Ding, Y, Zhu, F, Li, H, Parlikad, AK and Xie, M

A failure knowledge graph learning framework for offshore wind turbines with incomplete knowledge

Renewable and Sustainable Energy Reviews 215 (2025) 115561

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

A failure knowledge graph learning framework for offshore wind turbines with incomplete knowledge

Yi Ding [a], Feng Zhu [a], He Li [b,c,*], Ajith Kumar Parlikad [d], Min Xie [a,e]

[a] Department of Systems Engineering, City University of Hong Kong, Hong Kong Special Administrative Region
[b] School of Engineering, Liverpool John Moores University, 3 Byrom Street, Liverpool, L3 3AF, UK
[c] Centre for Marine Technology and Ocean Engineering (CENTEC), Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
[d] Institute for Manufacturing, Department of Engineering, University of Cambridge, Cambridge, UK
[e] Department of Data Science, City University of Hong Kong, Hong Kong Special Administrative Region

A R T I C L E I N F O

Keywords:
Offshore wind energy
Operation and maintenance
Failure knowledge graph
Knowledge transfer
Incomplete knowledge

A B S T R A C T

This study presents a novel framework for Failure Knowledge Graph (FKG) construction tailored for the safe operation and maintenance of offshore wind turbines. Specifically, Bidirectional Encoder Representations from Transformers (BERT) and Conditional Random Field (CRF) are combined for failure extraction, enhanced by iterative learning for failure data transfer from onshore to offshore wind turbines. Additionally, this framework incorporates a rule-based pseudo-label module and an innovative replacement-based pseudo-sample module to mitigate the impact of label errors and failure data imbalance during the iterative learning process. With the failure events extracted, the affiliate components and corresponding failure modes are identified to construct a tree-structured FKG automatically for offshore wind turbines. The feasibility and effectiveness of the proposed framework are validated by the presentation of an FKG regarding 313 offshore wind turbines recorded in the LGS-offshore dataset. Overall, the study provides the offshore wind sector with an intelligent framework for failure data analysis, presentation, and understanding and contributes to the safe operation of offshore wind turbines and wind farms.

Abbreviations

FKG	Failure Knowledge Graph
BERT	Bidirectional Encoder Representations from Transformers
CRF	Conditional Random Field
O&M	Operation and Maintenance
FMEA	Failure Mode and Effects Analysis
FMECA	Failure Mode, Effects, and Criticality Analysis
NER	Named Entity Recognition
NLP	Natural Language Processing
LSTM	Long Short-Term Memory

Nomenclature

$\mathscr{R}^{on}$	Set of records of onshore wind turbines	$\lvert\mathscr{R}^{on}\rvert$	Number of records of onshore wind turbines
$\mathscr{R}^{off}$	Set of records of offshore wind turbines	$\lvert\mathscr{R}^{off}\rvert$	Number of records of offshore wind turbines
r	Record	Y	Predicted label sequence of r
r_i^{on}	i th record of onshore wind turbines	Y_i^{on}	Labeled sequence of failure events of r_i^{on}

(continued on next column)

(*continued*)

r_j^{off}	j th record of offshore wind turbines	Y_j^{off}	Predicted sequence of failure events of r_j^{off}
$\mathscr{K}$	Knowledge database of onshore wind turbine	$\lvert\mathscr{K}\rvert$	Volume of the knowledge base of onshore wind turbine
com_k	k th components in the knowledge database $\mathscr{K}$	des_k	k th failure modes in the knowledge database $\mathscr{K}$
h_l	Hidden states obtained from BERT of the$l$ th token	$\lvert\mathscr{L}\rvert$	Length of record$r$
S_{l,y_l}	Score of the l th tag pertaining to the l th token	$A_{y_l,y_{l+1}}$	Score of a transition from l th to $l+1$ th toke
Y_r	Set of potential sequence combination predictions of r	$\widetilde{Y}$	Each possible sequence prediction of records r
$Y_{r_i}^{on}$	Set of potential sequence combination predictions of r_i^{on}	$\widetilde{Y_i^{on}}$	Each possible sequence prediction of record r_i^{on}
$Y_{r_j}^{off}$	Set of potential sequence combination predictions of r_j^{off}	$\widetilde{Y_j^{off}}$	Each possible sequence prediction of records r_j^{off}
e_j^n	n th failure event of record r_j^{off}	c_j^n	n th context of record r_j^{off}

(continued on next page)

* Corresponding author. School of Engineering, Liverpool John Moores University, 3 Byrom Street, Liverpool, L3 3AF, UK.
E-mail address: he.li@centec.tecnico.ulisboa.pt (H. Li).

https://doi.org/10.1016/j.rser.2025.115561
Received 1 April 2024; Received in revised form 19 August 2024; Accepted 25 February 2025
Available online 6 March 2025

(*continued*)

$\lvert n_j \rvert$	Number of failure events in r_j^{off}	E	Predicted failure events set of the offshore wind turbine
E'	Updated failure events set after module 2	$\overline{Y_j^{off}}$	Updated predicted label sequence of failure events of r_j^{off} after module 2
$score_j^{sc}$	Importance factor of r_j^{off}	$score_j^{r}$	Rareness factor of r_j^{off}
$f_{c_j^n}$	Number of records containing c_j^n	num_j	Number of resampling of r_j^{off}
$\overline{\overline{r}}_j^{\,off}$	j th updated record of offshore wind turbines after module 3	$\overline{\overline{Y}}_j^{\,off}$	Label sequence of failure events of $\overline{\overline{r}}_j^{\,off}$ after module 3
$\lvert \overline{\overline{\mathscr{R}}}^{off} \rvert$	Number of updated records of offshore wind turbines after module 3	θ	Model's parameters
θ_{tea}	Teacher Model's parameters	θ_{stu}	Student model's parameters
$\mathscr{T}$	Knowledge graph	$\mathscr{P}$	Node set
$\mathscr{A}$	Edge set	$\mathscr{I}$	Weight set (failure share)
p_{root}	Node of failure component	p	Node of failure mode
I_m	Child nodes with parent e_m		

1. Introduction

Harnessing the wind resources available distant from the coast and over deep waters, offshore wind farms offer remarkable potential for electricity generation with less human interference and an efficiency advantage [1,2]. Offshore wind farms can generate electricity at rates 1.5 to 2 times greater than onshore wind farms and boast a capacity factor of up to 50 % [3–5]. However, the economic feasibility of offshore wind farms is challenged by their lower accessibility, and thus, the Operation and Maintenance (O&M) cost constitutes 25 %–35 % of that of a wind turbine's lifecycle. On the contrary, the exact O&M cost of onshore wind turbines is generally no more than 25 % [6,7].

Over a prolonged O&M lifespan of 20–25 years, offshore wind turbines should undergo rigorous daily management, including operational control, failure prevention, production efficiency optimization, and maintenance activities management [8]. Effective O&M is pivotal for the resource management of wind farms and evaluating the technical, economic, and social benefits of a wind energy project. Accordingly, the realization of the design efficiency and economic advantages of offshore wind turbines relies heavily on developing practical and effective solutions aimed at the overall O&M cost reduction of wind energy projects [9]. To reduce O&M costs, an in-depth and proactive approach is required for the prediction, prevention, and rectification of incidents, especially for unexpected events that may lead to failure or reduced productivity [10,11]. This necessitates the implementation of effective failure data collection and analysis, which serves as the foundation of O&M cost-reduction-oriented investigations. Analysis of such semantic failure data (maintenance records) provides a comprehensive understanding of failure behaviors and the related maintenance actions of wind turbines, which can support risk assessment, failure analysis, reliability issues, and maintenance planning [12,13].

However, the rapid accumulation of failure data, particularly textual records pertaining to failures and maintenance, significantly challenges the traditional manual-based failure analysis approaches for diverse wind turbine configurations, including onshore, offshore, bottom-fixed, and floating. Especially for offshore wind turbines, the growth of failure data is accelerated by the ongoing offshore wind farms and the addition of new installations, thus imposing substantial constraints on the analysis of operational data. In light of these challenges, there is a pressing need for research to shift towards more effective analytical methodologies, which must be capable of handling, managing, and interpreting vast datasets in an intelligent, automated, and labor-efficient manner, thereby minimizing the reliance on human intervention.

Knowledge Graphs have emerged as a widely accepted tool for this purpose, which can graphically represent semantics by articulating entities and their relationships [14–16]. It is able to encapsulate failure information and associated maintenance actions derived from the failure data, thereby creating Failure Knowledge Graphs (FKGs) for wind turbines. Different from conventional approaches to failure knowledge representation, such as Failure Mode and Effect Analysis (FMEA) [17] and its updated version, Failure Mode, Effects, and Criticality Analysis (FMECA) [18], FKGs offer an innovative, intuitive, and user-friendly graphical format for the presentation of failure and maintenance data [19,20]. This approach necessitates fewer manual interventions, an advantage that becomes increasingly critical as the volume of data escalates to levels that are unmanageable through traditional data analysis and information mining methodologies.

However, FKG construction based on real field data for offshore wind turbines is hampered by several limitations, including incomplete domain knowledge, limited experiential insights, and a lack of comprehensive understanding concerning new failure modes toward offshore wind turbines. Additionally, the existing knowledge base is often inadequate for the automatic identification of all failure characteristics and their intrinsic relationships according to operation data reported by wind farms.

To this end, this study proposes a novel knowledge transfer-based framework for FKG construction for offshore wind turbines specifically tailored to scenarios with incomplete knowledge. This is achieved by harnessing knowledge transfer methodologies, where insights derived from the operational data analysis of the onshore wind sector are extrapolated to the offshore context. This framework aims to pioneer a new paradigm for FKG construction, bolstering performance and cost-efficiency in offshore wind energy through a systematic approach to intelligent failure data analysis and information mining.

The rest of study is organized as follows. Section 2 reviews the state-of-the-art and problems statement. Section 3 constructed the proposed FKG construction framework. The FKG for offshore wind turbines is constructed in Section 4, together with the validation and comparisons. Conclusions are provided in Section 5.

