9: “LOG IN” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE

The image shows a screenshot of a Windows-style dialog box titled "Form_login". The dialog box has a standard title bar with a close button (X) in the top right corner. Inside the dialog, there are two text input fields. The first field is labeled "Login ID" and the second field is labeled "Password". Below the input fields, there are two buttons: "OK" and "Cancel". The "Cancel" button has a dashed border, indicating it is disabled. The entire dialog box is centered on a light gray background.

APPENDIX 10: “WELCOME” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE

WELCOME

Existing Products

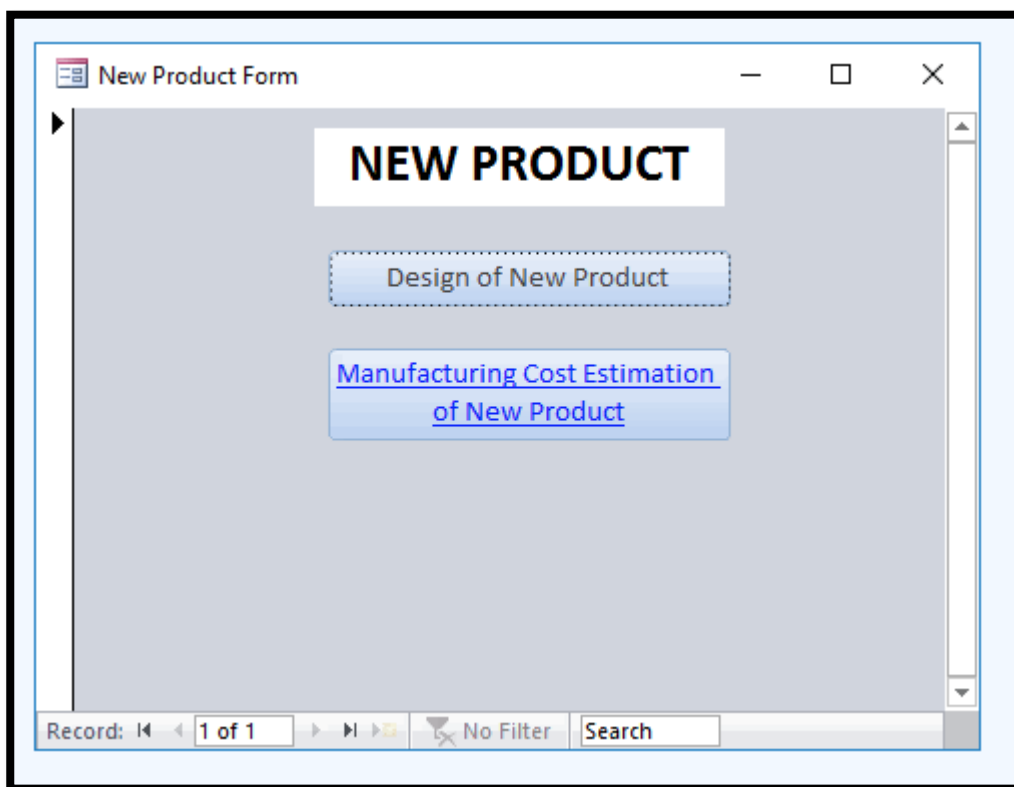
New Product

[Manufacturing Cost Calculation](#)

[Case Based Reasoning for Utilization Rate of a Machine](#)

Close the Database

APPENDIX 11: “NEW PRODUCT” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE



APPENDIX 12: “PRODUCT TYPES” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE



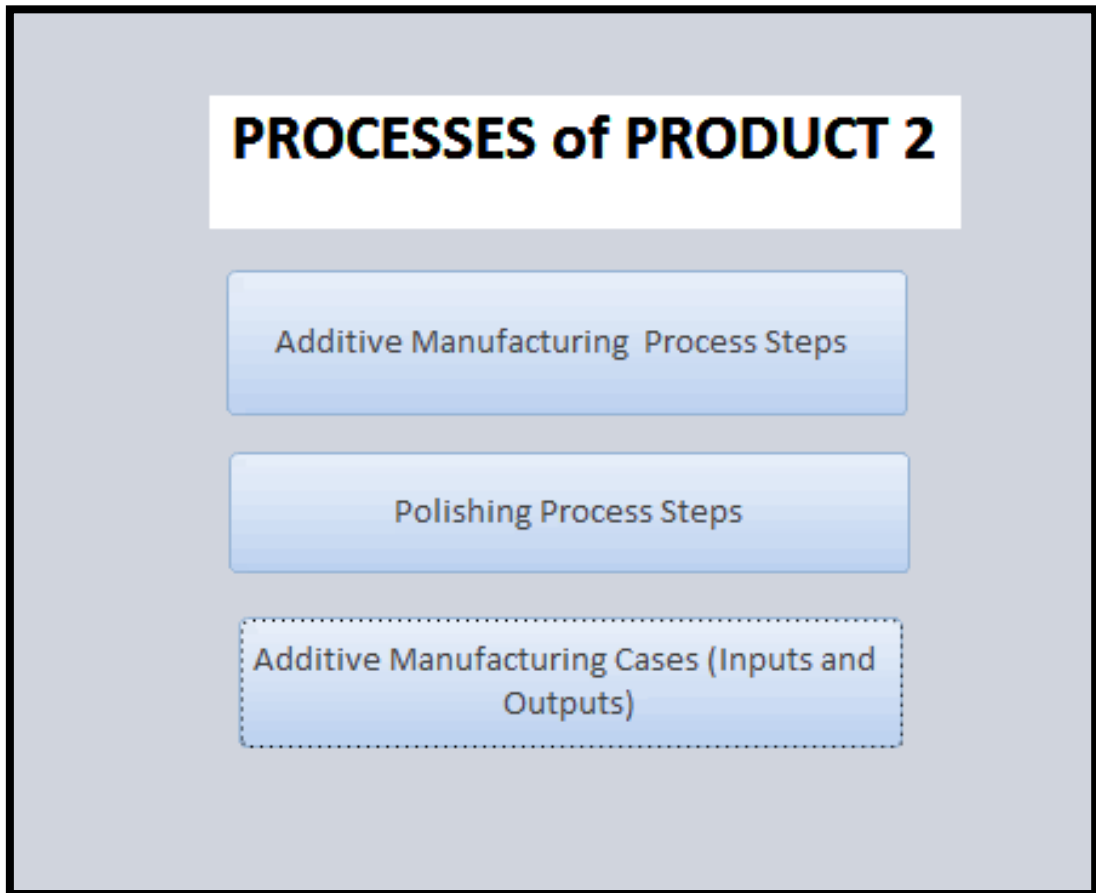
APPENDIX 13: “PRODUCT 2” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE

The image shows a screenshot of a software interface titled "PRODUCT 2 FORM". The interface has a dark blue background. At the top, there is a light blue header box containing the text "PRODUCT 2 FORM". Below this header, there are three stacked, light blue rectangular buttons with rounded corners. The first button is labeled "PRODUCT DRAWING" and has a dotted border. The second button is labeled "PROCESS INFORMATION". The third button is labeled "MANUFACTURING COST CALCULATION" and has a blue underline. In the bottom right corner of the interface, there is a small, light blue button labeled "Go Back".