2. State-of-the-art

Maintenance logs are typically collected during routine O&M activities and represent a common form of textual failure data. A tree-structured FKG initially involves extracting failure information from maintenance logs [21]. The associated failure components, failure mode, system configuration, and their relationships are then identified to define the FKG's intermediate nodes and arc, as shown in Fig. 1. Therefore, the initial phase of FKG construction is the extraction of failure events from maintenance logs. These events are then integrated with corresponding failure components and modes to form the nodes of the graph framework. One way to extract these failure events is Named Entity Recognition (NER), a regular Natural Language Processing (NLP) task that locates and classifies named entities presented in unstructured text [22]. Leveraging NER techniques, the unstructured records can be effectively analyzed to identify failure events, thereby establishing a fundamental basis for subsequent graph construction, reliability analysis, and operation management. NER has been extensively used in various domains, including the chemical industry [22] and clinical analysis [23], wherein it aids in identifying chemical elements, medical symptoms, and laboratory tests. The common models include Long Short-Term Memory and Conditional Random Field (LSTM-CRF) model [24], BERT model [23], BERT-Span model [25], and BERT-CRF model [26].

The utilization of NER models in identifying failure events within the records in the wind energy sector holds significant potential for knowledge graph construction [27]. Notably, NER models require failure information as labeled data for training [28]. However, experts within the offshore wind sector possess limited understanding, experience, and expertise regarding failures associated with offshore wind turbines [29–31]. Incomplete knowledge regarding existing failures hampers the accurate labeling process of failure information in operational records, which makes it impossible to use for failure identification

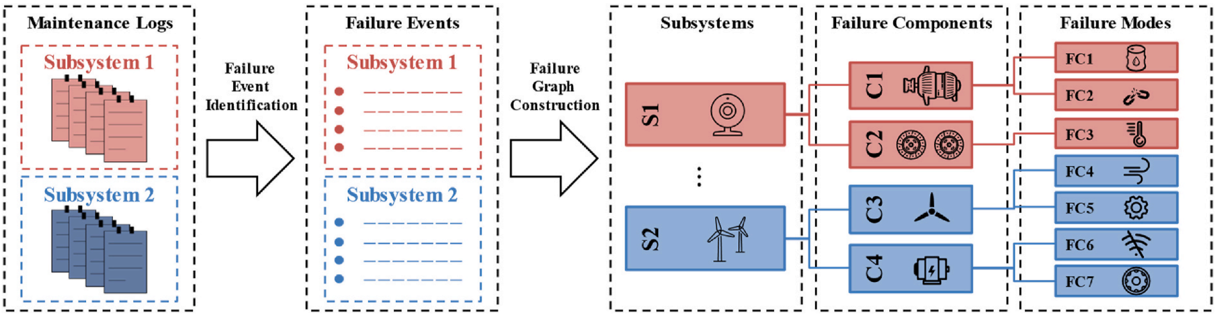

Fig. 1. Flow chart of FKG construction.

model training. In contrast, the knowledge base and research pertaining to onshore wind energy are considerably more abundant, which can provide sufficient labeled data for training.

By considering the relationships and similarities between onshore and offshore wind turbines, an intuitive and feasible approach is to introduce a shared NER model [32,33]. Trained based on onshore turbine data, this shared model is directly applied to offshore systems, aiming to extract pertinent failure information efficiently. However, such approaches overlook the knowledge gaps that persist between onshore and offshore wind turbines. Thus, the shared model inadequately captures all the necessary information on offshore wind turbines, which may significantly affect the model performance for failure information extraction. Therefore, leveraging the unlabeled data available in the offshore systems is a cost-effective approach that can yield valuable information. Iterative learning is a common semi-supervised technique that effectively incorporates unlabeled data to enhance model performance. Specifically, the shared NER model method with iterative learning transfers the semantic spaces of failure information between onshore and offshore wind turbines. The NER model maps labeled onshore data into the semantic space. Through feature transfer, iterative learning is applied to use unlabeled offshore data to adapt the model for offshore failure recognition. This process, following an encoding-transfer-decoding paradigm, is widely applicable in information extraction tasks [34,35]. Nonetheless, applying this technique to the wind energy sector still encounters several challenges due to the unique properties of the collected maintenance logs, which impact failure information extraction and knowledge graph construction. Offshore wind turbines feature a more complex array of subsystems, resulting in different failure modes. Thus, it hampers model accuracy when trained solely on onshore data. During the iterative cycle, an initial error that is too large may cause the model to fail to converge. Besides, failure events are unevenly distributed, with some components exhibiting high failure rates and others demonstrating robust reliability. This issue hinders the model's ability to detect infrequent failures and those related to less common contexts, thereby constraining overall performance [36].

To this end, this study proposes a novel knowledge transfer learning-based framework with iterative learning as a basis to construct failure knowledge graph construction of offshore wind turbines. Additionally, the framework incorporates a pseudo-label module and a pseudo-sample module to mitigate error amplification and address the issue of imbalanced data during the iterative training process to adapt the model for effective unlabeled offshore failure recognition. The contributions of this study include.

(1) Propose a novel framework with iterative learning that incorporates the pseudo-label module and pseudo-sample module to facilitate the effective transfer of failure knowledge from onshore to offshore.
(2) Design a rule-based pseudo-label module to correct and update the labels associated with failure events in offshore records.
(3) Develop a replacement-based pseudo-sample module with a novel resample strategy to generate balanced data with highly reliable records.

Overall, this study proposes a knowledge transfer learning-based failure knowledge graph construction framework for offshore wind turbines with incomplete knowledge. It contributes to knowledge transfer development under errored labels and imbalanced data. Additionally, the study provides the wind energy sector with a reliable tool for intelligent failure data management.

3. Methodology

This section presents the proposed knowledge graph construction approach through knowledge transfer from onshore wind turbines with sufficient failure knowledge to offshore devices with unlabeled failure data. The proposed approach assumes the presence of onshore wind turbine records, inclusive of onshore wind turbine knowledge encompassing failure components and modes, alongside offshore wind turbine records as inputs. An overview of the proposed approach is shown in Fig. 2.

Module 1: Onshore Knowledge Learning (Section 3.1). Construct a BERT-CRF-based sequence labeling model and train the model using onshore failure data and known failure events (labels). It is used to predict failure events of offshore wind turbines. Failure events in offshore records can be obtained.

Module 2: Pseudo-Label Module (Section 3.2). Design a pseudo-label generator to correct and update the labels associated with failures reflected by offshore records. Updated failure events with its corresponding offshore records can be obtained.

Module 3: Pseudo-Sample Module (Section 3.3). Construct a pseudo-sample generator to increase the diversity of labeled offshore records. A balanced dataset that combines failure events with corresponding offshore records can be obtained.

Module 4: Iterative Learning (Section 3.4). Develop a teacher-student model to facilitate iterative learning. This enables a progressive learning process, where the teacher model guides the student model to fully capture the knowledge of the offshore wind. A trained BERT-CRF model for effective failure event identification in the offshore wind field can be obtained.

The overview of the outputs of modules is shown in Fig. 3. The trained model is used to infer failure events in the records. After obtaining failure events, failure modes and components of corresponding offshore wind turbines can be extracted. Then, a tree-structured FKG can be constructed.

3.1. Onshore Knowledge Learning

Denote a collection of onshore maintenance records which describes the failure information and property, as $\mathscr{R}^{on} = \{r_i^{on}\}_{i=1}^{|\mathscr{R}^{on}|}$, where r_i^{on} and $|\mathscr{R}^{on}|$ are the i th record of the onshore wind turbines and its number.

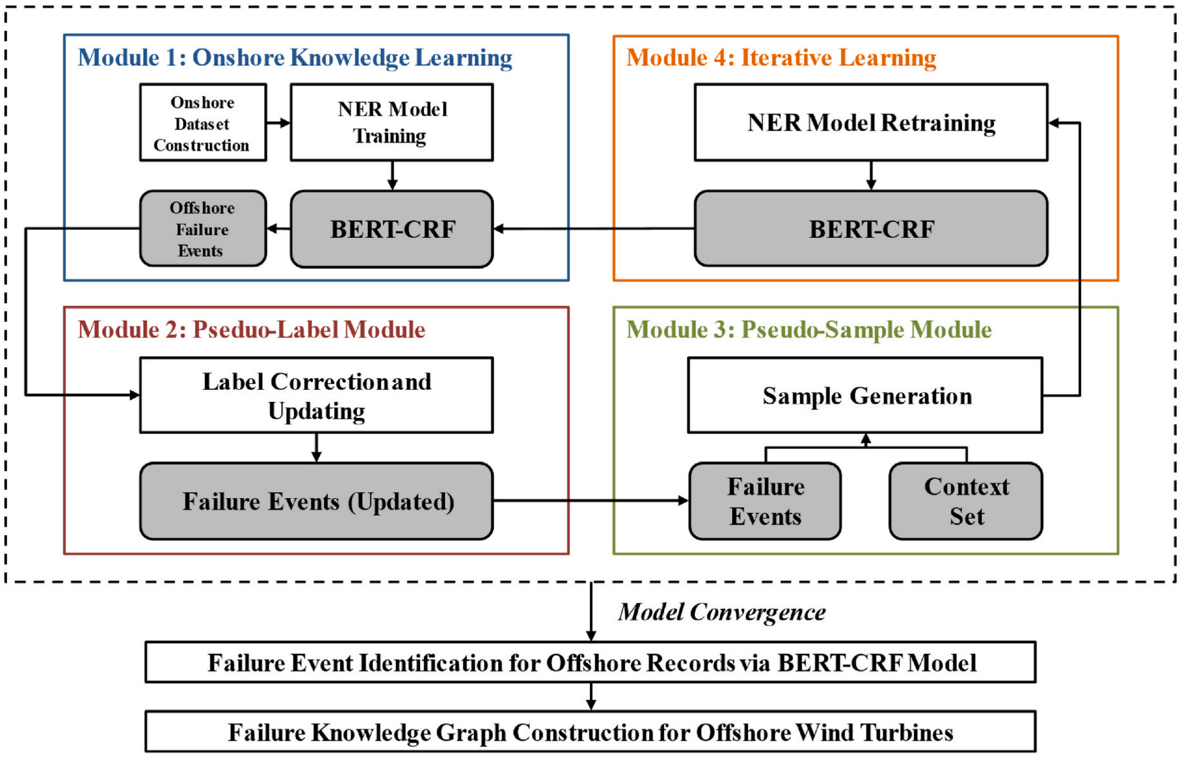

Fig. 2. Overview of the proposed knowledge graph construction approach.

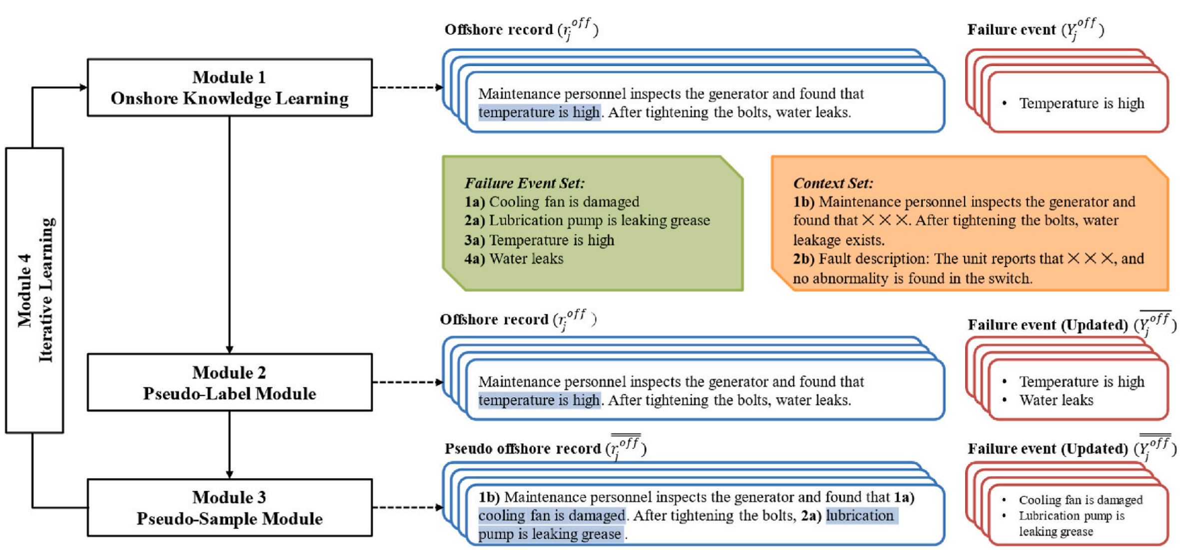

Fig. 3. Overview of the of the outputs of modules.

Define the knowledge database of onshore wind turbine, as $\mathscr{K} = \{com_k, des_k\}_{k=1}^{|\mathscr{K}|}$, contains components and failure modes, where $|\mathscr{K}|$ is the volume of the knowledge base.

Given $\mathscr{R}^{on}$ and $\mathscr{K}$, if the failure description $\{com_k, des_k\}$ in the knowledge base matches in the record r_i^{on}, denote the failure events of record r_i^{on} as $com_k + des_k$. For constructing the label of the records, the first token (word in records) describing a failure is labeled as B-failure, while the remaining tokens within the description are labeled as I-failure. All other tokens are assigned as O. Hence, the dataset $\{r_i^{on}, Y_i^{on}\}_{i=1}^{|\mathscr{R}^{on}|}$ for sequence labeling model training is established, where Y_i^{on} represents the label sequence of record r_i^{on}. As illustrated in Fig. 4, for the input "the staff reported a pitch activation failure," the corresponding label sequence is "O O O O B-failure I-failure I-failure O."

Then, the common sequence labeling model, the BERT-CRF model, which is shown in Fig. 4, is used. The BERT-CRF model amalgamates two crucial components: BERT for feature extraction and CRF for sequence labeling. BERT, built upon the Transformer architecture, serves as a pre-trained model that extracts representations of input text to encapsulate semantic and contextual information within the text. Following BERT processing, tokenized text undergoes conversion into corresponding hidden states. For instance, considering the sentence "The staff reported a pitch activation failure" in Fig. 4, each token is transformed into token embeddings and fed into BERT, resulting in h_l ($l = 1$ *to* 8), signifying the transformation from discrete tokens to continuous hidden states. Subsequently, CRF is employed for sequence labeling, not only ensuring

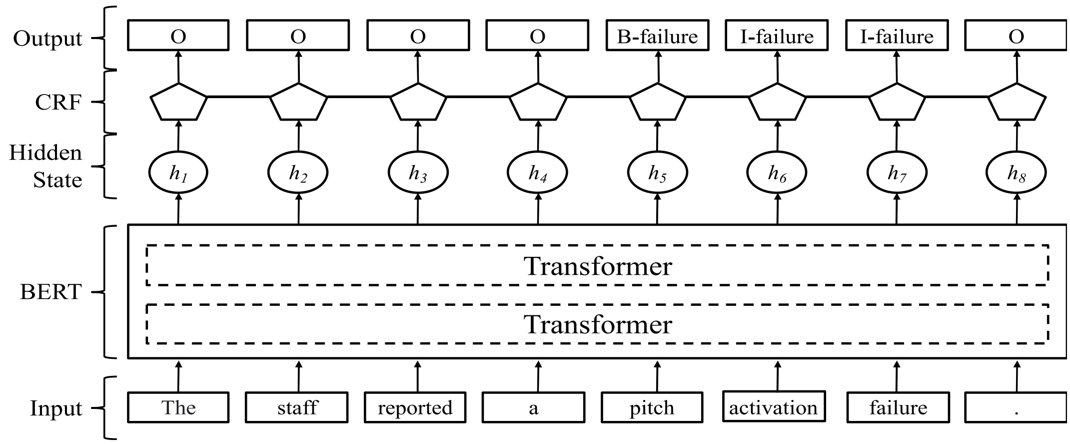

Fig. 4. Overview of the BERT-CRF model.

relationships between output labels but also enforcing constraints or rules among the labels. In the CRF process, each token embedding is passed through a linear layer to derive the score of the l th tag pertaining to the l th token, denoted as S_{l,y_l}. For every input record r that generates a predicted label sequence Y (e.g., "O O O B-failure I-failure I-failure O" in Fig. 4), the scoring function s_θ is defined as follows:

$$s_\theta(r,Y)=\sum\nolimits_{l=1}^{|\mathscr{L}|-1}A_{y_l,y_{l+1}}+\sum\nolimits_{l=1}^{|\mathscr{L}|}S_{l,y_l} \tag{1}$$

where $|\mathscr{L}|$ is the number of the tokens in the record. A is a transition matrix such that $A_{y_l,y_{l+1}}$ represents the score of a transition from l th to $l+1$ th token. This function imposes constraints on the relationships between labels. For example, "I-failure" cannot appear before "B-failure." Afterward, for each training sample, scores for all potential labeling sequences are calculated using the scoring function, which is then followed by the normalization of these scores.

$$P(Y|r)=\frac{e^{s_\theta(r,Y)}}{\sum\limits_{\tilde{Y}\in Y_r}e^{s_\theta(r,\tilde{Y})}} \tag{2}$$

where Y_r is the set of potential sequence combination predictions. $\widetilde{Y}$ is each possible sequence prediction of the records within the entire set. The loss function is determined using the maximum likelihood method of the scores, as:

$$\max_\theta s_\theta(r,Y)-\log\left(\sum_{\tilde{Y}\in Y_r}e^{s_\theta(r,\tilde{Y})}\right)=\min_\theta-\left(s_\theta(r,Y)-\log\left(\sum_{\tilde{Y}\in Y_r}e^{s_\theta(r,\tilde{Y})}\right)\right) \tag{3}$$

Subsequently, model training is performed using gradient descent to update parameters θ until the training process is completed.

Specifically, the BERT-CRF model is trained based on the labeled onshore data, as minimizing the loss function by following:

$$\min_\theta L(\mathscr{R}^{on})=\min_\theta\sum\nolimits_{i=1}^{|\mathscr{R}^{on}|}-\left(s_\theta\left(r_i^{on},Y_i^{on}\right)-\log\left(\sum\nolimits_{\tilde{Y}_i^{on}\in Y_{r_i}^{on}}e^{s_\theta\left(r_i^{on},\tilde{Y}_i^{on}\right)}\right)\right) \tag{4}$$

where $Y_{r_i}^{on}$ denotes a set of potential sequence combination predictions for the record r_i^{on}. $\widetilde{Y_i^{on}}$ signifies each possible sequence prediction of the records within the entire set $Y_{r_i}^{on}$.

The learned model is then exploited for identifying failure events of offshore record $r_j^{off}\in\mathscr{R}^{off}$. Denote a collection of maintenance records of offshore wind farms as $\mathscr{R}^{off}=\left\{r_j^{off}\right\}_{j=1}^{|\mathscr{R}^{off}|}$, where $|\mathscr{R}^{off}|$ is the number of offshore records. Specifically, for the record r_j^{off}, the labeling sequence of failure events Y_j^{off} is calculated by:

$$Y_j^{off}=\arg\max_{\widetilde{Y_j^{off}}\in Y_{r_j}^{off}}s_\theta\left(r_j^{off},\widetilde{Y_j^{off}}\right) \tag{5}$$

where $Y_{r_j}^{off}$ denotes a set of potential sequence combination predictions for the record r_j^{off}. $\widetilde{Y_i^{off}}$ signifies each possible sequence prediction of the records within the entire set $Y_{r_j}^{off}$.

3.2. Pseudo-label module

The model in the first module is likely to lead to incorrect predictions as labels due to the divergence in failures between onshore and offshore wind turbines. Accordingly, a crucial aspect of improving the model performance lies in the provision of a more reliable offshore dataset for supervised training. To address this issue, a pseudo-label module is designed for label correction and updating, as shown in Fig. 3.

Generally, failure events can be characterized by different Parts-of-Speech (POS) type combinations, which can be applied to refine and rectify the labels associated with each failure event prediction [37]. Four rules are formulated to determine the correction of predicted failure events. These rules encompass the following patterns: (1) Noun + Verb, (2) Noun + Verb + Noun, (3) Noun + Verb + Adjective, and (4) Noun + Verb + Prepositional Phrase. Table I gives four examples of these rules. It can be shown that the labeled failure event follows the given rules. Predictions that do not adhere to these rules have a high probability of being incorrect. Therefore, these four rules can be used to broadly filter

Table 1
Examples of four rules.