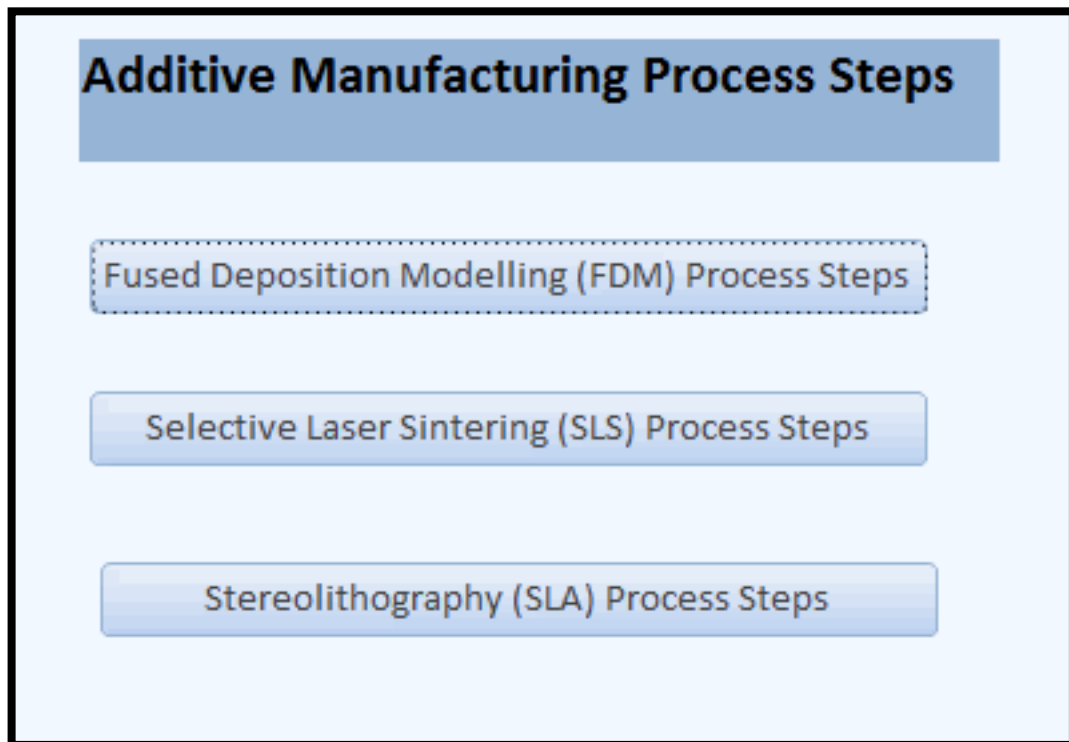
**APPENDIX 14: "PRODUCT DRAWING" FORM OF THE KNOWLEDGE SUPPORT SYSTEM
DATABASE**

Enquiry/Order Form of Perforated Plate Basket (Click on the right box):	\\jmu.ac.uk\pfs\HS04G\Store08\HS215225\My Documents\CROFT PRODUCT DRAWINGS\30737- 1 off
Enquiry/Order Form of Cone Filter (Click on the right box):	\\jmu.ac.uk\pfs\HS04G\Store08\HS215225\My Documents\CROFT PRODUCT DRAWINGS\30737- 1 off

APPENDIX 15: “PROCESSES OF PRODUCT 2” FORM OF THE KNOWLEDGE SUPPORT SYSTEM DATABASE



APPENDIX 16: “ADDITIVE MANUFACTURING PROCESS STEPS” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE



**APPENDIX 17: “FUSED DEPOSITION MODELLING PROCESS STEPS” FORM of the
KNOWLEDGE SUPPORT SYSTEM DATABASE**

Fused Deposition Modelling Process Steps

- Create the part model in STL file with a design program and calculate a path to extrude thermoplastic and any necessary support material.
- Wear safety equipment
- Create the supports and the tool paths if they are needed
- FDM process starts through semi-liquid material feeds into the temperature-controlled FDM extrusion head
- After the process, a solid three-dimensional part is obtained.
- Break away support material or dissolve it in detergent and water.

[Go Back](#)

APPENDIX 18: “ADDITIVE MANUFACTURING CASES” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE

Additive Manufacturing Cases


Fused Deposition Modelling (FDM) Cases

Selective Laser Sintering Process (SLS) Cases

Stereolithography (SLA) Cases

APPENDIX 19: “FUSED DEPOSITION MODELLING CASES” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE

Cases (FDM)

Case Name	<input type="text" value="Ball Mount"/>
Manufacturing Technology	<input type="text" value="FDM"/>
Material	<input type="text" value="ABS"/>
Machine	<input type="text" value="The Replicator 2"/>
Software	<input type="text" value="MakerBot"/>
Type of File	<input type="text" value=".STL"/>
Drawings	

[Go Back](#)