Rule	Record	Labeled failure event	Predicted failure event
Noun + Verb	The current overloads, leading to a temporary power outage.	current overloads	overloads (×)
Noun + Verb + Noun	It is reported that the device loses connection.	device loses connection	reported that the device (×)
Noun + Verb + Adjective	It was found that the cooling fans operates abnormally indicating a failure.	cooling fans operates abnormally	operates abnormally (×)
Noun + Verb + Prepositional Phrase	The cooling system overheats under load, causing the equipment to shut down.	cooling system overheats under load	The cooling system overheats under load (✔)

correct predicted failure events.

Specifically, the record r_j^{off} is decomposed into failure events and contexts, formulated by $r_j^{off} = \left\{e_j^n \cup c_j^n\right\}_{n=1}^{|n_j|}$, where e_j^n, c_j^n are the n th failure event and context of the j th record. $|n_j|$ is the number of failure events in the record. All failure events are grouped as $E = \{e_j^n | n = 1, \ldots, |n_j|, j = 1, \ldots, |\mathscr{R}^{off}|\}$. If a failure event meets the four rules, update the failure events set by

$$E' = \{e_j^n \,|\, rule\left(e_j^n\right) = TRUE, n = 1, \ldots, |n_j|, j = 1, \ldots, |\mathscr{R}^{off}|\} \tag{6}$$

The failure events in the updated set E' are preserved as corresponding labels.

In some scenarios, failure events identified in certain logs may not be recognized in others. However, when a failure event is identified in one record, it also applies to other records. Hence, if a failure event belongs to E' but has no indications in labels in the record r_j^{off}, update the label by

$$e_j^{|n_j|+1} = \left\{e \,\middle|\, e \in E', e \text{ in } r_j^{off}, e \neq e_j^n, n = 1, \ldots, |n_j|\right\} \tag{7}$$

Following these rules, the offshore dataset is updated as $\left\{r_j^{off}, \overline{Y_j^{off}}\right\}_{j=1}^{|\mathscr{R}^{off}|}$.

3.3. Pseudo-sample module

Following the pseudo-label module, a pseudo-sample module is built to generate new records by replacing the failure events as data augmentation to expand and achieve a more balanced offshore record dataset for model training. Introducing pseudo-sample generation for data augmentation is a conventional approach to address the imbalanced dataset challenge [38]. Following the work [39], to determine the number of resampling records for generating new records for data augmentation, the importance and the rareness of the records are considered. To measure the importance of the records, the idea is that a record containing more failure events is more important than one with no failure events. The importance factor of the j th record is defined as

$$score_j^{sc} = |n_j|. \tag{8}$$

where $|n_j|$ is the number of failure events in the record r_j^{off}.

To incorporate the rareness factor, the general idea says that the rarer a minority context is, the fewer failure events corresponding to that context will be identified, and the more times records containing this context should be resampled. The rareness factor of the j th record is defined by

$$score_j^{r} = -\log_2\left(\max\left\{f_{c_j^n} | n = 1, \ldots, |n_j|\right\} \Big/ |\mathscr{R}^{off}|\right) \tag{9}$$

where $f_{c_j^n}$ refers to the number of records containing c_j^n, $|\mathscr{R}^{off}|$ is the number of records.

A function incorporating these two factors is designed to determine the number of resampling of j th records as

$$num_j = 1 + \left\lceil \sqrt{score_j^{sc} * score_j^{r}} \right\rceil. \tag{10}$$

The add-one avoids removing failure-events-less records from the training set in the following step by guaranteeing all training sentences can be resampled at least once.

The binomial distribution is used to determine the replacement or not a failure event in the resampled record. Specifically, for the failure event e_j^n in record r_j^{off}, a binomial distribution decides whether a failure event e should be replaced randomly from set E', as $\widehat{e_j^n}$ replaces e_j^n.

Following the step, a new record $\overline{\overline{r}}_j\, off = \left\{\widehat{e_j^n} \cup c_j^n\right\}_{n=1}^{|n_j|}$ can be generated to update the BIO-label sequence accordingly and a new dataset $\left\{\overline{\overline{r}}_j\, off, \overline{\overline{Y}}_j\, off\right\}_{j=1}^{|\overline{\overline{\mathscr{R}^{off}}}|}$ is gained. The new dataset is more reliable and improves the model generalization in extracting offshore failures.

3.4. Iterative learning

A teacher-student model is established to learn knowledge from offshore data, where the teacher model guides the student model to fully capture the knowledge of the offshore wind. Initially, a teacher model with parameters θ_{tea} and a student model with parameters θ_{stu} is constructed based on the BERT-CRF model. These models are initialized using θ obtained from onshore data through equation (4) and setting parameters as $\theta_{tea} = \theta_{stu} = \theta$.

Train the student model using the updated offshore wind records, along with their corresponding labels $\left\{\overline{\overline{r}}_j\, off, \overline{\overline{Y}}_j\, off\right\}_{j=1}^{|\overline{\overline{\mathscr{R}^{off}}}|}$, to compute the updated θ_{stu} by minimizing the loss function $\min_\theta L\left(\overline{\overline{\mathscr{R}^{off}}}\right)$. The parameters of the teacher model θ_{tea} are updated following $\theta_{tea} = \theta_{stu}$. The parameters of the student model θ_{stu} is then updated by $\theta_{stu} = \theta$. Subsequently, the teacher model is employed to predict the failure events of the original offshore records $\mathscr{R}^{off}$, resulting in the generation of $\left\{r_j^{off}, Y_j^{off}\right\}_{j=1}^{|\mathscr{R}^{off}|}$.

The process of offshore data inference, pseudo-label generation, pseudo-sample generation, student model training, and teacher model update is repeated until model convergence.

3.5. Failure graph construction

After obtaining the failure event identification model, failure events can be extracted from the records. A tree-structured FKG represented by $\mathscr{T} = (\mathscr{P}, \mathscr{A}, \mathscr{I})$ is constructed with the failure modes and components of offshore wind identified accordingly from events. Specifically, components and failure modes are denoted by $p_{\text{root}}, p \in \mathscr{P}$. An arc $\langle p_{\text{root}}, p\rangle$ will be added to the arc set $\mathscr{A}$ in case a failure component p_{root} and mode p are extracted in the same events. The FKG is constructed in a top-down manner with the hierarchical structure of subsystem - failure components – failure modes. Failure share $C(m,n) \in \mathscr{I}$ reflects the probability of a failure, as:

$$C(m,n) = \frac{N(m,n)}{\sum_{n' \in I_m} N(m,n')} \tag{11}$$

where $N(m,n)$ is the number of occurrences of n. I_m reflect child nodes with parent e_m. When $I_n = \varnothing$, e_n is a leaf node. $N(m,n)$ is expressed as the frequency of n directly. Otherwise, $N(m,n)$ is calculated by $\sum_{o' \in I_n} N(n, o')$.

4. Case study

4.1. Data description

The records utilized in this study include 2022 failure data of 119 onshore wind farms with 1257 wind turbines and 1755 failure data of 6 offshore wind farms with 313 offshore wind turbines (LGS-Offshore dataset). The offshore wind turbines are decomposed into the support structures, pitch system, energy production system, cooling system, and auxiliary system. To be detailed, the records provide the failure calendar of wind turbines and contain a wide range of information such as failure scenarios, failure modes, components, maintenance times, and maintenance actions, see Table II. Failure descriptions in records are regarded as samples for model training. The failure descriptions of records are typically unstructured text due to the inconsistent language habits and unstrict recording rules followed by maintenance personnel, which results in multiple failure descriptions towards the same failure mode and introduces additional restrictions to failure identification and analysis. In addition, offshore devices incorporate more subsystems like dehumidification systems, contributing to distinct failure modes. Hence, it is imperative to emphasize the presence of specific components exhibiting different symptoms between onshore and offshore wind turbines.

Regarding data preprocessing, the initial step involves filtering out numerical values, English characters, and irregular symbols from the records. Subsequently, each sentence is segmented into individual words. A dictionary file[1] https://github.com/dbiir/UER-py/blob/master/models/google_zh_vocab.txtis leveraged, containing 21,128 Chinese characters provided by Google. After segmentation, only words present in this vocabulary are retained to generate the processed record, while eliminating stop words sourced from the repository[2] https://github.com/goto456/stopwords. This meticulous preprocessing approach is crucial for addressing recording inaccuracies and ambiguities. Each word in the processed record undergoes encoding through a combination of word embedding, segment embedding, and position embedding, preparing it for input into the failure event identification model, thereby ensuring data integrity and enhancing the model's text interpretation capabilities.

4.2. Experiment setting

For hyperparameter adjustment and model validation, data are divided into training, validation, and test sets for failure event identification. The training set, being the largest subset, is employed to train the model on data patterns. The validation set plays a crucial role in hyperparameter tuning and unbiased model evaluation. Lastly, the test set, offers an unbiased estimation of the model's performance on entirely new data, thereby facilitating an assessment of the model's generalization capability. This diving method ensures effective model generalization, guards against overfitting, aids in hyperparameter optimization, and provides a reliable gauge of the model's real-world efficacy. Specifically, a randomly selected 10 % failure data of offshore wind is employed as the validation set, with an additional 20 % allocated for model testing, and the remaining failure data of offshore wind turbines and onshore wind turbines serves as the training set for failure event identification. Both validation and testing sets are manually labeled to facilitate a rigorous model performance evaluation.

The hyperparameters of the BERT-CRF model are batch size and learning rate. The batch size is set to grid search within the range [16, 32, 64], while the learning rate is set to grid search within the range [1e-7, 1e-6, 1e-6]. The experiment was conducted on the Google Colab, with NVIDIA Tesla T4. The environment is Python 3.10.11, CUDA 12.0, Pytorch 2.0.1.

To assess the quality of the knowledge graph, evaluating the performance of failure events extraction is crucial, as constructing a knowledge graph depends on accurately extracting entities and relationships from text. The quality of this construction hinges on precise identification of these failure events [40,41]. Three performance metrics are used to evaluate the performance of failure event identification, including precision (P), recall (R), and F1-score ($F1$).

- **Precision:** It measures the accuracy of correctly predicted failure events out of the total instances predicted. A higher precision indicates a higher level of confidence in predictions. It is defined by:

$$P = \frac{|prediction \cap true|}{|prediction|} \tag{12}$$

where *prediction* and *true* indicate the predicted and true failure events extracted according to the labeling sequence. $|prediction \cap label|$ refers to the number of words that overlap in the predicted and labeled failure events. $|prediction|$ represents the number of words of the predicted failures.

- **Recall:** It assesses the model's ability to identify the failure events from the total labeled instances correctly. A higher recall signifies that the model identifies a more significant proportion of failure events. It is calculated by

$$R = \frac{|prediction \cap true|}{|true|} \tag{13}$$

where $|true|$ represents the number of words of the labeled failures.

- **F1-score:** It combines precision and recall to offer a balanced assessment of the model's performance. A higher F1-score shows the superior overall performance of the model in terms of precision and recall, and it is defined as

$$F1 = \frac{2^*P^*R}{P+R}. \tag{14}$$

4.3. Results

The knowledge graph of offshore wind turbines constructed and the failure share calculated is shown in Fig. 5 and Table III. It illustrates: (i) Failure behavior, is a systematic description of failure properties, including components and failure modes. Overall, 18 components with 61 failure modes of offshore wind turbines are determined, which support the deep and comprehensive understanding of the failure mechanisms and behaviors of offshore wind turbines; (ii) Failure share, represents the ratio of failure modes. Take the lubrication system failure of the pitch system as an example. Lubrication system failures are caused by unqualified lubricating oil (36 %) and lubrication pump power failure (64 %). The failure shares identified benefits for most likely failure behaviors identification, which further supports fast failure location and diagnosis as well as maintenance strategy planning of offshore wind farms. The failure characteristics of offshore wind turbines are summarized by systems as follows.

(1) The cooling system

It is pointed out that the cooling system is responsible for the heat dissipation of multiple electrical and electromechanical components. The cooling fan and cooling water pump are the most prevalent components, contributing 94 % of the cooling system's failures, followed by the water tank and water pipe. The damaged cooling fan severely impacts the cooling system, leading to a chain of failures and disrupting the operation of offshore wind turbines.

The malfunctions/failures of the cooling system give rise to common cause failures in such systems and components. Regular inspection and quality control of cooling fans and pumps are required to prevent a minor cooling system failure from propagating to related components in the driving train.

(2) The auxiliary system

The auxiliary system is a combination of blades and supportive

[1] https://github.com/dbiir/UER-py/blob/master/models/google_zh_vocab.txt

[2] https://github.com/goto456/stopwords

Table 2
Examples of unstructured records in maintenance logs.

	Failure Time	Failure Description	System	Maintenance Start	Maintenance Complete
Onshore	2019-10-09 09:00:00	The cooling air temperature is over 65 °C. After inspection, we find that the fan runs normally, but the cooling water pump works abnormally. The pump is replaced.	Cooling	2019-10-09 10:10:00	2019-10-09 18:30:00
	2019-10-28 08:00:00	When cooperating with the generator manufacturer to measure the generator bearing insulation, it is found that the generator bearing insulation of 14# and 37# units fail, and the resistance value is 0.	Generator	2019-10-28 09:10:00	2019-10-29 08:00:00
Offshore	2020-05-02 10:00:00	During the idling debugging process of the 42# unit, the personnel report a pitch activation failure.	Pitch	2020-05-02 11:20:00	2020-05-02 11:50:00
	2019-06-03 12:00:00	Fault phenomenon description: During the unit's operation, the inverter's activation fault is reported. Replace the DSP control board, and the fault is eliminated.	Generator	2019-06-03 14:00:00	2019-06-03 15:00:00

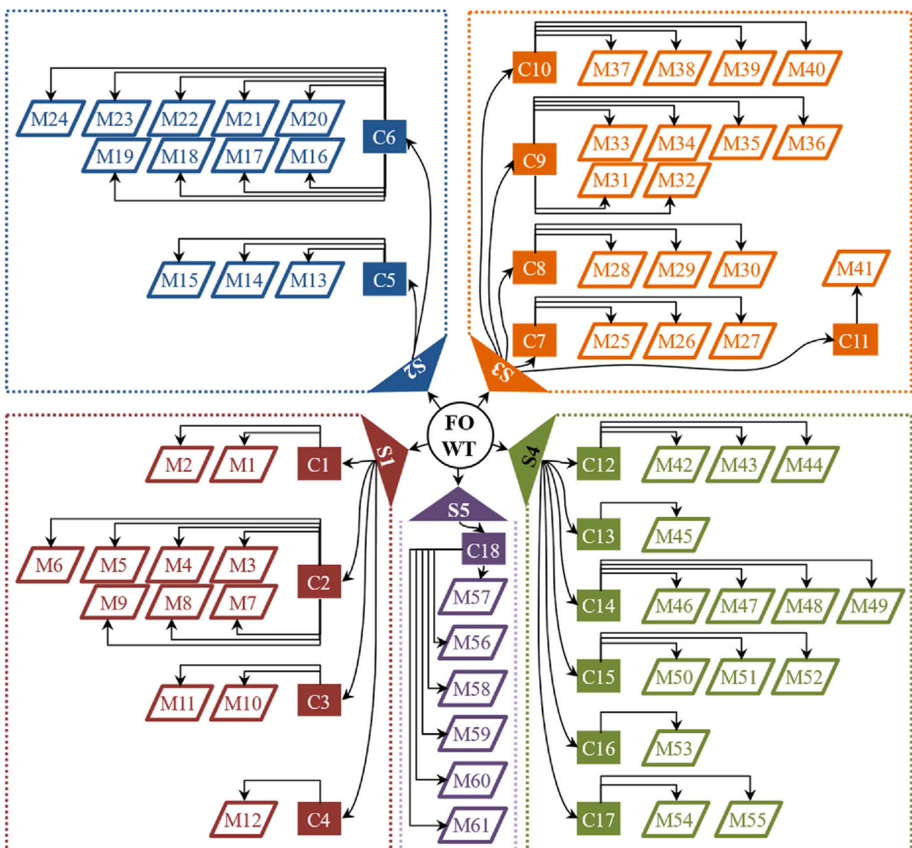

Fig. 5. The FKG of offshore wind turbines (Constructed upon the LGS-Offshore dataset).

elements like yaw system. In this system, blades are responsible for 57 % of the failures. Broken bolts account for 52 % of these failures, followed by damaged blades (27 %) and stuck blades (21 %). Strong winds, large waves, and typhoons (in specific sea locations) are primary contributors to blade failures. The remaining failures in the auxiliary system are related to the yaw system, displaying various symptoms.

It is pointed out that blades are expensive components for manufacturing, installation, operation, and maintenance. The wind farms should ensure zero catastrophic failure, which requires the replacement of the overall lifecycle of offshore wind turbines. To this end, blades' health state monitoring in extreme weather conditions is needed, and only non-destructive monitoring tools and concepts are acceptable.

(3) The energy production system

The energy production system is known by driving train of offshore wind turbines. It composes a selection of components that directly involved in electricity production such as main bearing, inverter, convertor and so forth. In terms of failure of the energy production system, bearing failure is critical and contributes to 51 % of overall failures. Other notable failures include inverter (27 %), lubrication system (11 %), converter (7 %), and carbon brush (4 %). Concerning bearing failures, high temperatures, damaged bolts, and abnormal noise are dangerous states that engineers should be vigilant about. In terms of inverters, the primary failures are alarm activation and current overload.

Unlike other systems that play a supportive role in the energy production process, the failures of any components of the energy production system will give rise to a decrease in energy generation efficiency or shutdown of offshore wind turbines. This highlights the importance of stringent quality control for the electrical and electromechanical

Table 3
Items definition of the FKG for offshore wind turbines.

S1: Cooling system			
C1	Cooling fan (60 %)	M5	Leaky water pump (12 %)
C2	Cooling water pump (34 %)	M6	Damaged airbag (12 %)
C3	Water tank (4 %)	M7	Damaged elastic support (9 %)
C4	Water pipe (2 %)	M8	Damaged sensor (3 %)
M1	Damaged fan (57 %)	M9	Abnormal temperature (3 %)
M2	Overloaded current (43 %)	M10	Leaky water tank (75 %)
M3	Overloaded current (36 %)	M11	Damaged water tank (25 %)
M4	Abnormal pressure (25 %)	M12	Deformed water pipe (100 %)
S2: Auxiliary system			
C5	Blade (57 %)	M18	Abnormal brake (16 %)
C6	Yaw system (43 %)	M19	Damaged motor (16 %)
M13	Broken bolts (52 %)	M20	Abnormal pressure (10 %)
M14	Damaged blades (27 %)	M21	Damaged fuse (10 %)
M15	Stuck blades (21 %)	M22	Damaged resistance box (4 %)
M16	Inverter failure (22 %)	M23	Oil spill (3 %)
M17	Communication failure (17 %)	M24	Burnt black brake unit (2 %)
S3: The energy production system			
C7	Bearing (51 %)	M31	Leaky oil (44 %)
C8	Inverter (27 %)	M32	Damaged lubrication pump (20 %)
C9	Lubrication system (11 %)	M33	Low fluid level (10 %)
C10	Converter (7 %)	M34	Power supply failure (10 %)
C11	Carbon brush (4 %)	M35	Damaged motor (8 %)
M25	High temperature (42 %)	M36	Damaged sensor (8 %)
M26	Damaged bolts (30 %)	M37	Converter overload (46 %)
M27	Abnormal noise (28 %)	M38	High temperature (27 %)
M28	Alarm activation failure (54 %)	M39	Converter activation failure (15 %)
M29	Overloaded current (39 %)	M40	High pressure (12 %)
M30	Precharge failure (7 %)	M41	Carbon brush failure (100 %)
S4: Pitch system			
C12	Power supply (30 %)	M46	Worn encoder (56 %)
C13	Drive (19 %)	M47	High current (20 %)
C14	Motor (17 %)	M48	High temperature (15 %)
C15	Controller (16 %)	M49	Motor stop failure(9 %)
C16	Limit switch (10 %)	M50	Safety chain failure (62 %)
C17	Lubrication system (8 %)	M51	Damaged cabinet door (21 %)
M42	Abnormal power module (49 %)	M52	Controller failure (17 %)
M43	Damaged capacitor cabinet (28 %)	M53	Smashed limit switch (100 %)
M44	Low capacitor voltage(23 %)	M54	Lubrication pump power failure (64 %)
M45	Damaged drive (100 %)	M55	Unqualified lubricating oil (36 %)
S5: Support structure			
C18	Tower system (100 %)	M59	Dehumidifier failure (6 %)
M56	Control cabinet failure (43 %)	M60	Lighting system failure (6 %)
M57	Switch failure (25 %)	M61	Stuck screw (3 %)
M58	Communication failure (17 %)		

SX: XX – Subsystem; CX XX(xx%) – Failure component (Failure share); MX XX(xx %) – Failure mode (Failure share).

components of the system during the manufacturing process to ensure its adequate reliability. It is worth mentioning that electrical components fail more frequently than others due to degradation, such as the lubrication system, and wear, such as carbon brushes. The findings indicate that such components' supply chain control and backups are essential for offshore wind farm management.