APPENDIX 20: “MANUFACTURING COST CALCULATION” FORM of the KNOWLEDGE SUPPORT SYSTEM DATABASE

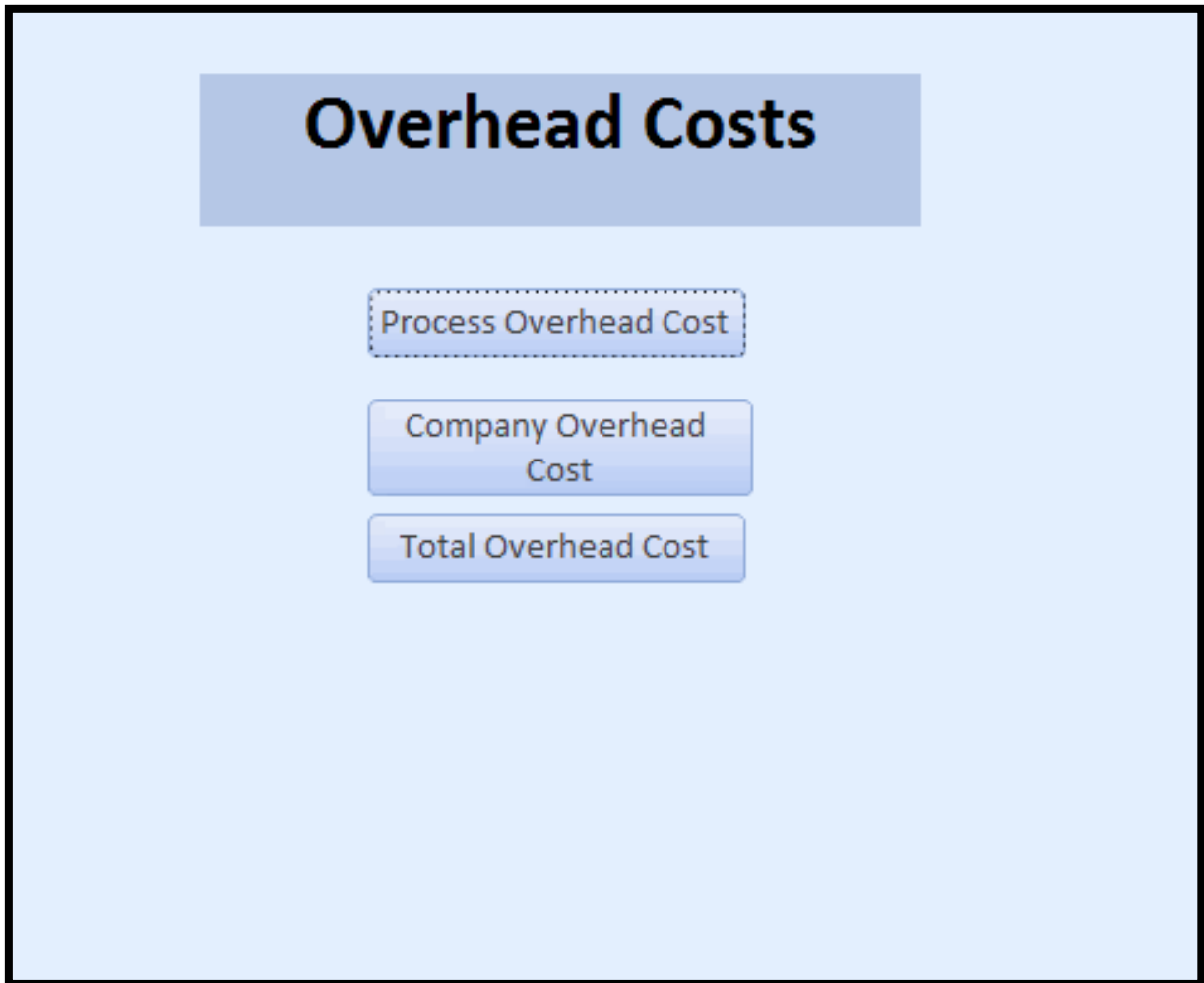
The image shows a screenshot of a software application window titled "Manufacturing Cost Calculation". The window has a standard Windows-style title bar with a minimize button, a maximize button, and a close button. On the left side of the window, there is a vertical scrollbar. The main content area of the window contains a menu with four items, each represented by a blue button with white text:

- Manufacturing Cost Calculation** (highlighted in a darker blue)
- Generic Database
- Product 2 Cost (the button has a dashed border)
- Product 3 Cost

APPENDIX 21: “WELCOME” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE

The image shows a 'WELCOME' form with a light blue background and a black border. At the top, there is an orange rectangular box containing the word 'WELCOME' in bold black capital letters. Below this, there are five buttons arranged vertically. The first button is 'Overhead Costs', the second is 'Direct Process Level Cost', the third is 'Total Unit Manufacturing Cost' (highlighted in yellow), and the fourth is 'Quit'. All buttons have a light blue gradient and rounded corners.

APPENDIX 22: “OVERHEAD COSTS” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE



The diagram shows a light blue rectangular area with a black border. At the top center is a darker blue rounded rectangle containing the text "Overhead Costs". Below this, three rounded rectangular boxes are arranged vertically. The top box has a dashed border and contains "Process Overhead Cost". The middle box has a solid border and contains "Company Overhead Cost" on two lines. The bottom box has a solid border and contains "Total Overhead Cost".

APPENDIX 23: "PROCESS OVERHEAD COST" FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE

Process Overhead Cost

Depreciation of the machine	<input type="text"/>	£0.00
Computer costs	<input type="text"/>	£0.00
Post and stationery	<input type="text"/>	£0.00
Staff Training	<input type="text"/>	£0.00
Total Process Overhead Cost	<input type="text"/>	£0.00
Utilization Rate	<input type="text" value="100"/>	
	<input type="button" value="Calculate from Power Signals"/>	<input type="button" value="Estimate from Case Based Reasoning"/>
Process Overhead Cost after Considered Utilization Rate	<input type="text"/>	£0.00

APPENDIX 24: “COMPANY OVERHEAD COST” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE

Company Overhead Cost			
Period	<input type="text" value="0"/>	Insurance	<input type="text" value="£0.00"/>
Debtors	<input type="text" value="£0.00"/>	Total Gas Cost	<input type="text" value="£0.00"/>
Pension costs and other post-retirement benefits	<input type="text" value="£0.00"/>	Standing Electricity Charges	<input type="text" value="£1.00"/>
Security Charges	<input type="text" value="£0.00"/>	Estate Taxes	<input type="text" value="£0.00"/>
Staff Welfare	<input type="text" value="£0.00"/>	Rent of the facility	<input type="text" value="£0.00"/>
Trade subscriptions	<input type="text" value="£0.00"/>	Maintenance Cost	<input type="text" value="£0.00"/>
Courier Services	<input type="text" value="£0.00"/>	Total Company Overhead Cost	<input type="text" value="£1.00"/>
Medical insurance	<input type="text" value="£0.00"/>		
Light and Heat	<input type="text" value="£0.00"/>		
Accountancy	<input type="text" value="£0.00"/>		
Legal and Professional fees	<input type="text" value="£0.00"/>		
Total Annual Salary of Administrative personnel (Management Fees)	<input type="text" value="£0.00"/>		
Non-direct labour working hours	<input type="text" value="10"/>		
Number of non-direct labours	<input type="text" value="10"/>		
Hourly non-direct labour salaries	<input type="text" value="0"/>		
Administrative personnel for process	<input type="text" value="0"/>		
Total cycle time of the product	<input type="text" value="0"/>		
Total non-direct labour cost	<input type="text" value="£0.00"/>		
Advertising	<input type="text" value="£0.00"/>		
Telephone	<input type="text" value="£0.00"/>		
Travelling	<input type="text" value="£0.00"/>		

**APPENDIX 25: "TOTAL OVERHEAD COST" FORM of the GENERIC COST SUB-DATABASE of
KNOWLEDGE SUPPORT SYSTEM DATABASE**

Total Overhead Cost	
Total Company Overhead Cost	<input type="text" value="£1.00"/>
Process Overhead Cost after Considered Utilization Rate	<input type="text" value="0"/>
Total Overhead Cost	<input type="text" value="£1.00"/>

APPENDIX 26: "DIRECT PROCESS COST" FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE

Direct Process Cost	
Period	<input type="text" value="2016"/>
Work Expenses	<input type="text" value="£200.00"/>
Annual Electricity Cost	<input type="text" value="£0.00"/>
Total Annual Salary of the Operators	<input type="text" value="£20,000.00"/>
Annual working hours of the operators	<input type="text" value="5"/>
Number of Staff	<input type="text" value="1"/>
Hourly Salary of the Operators	<input type="text" value="4000"/>
Number of Operators for Process	<input type="text" value="1"/>
Operator Cost for Process	<input type="text" value="£20,000.00"/>
Raw Material Cost	<input type="text" value="£1,000.00"/>
Packaging Price	<input type="text" value="£200.00"/>
Packaging Time	<input type="text" value="1"/>
Packaging Cost	<input type="text" value="£200.00"/>
Total Process Cost	<input type="text" value="£21,400.00"/>

APPENDIX 27: “TOTAL UNIT MANUFACTURING COST” FORM of the GENERIC COST SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE

Total Unit Manufacturing Cost	
Total Direct Process Cost	£21,400.00
Total Overhead Cost	£1.00
Total Unit Manufacturing Cost	£21,401.00

APPENDIX 28: “WELCOME” FORM of the CASE BASED REASONING SYSTEM SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE

The image shows a graphical user interface for a 'WELCOME' form. It consists of a central gray rectangular area with a black border. At the top of this area is a green rectangular button with the word 'WELCOME' in bold black capital letters. Below this are four light blue rectangular buttons with a slight gradient and rounded corners, each containing text in bold black capital letters. The buttons are stacked vertically and separated by small gaps. The text on the buttons, from top to bottom, is: 'Case Based Reasoning for Design', 'Case Based Reasoning for Manufacturing', 'Case Based Reasoning for Costing', and 'Quit'.

APPENDIX 29: "CASE BASED REASONING" FORM of the CASE BASED REASONING SYSTEM SUB- DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE

Case Based Reasoning Form

Measure Existing Product	<input type="text" value="0"/>
Purchase Wedge Wire	<input type="text" value="0"/>
Buy in Screw Fitting Sectio	<input type="text" value="0"/>
Buy Filter Bags	<input type="text" value="0"/>
Buy in Plug Section	<input type="text" value="0"/>
Punching	<input type="text" value="1"/>
Meshing	<input type="text" value="0"/>
Cutting	<input type="text" value="0"/>
Welding	<input type="text" value="0"/>
Grinding	<input type="text" value="1"/>
Rolling(Forming)	<input type="text" value="0"/>
Checking	<input type="text" value="0"/>
Packing	<input type="text" value="1"/>

Please write "1" next to the processes above if the manufacturing process of the product has these related processes. Otherwise, please write "0" next to the processes above.

APPENDIX 30: "PREVIOUS CASES" FORM of the CASE BASED REASONING SYSTEM SUB-DATABASE of KNOWLEDGE SUPPORT SYSTEM DATABASE

Previous Cases			
ID	<input type="text"/>	SimilarityTotal	<input type="text" value="0.846153846153846"/>
Enquiry ID	<input type="text" value="542"/>	Utilization Rate	<input type="text" value="75"/>
Products	<input type="text" value="Mild Steel Welded Mesh Baskets to drg & as prev supp"/>		
Measure Existing P	<input type="text" value="0"/>		<input type="button" value="NOTES"/>
Purchase Wedge W	<input type="text" value="0"/>		<input type="button" value="Go Back"/>
Buy in Screw Fitting	<input type="text" value="0"/>		
Buy Filter Bags	<input type="text" value="0"/>		
Buy in Plug Section	<input type="text" value="0"/>		
Punching	<input type="text" value="1"/>		
Meshing	<input type="text" value="0"/>		
Cutting	<input type="text" value="0"/>		
Welding	<input type="text" value="1"/>		
Grinding	<input type="text" value="1"/>		
Rolling(Forming)	<input type="text" value="1"/>		
Checking	<input type="text" value="0"/>		
Packing	<input type="text" value="1"/>		
Total Machine-on-t	<input type="text" value="16"/>		
Total Working Time	<input type="text" value="12"/>		