(4) The pitch system

The pitch system is of importance for electricity production efficacy improvement and safety control in extreme sea conditions by adjusting attack angles. The system is independently powered and consists of powers (with backups), motors, and activators. Specifically, the power supply represents the most significant contributor to pitch system failures, accounting for 30 % of failures to the total, considering that motor and controller failures account for approximately 20 % of each, respectively. Additionally, the limit switch and lubrication system are also observed in failures of the pitch system.

Even though the powers have been equipped with backups, the failure rate is still high. It calls for a better design of the scheduling algorithm for backup units from the software's point of view. From the hardware's perspective, real-time monitoring of power units is also essential. Additionally, timely failure prevention of these components should be integrated into the maintenance strategy planning of offshore wind farms.

(5) The support structures

The support structures are specially designed for offshore wind turbines to support such devices stand on the sea space, representing the unique differences between onshore and offshore wind turbines. The obvious reasons for the failures of this system can be traced back to the control cabinet failure (43 %), switch failure (25 %), and communication failure (17 %). Additionally, minor failures include lighting prevention system failure (6 %), dehumidifier failure (6 %), and stuck screws (3 %).

It is pointed out that LGS-Offshore datasets creatively provide the wind energy sector with a basic data foundation and understanding of failures, the populations which, however, are 2–5 years old. It indicates that structural failures do not appear in these young wind turbines, resulting in this analysis merely reporting functional failures but not structural ones. However, the structural failures introduced by harsh sea conditions should be emphasized in wind farms' operation and maintenance process. Like several systems mentioned, the failure of electrical elements is the most frequent of the support structures. Robust designs and stringent electrical component selection can be solutions for unwanted failure prevention for this system.

To conclude, from the methodologies' point of view, the proposed method serves as a valuable tool for acquiring failure information and conducting reliability analysis for offshore wind turbines through knowledge transfer. It enables the identification of failure modes of offshore wind turbines by leveraging operational knowledge and failure data from analogous onshore wind turbines. Notably, the method excels in extracting unique symptoms of components specific to offshore wind turbines, even when faced with different failure systems. Addressing the domain gaps between onshore and offshore wind farms, the proposed method effectively transfers knowledge to extract new failure information, providing a comprehensive understanding of failures. The resulting FKG facilitates querying of failure mechanisms and supports optimized operation and maintenance of offshore wind turbines by implementing failure localization and diagnosis. The presented information is designed to be easily understandable and user-friendly. Another contribution of the study is that it provides an automatic tool for failure information extraction, which requires no human interactions and supports future big data scenarios.

4.4. Discussions

4.4.1. Iterative learning

In the process of iterative learning, the generation of pseudo-samples needs to determine the number of records to resample. The range of importance score is $[0,5]$, while the range of rareness score is $[2.1,10.1]$. And the range of combined score, which is the number of resampling of records, is $[1,9]$. Considering the sampling uncertainty of failure events for replacement in iterative learning, the mean and variance of the criteria are calculated to evaluate the performance of the proposed method.

Fig. 6 shows the precision, recall, and F1-score of the proposed model during training processes. The mean and variance of the criteria are calculated over 20 runs. As the training set has no label information, the validation set is used to calculate during the training stage. Precision and recall converge stably with less fluctuation, which demonstrates the convergence of the proposed method. The F1-score, which symbolizes the overall performance of the proposed model, gradually increases and

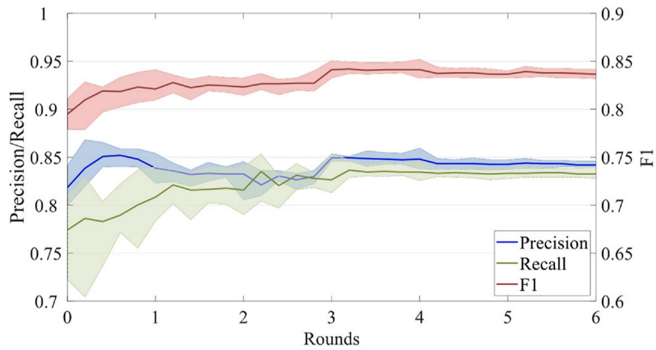

Fig. 6. Model performance of the proposed method over the training rounds.

stabilizes during iterative learning. It verifies the effectiveness of the proposed model. The results also illustrate that the performance of the model increases less after the first four runs, and it proves the calculation efficiency of the proposed model and guides the calculation of the model when applied to other data analysis circumstances.

4.4.2. *Impacts of sample size*

The sample sizes of both labeled onshore data and unlabeled offshore data, which are inputs of the failure event identification model, are pivotal in determining the performance.

A larger sample size of labeled onshore data provides a diverse set of failure events of onshore data, enhancing the model's learning and generalization. This results in improved knowledge transfer and identification of offshore data. Understanding the impact of sample size of labeled onshore data helps determine adequate training onshore data for a robust model that can improve the method's performance. Fig. 7 validates the performance of the proposed method on different sample sizes of labeled onshore data. The results demonstrate a gradual improvement in the performance of the method as the sample sizes of onshore maintenance records increase. This improvement can be attributed to the fact that a larger sample size of onshore maintenance records contains more failure information about onshore wind farms, which in turn facilitates the extraction of failure information of offshore wind farms via knowledge transfer.

Fig. 7 also indicates that the increase in accuracy is in line with the increase in sample size. It can be an asset for model accuracy estimation and making the best decisions regarding failure data analysis by balancing accuracy and computational efficiency.

A larger unlabeled offshore data samples facilitate the extraction of failure information during iterative learning for improved model adaptation in offshore wind farm contexts. Fig. 8 validates the performance of the proposed method on different sample sizes of offshore data. It is shown that expanding sample sizes of unlabeled offshore records results in a consistent method of performance enhancement. This improvement

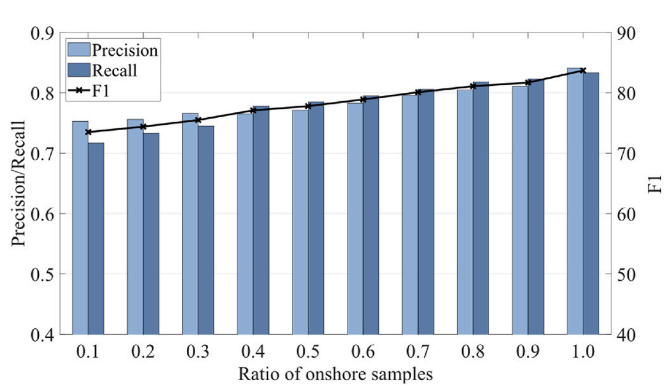

Fig. 7. Model performance with different sample sizes of labeled onshore data.

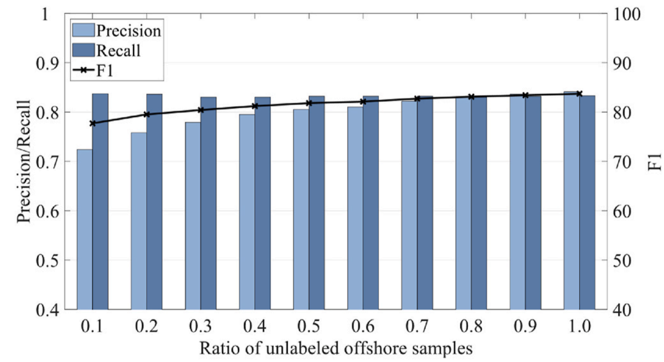

Fig. 8. Model performance with different sample sizes of unlabeled offshore data.

results from more failure information being extracted during iterative learning, leading to better model adaptation in offshore wind farm contexts. Noteworthy is the significant impact of labeled onshore data on the model's performance compared to unlabeled offshore data. This underscores the importance of leveraging onshore data and knowledge base for onshore labeled data obtained for refining the model towards optimal performance.

4.4.3. *Ablation analysis*

The ablation analysis is designed to understand the influence of modules of the proposed method. The insights into the relative significance of different elements can be verified by selectively disabling elements and comparing the performance before and after the ablation. The results promote a full understanding of the capability of each module in the proposed method. Different combinations of modules for the failure identification are examined, including pseudo-label module, pseudo-sample module, and iterative learning module, see Table IV.

The results demonstrate that the pseudo-label module and pseudo-sample module with iterative learning can significantly improve the model performance. The proposed method can dig out new knowledge better and more profoundly.

Specifically, compared to the baseline, the inclusion of the pseudo-label module (module 2) results in a notable improvement in F1 value by 1.8 %. This improvement indicates that the presence of incorrectly labeled offshore data due to the knowledge gaps in the transfer learning can hinder performance, and the proposed pseudo-label module effectively corrects labels for offshore data, establishing a solid foundation for training. Additionally, the pseudo-sample module (module 3) improved the F1 value by 1.1 %, suggesting that generating a more diverse range of samples enhances the model's overall performance. Both generators (pseudo-label module and pseudo-sample module) pull up the F1 value by 5.1 %. This improvement is more pronounced compared to the use of either element individually due to the presence of incorrect labels of offshore data results from the knowledge transfer, which continues to limit performance. The proposed pseudo-label module can provide reliable data for subsequent pseudo-sample

Table 4
Model performance of different module combinations.

Models	*Precision*	*Recall*	*F1*
Module 1 (baseline)	0.774	0.748	0.761
Module 1 + Module 2	0.760 (0.014↓)	0.799 (0.051↑)	0.779 (0.018↑)
Module 1 + Module 3	0.757 (0.017↓)	0.787 (0.039↑)	0.772 (0.011↑)
Module 1 + Module 2 + Module 3	0.838 (0.064↑)	0.802 (0.054↑)	0.812 (0.051↑)
The proposed method	**0.841 (0.067↑)**	**0.833 (0.085↑)**	**0.837 (0.076↑)**

generation, resulting in diverse, reliable, and balanced offshore data for training.Besides, the iterative learning approach (module 4), as evident from the final lines in the table, enables continuous learning from offshore data. This iterative learning process further contributes to extracting and utilizing valuable knowledge from the offshore data.

4.5. Comparisons

4.5.1. Comparison – methodology

Five models are considered in the comparison to validate the performance of the proposed method of failure event identification: (i) Rule-based model [42]. The rule-based model extracts the contexts with high frequencies as rules and uses them as indicators to discover the failure events; (ii) BERT-Span model [25]. We use labeled data in the onshore wind farms to generate pseudo-labeled data in the offshore data for training the BERT-Span model; (iii) BERT model [23]. The training process is like the second model; (iv) LSTM-CRF model [24]. The training process is like the second model; (v) Multi-head teacher-student model [43]. The labeled onshore data is applied to generate pseudo-labeled data for that of offshore data for training. The pseudo-labeled offshore data is then randomly divided into several parts to train corresponding models as teacher models, and then use the union of these models to infer results, see Fig. 9.

It is concluded that the proposed method shows the highest F1-score, indicating that it holds the best performance among others. The rule-based model shows worse performance due to its template matching mechanism, which ignores the hidden contextual information on the unstructured text. The BERT-Span model, BERT model, and LSTM-CRF model perform worse due to the ignorance of knowledge gaps between offshore and onshore wind turbines, which results in unreliable training labels and misleading the training process; the multi-head teacher-student model has limited capability of failure identification as it lacks a continuous learning process to further improve the performance of the model.

It is also noted that the rule-based method is more effective for small sample size conditions. This is because the other methods employed are based on deep learning, which requires a larger dataset for training to obtain a robust and high-performance model. According to the results from the comparison study, the proposed method can guarantee good performance without high labeling and annotation efforts and achieve effective knowledge transfer from the high-knowledge onshore domain when given a large volume of onshore failure data.

This study provides a novel solution for textual failure data analysis in the wind energy sector, addressing challenges posed by rapidly accumulating and diverse data forms. The proposed FKG learning framework can reduce the overarching operation and maintenance costs of the offshore wind energy sector as well as the labor request. Meanwhile, by providing a practical FKG construction schedule, this study contributes to better utilization of failure data. Overall, the outcomes of this study benefit both industry and academia, providing a novel systematic framework and operational procedure for failure data analysis and FKG construction.

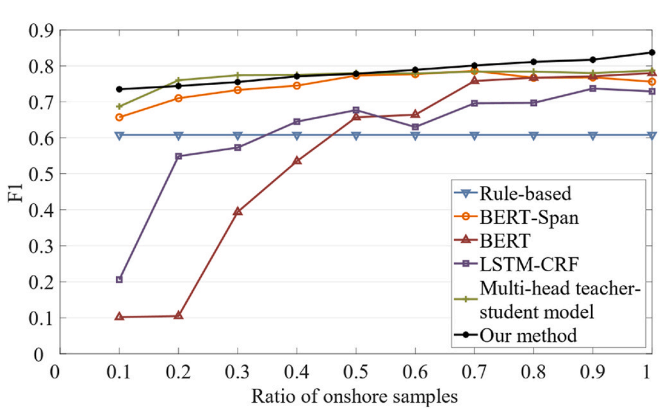

Fig. 9. Benchmarking results of the proposed method.

4.5.2. Comparison – offshore wind turbines

A comparison is conducted to examine the failure modes of offshore wind turbines extracted in various research for failure analysis, including: Risk-based failure analysis [44], FMEA [45,46], Correlation-FMEA [47], FMECA [48], cost-based FMEA [49], AHP-FMEA [7], Two-stage FMEA [50], see Table V.

Notably, the proposed methodology in this study identifies 61 failure modes of 24 components, which represents more failures identified than existing understandings published before. On the one hand, it verifies the comprehensiveness of the proposed method in revealing the failure properties of offshore wind turbines. It also provides the sector with a more detailed and extensive understanding in terms of offshore wind turbine failures, which act as the foundation of failure prevention and control, reliable operation and maintenance issues, as well as cost savings for offshore wind farms.

It is worth mentioning that this analysis delved into a more detailed investigation of failure modes within the cooling system and energy production system, as well as their corresponding components, such as the cooling fan, cooling water pipe, carbon crush, and lubrication system. This emphasis arises from the high frequency of failure modes observed in these subsystems within the onshore system. Leveraging

Table 5
Components and failure modes of offshore wind turbines.

Components	This study	[a]	[b]	[c]	[d]	[e]	[f]	[g]	[h]
Cooling system	✓	✓		✓	✓			✓	✓
Cooling fan	✓								
Cooling water pump	✓								
Water tank	✓								
Water pipe	✓								
Auxiliary system	✓		✓	✓				✓	✓
Hub				✓			✓	✓	✓
Blade	✓		✓	✓			✓	✓	✓
Yaw system	✓		✓	✓	✓			✓	✓
Energy production system	✓				✓				
Generator			✓	✓	✓			✓	
Transformer				✓	✓			✓	✓
Bearing	✓	✓	✓		✓	✓		✓	✓
Inverter	✓								
Lubrication system	✓	✓		✓		✓	✓		
Converter	✓			✓	✓			✓	✓
Carbon brush	✓								
Main Shaft and/or Coupling				✓			✓	✓	✓
Pitch system	✓		✓	✓				✓	✓
Power supply	✓								
Drive	✓								
Motor	✓								
Controller	✓		✓	✓	✓			✓	✓
Limit switch	✓								
Lubrication system	✓								
Support structure	✓		✓	✓	✓			✓	✓
Tower	✓		✓	✓				✓	✓
Others		✓	✓	✓				✓	✓
Number of failure modes	61	36	16	31	9	7	8	32	42

[a] Sinha and Steel [44]; [b] Arabian-Hoseynabadi, Oraee [45]; [c] Bharatbhai [46]; [d] Kang, Sun [47]; [e] Du [48]; [f] Tazi, Châtelet [49]; [g] Li, Teixeira [50]; [h] Li, Díaz [7].

knowledge transfer based on failure data from onshore systems, the proposed method became sensitive to failure information related to these components. Consequently, the proposed method excels at identifying failure modes specific to these components. Meanwhile, existing studies focus on understanding the failure modes of generators and transformers. However, this study reveals less failure (by proportion) of these aspects. The reason lies in the nature of the records, where failure descriptions involving generators and transformers often overlap with other systems, such as the cooling system.

Except for the discussed, in the algorithm perspective, traditional methods like FMEA and FMECA rely on manual information extraction and analysis by expert teams. This process is time-consuming, costly, and particularly challenging for large-scale systems. These methods may also struggle to fully explore the potential interrelationships within the records between different types of wind turbines. In contrast, the proposed method automates the failure information identification to construct a knowledge graph for failure analysis that does not require human involvement. Leveraging a comprehensive knowledge base of onshore wind turbines, the proposed approach can use this as a reference label for failure event identification model training for offshore wind turbines, enabling the exploration of semantic correlations between onshore and offshore wind turbine records. This feature helps the knowledge transfer, enhancing the efficiency of the analysis process. However, the proposed method need adequate labeled onshore data by adopting an onshore wind turbine knowledge base for effective BERT-CRF model training. Large deep learning models like BERT-CRF demand substantial computational resources for training and inference, potentially posing challenges in terms of hardware.

In conclusion, this study's main contribution is that it provides an intelligent tool for failure data analysis of offshore wind turbines and can be extended to other scenarios. The approach can be generalized to various wind turbines because the methodology is to establish a transfer between failure events of the onshore wind turbines and offshore wind turbines in the semantic spaces. A deep learning-based method is adopted to transfer the semantic spaces of failure events of both. The failure event identification model BERT-CRF maps labeled onshore data into the semantic space. Through feature transfer, a rule-based pseudo-label module, pseudo-sample module, and iterative learning are applied to adapt the model for unlabeled offshore failure recognition. Consequently, failure components and modes can be decomposed from the events to construct the knowledge graph. As the records of various wind turbines have correlations in semantic perspective, the method's ability to transfer and learn semantic features within this semantic space enables its applicability to broader Wind Farm scenarios, showcasing its potential for generalization across diverse wind turbine settings.

5. Conclusion

The study introduces a novel knowledge transfer learning framework, utilizing incomplete knowledge to construct a Failure Knowledge Graph for offshore wind turbines based on onshore data. The proposed method employs the BERT-CRF model with iterative learning for deep extraction of failure information in offshore wind turbines. It also incorporates two key components: a rule-based pseudo-label module for refining failure information labels during learning and a pseudo-sample module with an innovative resampling strategy to generate a balanced dataset with more reliable records. A Failure Knowledge Graph is constructed as a basis to represent the failure information of offshore wind turbines, providing understanding for failure analysis, understanding, and prevention of failures. According to the Failure Knowledge Graph, the failure properties of offshore wind turbines are presented by their components. Overall, the outcomes of this study contribute to performance improvement and cost-saving of offshore wind turbines by providing a novel systematic framework for intelligent failure data analysis, presentation, and understanding.

Although the proposed method outperforms the knowledge graph construction task, it still can be further improved considering the practical applications. Firstly, since the method requires a amount of labeled onshore data, it is crucial to use large language models with prior knowledge to reduce data dependency. Secondly, exploring model optimization methods that require fewer computational resources is worthwhile. In the future, we will continue to explore these two directions.

CRediT authorship contribution statement

Yi Ding: Methodology, Writing – original draft. **Feng Zhu:** Formal analysis. **He Li:** Methodology, Writing – original draft, Writing – review & editing. **Ajith Kumar Parlikad:** Methodology, Supervision. **Min Xie:** Writing – review & editing, Supervision.

Declaration of competing interest

There are no Conflicts of Interest.

Acknowledgment

This work is supported by National Natural Science Foundation of China (72371215, 72301299) and by Research Grant Council of Hong Kong (11201023, 11202224). It is also funded by Hong Kong Innovation and Technology Commission (InnoHK Project CIMDA), the Horizon Europe Marie Skłodowska-Curie Postdoctoral Fellowship (DROMS-FOWT–101146961), and UKRI (EPSRC EP/Z001501/1).

Data availability

Data will be made available on request.

References

[1] Lüth A, Keles D. Risks, strategies, and benefits of offshore energy hubs: a literature-based survey. Renew Sustain Energy Rev 2024;203:114761.
[2] Edwards EC, Holcombe A, Brown S, Ransley E, Hann M, Greaves D. Evolution of floating offshore wind platforms: a review of at-sea devices. Renew Sustain Energy Rev 2023;183:113416.
[3] Ohunakin OS, Matthew OJ, Adaramola MS, Atiba OE, Adelekan DS, Aluko OO, et al. Techno-economic assessment of offshore wind energy potential at selected sites in the Gulf of Guinea. Energy Convers Manag 2023;288:117110.
[4] Irena S. Offshore renewables: an action agenda for deployment. Int Renew Energy Agency; 2021. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2021/Jul/IRENA_G20_Offshore_renewables_2021.pdf.
[5] Li H, Huang C-G, Soares CG. A real-time inspection and opportunistic maintenance strategies for floating offshore wind turbines. Ocean Eng 2022;256:111433.
[6] Martinez A, Iglesias G. Site selection of floating offshore wind through the levelised cost of energy: a case study in Ireland. Energy Convers Manag 2022;266:115802.
[7] Li H, Díaz H, Soares CG. A failure analysis of floating offshore wind turbines using AHP-FMEA methodology. Ocean Eng 2021;234:109261.
[8] Ren Z, Verma AS, Li Y, Teuwen JJ, Jiang Z. Offshore wind turbine operations and maintenance: a state-of-the-art review. Renew Sustain Energy Rev 2021;144:110886.
[9] Garcia-Teruel A, Rinaldi G, Thies PR, Johanning L, Jeffrey H. Life cycle assessment of floating offshore wind farms: an evaluation of operation and maintenance. Appl Energy 2022;307:118067.
[10] Li X, Zhou S, Zhao Y. Onshore and offshore wind power generation forecasting using a novel flexible time-varying fractional nonlinear grey model. Energy Convers Manag 2023;297:117695.
[11] McMorland J, Flannigan C, Carroll J, Collu M, McMillan D, Leithead W, et al. A review of operations and maintenance modelling with considerations for novel wind turbine concepts. Renew Sustain Energy Rev 2022;165:112581.
[12] Nguyen TAT, Chou S-Y. Maintenance strategy selection for improving cost-effectiveness of offshore wind systems. Energy Convers Manag 2018;157:86–95.
[13] Li H, Peng W, Huang C-G, Guedes Soares C. Failure rate assessment for onshore and floating offshore wind turbines. J Mar Sci Eng 2022;10(12):1965.
[14] Ji S, Pan S, Cambria E, Marttinen P, Philip SY. A survey on knowledge graphs: representation, acquisition, and applications. IEEE Trans Neural Netw Learn Syst 2021;33(2):494–514.
[15] Chen X, Jia S, Xiang Y. A review: knowledge reasoning over knowledge graph. Expert Syst Appl 2020;141:112948.
[16] Ding Y, Li H, Zhu F, Wang Z, Peng W, Xie M. A semi-supervised failure knowledge graph construction method for decision support in operations and maintenance. IEEE Trans Industr Inform 2023;20(3):3104–14.

[17] Li H, Diaz H, Soares CG. A developed failure mode and effect analysis for floating offshore wind turbine support structures. Renew Energ 2021;164:133–45.
[18] Leimeister M, Kolios A. A review of reliability-based methods for risk analysis and their application in the offshore wind industry. Renew Sustain Energy Rev 2018; 91:1065–76.
[19] Xu X, Liu Z, Liu W, Pei C, Wu X, Nie Z. A sustainable development benchmarking framework for energy companies based on topic mining and knowledge graph: the case of oil and gas industry. Renew Sustain Energy Rev 2024;196:114350.
[20] Azam A, Ahmed A, Kamran MS, Hai L, Zhang Z, Ali A. Knowledge structuring for enhancing mechanical energy harvesting (MEH): an in-depth review from 2000 to 2020 using CiteSpace. Renew Sustain Energy Rev 2021;150:111460.
[21] Brundage MP, Sexton T, Hodkiewicz M, Dima A, Lukens S. Technical language processing: unlocking maintenance knowledge. Manuf Letters 2021;27:42–6.
[22] Li J, Sun A, Han J, Li C. A survey on deep learning for named entity recognition. IEEE Trans Knowl Data Eng 2020;34(1):50–70.
[23] Sun C, Yang Z, Wang L, Zhang Y, Lin H, Wang J. Deep learning with language models improves named entity recognition for PharmaCoNER. BMC Bioinform 2021;22(1):1–16.
[24] Kadari R, Zhang Y, Zhang W, Liu T. CCG supertagging via Bidirectional LSTM-CRF neural architecture. Neurocomputing 2018;283:31–7.
[25] Zhong J, Xuan Z, Wang K, Cheng Z. A BERT-Span model for Chinese named entity recognition in rehabilitation medicine. PeerJ Comput Sci. 2023;9:e1535.
[26] Zheng X, Wang B, Zhao Y, Mao S, Tang Y. A knowledge graph method for hazardous chemical management: ontology design and entity identification. Neurocomputing 2021;430:104–11.
[27] Sharma R, Morwal S, Agarwal B, Chandra R, Khan MS. A deep neural network-based model for named entity recognition for Hindi language. Neural Comput Appl 2020;32:16191–203.
[28] Shen H, Cao H, Sun G, Chen D. Research on Chinese semantic named entity recognition in marine engine room systems based on BERT. J Mar Sci Eng 2023;11 (7):1266.
[29] Li H, Soares CG, Huang H-Z. Reliability analysis of a floating offshore wind turbine using Bayesian Networks. Ocean Eng 2020;217:107827.
[30] Cao J, Qin Z, Ju Y, Chen Y, Shen WZ, Shen X, et al. Study of air compressibility effects on the aerodynamic performance of the IEA-15 MW offshore wind turbine. Energy Convers Manag 2023;282:116883.
[31] Liu Y. The effect of vertical arrangement on performance and wake characteristics of two tandem offshore wind turbines under various operating conditions. Energy Convers Manag 2023;278:116743.
[32] Hung JC, Chang J-W. Multi-level transfer learning for improving the performance of deep neural networks: theory and practice from the tasks of facial emotion recognition and named entity recognition. Appl Soft Comput 2021;109:107491.
[33] Liu P, Guo Y, Wang F, Li G. Chinese named entity recognition: the state of the art. Neurocomputing 2022;473:37–53.
[34] Francis S, Van Landeghem J, Moens M-F. Transfer learning for named entity recognition in financial and biomedical documents. Information 2019;10(8):248.
[35] MulDA: a multilingual data augmentation framework for low-resource cross-lingual NER. In: Liu L, Ding B, Bing L, Joty S, Si L, Miao C, editors. Proc annual meeting assoc comput linguistics and int joint conf nat lang process, vol. 1. Long Papers); 2021.
[36] Nguyen T, Nguyen D, Rao P. Adaptive name entity recognition under highly unbalanced data. 2020. arXiv preprint, https://arxiv.org/abs/2003.10296.
[37] Alfred R, Leong LC, On CK, Anthony P. Malay named entity recognition based on rule-based approach. Int J Mach Learn Comput 2014;4(3).
[38] Nie Y, Tian Y, Wan X, Song Y, Dai B, editors. Named entity recognition for social media texts with semantic augmentation. Proc conf empirical methods nat lang process (EMNLP); 2020.
[39] Wang X, Wang Y, editors. Sentence-level resampling for named entity recognition. Proc Conf North Amer Chapter Assoc Comput linguistics: Hum Lang Technol; 2022.
[40] Wu H, Chen P, Li W, Dai Y, Jiang C, Li J, et al., editors. Construction of power communication network knowledge graph with BERT-BiLSTM-CRF model based entity recognition. 2021 IEEE int conf comput commun syst (ICCCS). IEEE; 2021.
[41] Guo L, Yan F, Li T, Yang T, Lu Y. An automatic method for constructing machining process knowledge base from knowledge graph. Robot Comput-Integr Manuf. 2022;73:102222.
[42] Domain adaptation of rule-based annotators for named-entity recognition tasks. In: Chiticariu L, Krishnamurthy R, Li Y, Reiss F, Vaithyanathan S, editors. Proc conf empirical methods nat lang process; 2010.
[43] Wu Q, Lin Z, Karlsson B, Lou J-G, Huang B, editors. Single-/Multi-source cross-lingual NER via teacher-student learning on unlabeled data in target language. Proc Annual Meeting Assoc Comput Linguistics 2020:6505–14.
[44] Sinha Y, Steel JA. A progressive study into offshore wind farm maintenance optimisation using risk based failure analysis. Renew Sustain Energy Rev 2015;42: 735–42.
[45] Arabian-Hoseynabadi H, Oraee H, Tavner P. Failure modes and effects analysis (FMEA) for wind turbines. Int J Electr Power Energy Syst 2010;32(7):817–24.
[46] Bharatbhai MG. Failure mode and effect analysis of repower 5M wind turbine. Int J Adv Res Sci Eng Technol 2015;2(5):2394–444.
[47] Kang J, Sun L, Sun H, Wu C. Risk assessment of floating offshore wind turbine based on correlation-FMEA. Ocean Eng 2017;129:382–8.
[48] Du M. An improved FMECA method for wind turbines health management. Energy Power Eng 2017;9(4):36.
[49] Tazi N, Châtelet E, Bouzidi Y. Using a hybrid cost-FMEA analysis for wind turbine reliability analysis. Energies 2017;10(3):276.
[50] Li H, Teixeira AP, Soares CG. A two-stage Failure Mode and Effect Analysis of offshore wind turbines. Renew Energ 2020;162:1438–61.