Towards the Discrimination of Primary and Secondary Headache: An Intelligent Systems Approach

Robert Keight*, Dhiya Al-Jumeily*, Abir Jaafar Hussain*, Mohammed Al-Jumeily†, Conor Mallucci‡

*Applied Computing Research Group,
School of Computing and Mathematical Sciences,
Liverpool John Moores University,
Byrom Street, Liverpool, L3 3AF, UK
R.Keight@2015.ljmu.ac.uk, {D.Aljumeily, A.Hussain}@ljmu.ac.uk
†Selman Al Habib Hospital, Dubai, UAE
maljumaily@yahoo.fr
‡Department of Neurosurgery,
The Royal Liverpool Childrens NHS Trust,
Alder Hey, Liverpool
cmallucci@me.com

Abstract—We consider the use of intelligent systems to address the long-standing medical problem of diagnostic differentiation between harmful (secondary) and benign (primary) headache conditions. In secondary headaches, the condition is caused by an underlying pathology, in contrast to primary headaches where the production of pain represents the sole constituent of the disorder. Conventional diagnostic paradigms carry an unacceptable risk of misdiagnosis, leaving patients open to potentially catastrophic consequences. Intelligent systems approaches, grounded in artificial intelligence, are adopted in this study as a potential means to unite contributions from multiple settings, including medicine, the life sciences, pervasive computation, sensor technologies, and autonomous intelligent agency, in the fight against headache uncertainty. In this paper, we therefore present the first steps in our research towards a data intensive, unified approach to headache dichotomisation. We begin by presenting a background to headache and its classification, followed by analysis of the space of confounding symptoms, in addition to the problem of primary and secondary condition discrimination. Finally, we proceed to report results of a preliminary case study, in which the epileptic seizure is considered as a manifestation of a headache generating neuropathology. It was found that our classification approach, based on supervised machine learning, represents a promising direction, with a best area under curve test outcome of 0.915. We conclude that intelligent systems, in conjunction with biosignals, could be suitable for classification of a more general set of pathologies, while facilitating the medicalisation of arbitrary settings.

I. INTRODUCTION

Forms of head pain have troubled human populations as far into the past as the historical record extends [1]–[6]. Today, known by the medical name Cephalalgia, headache is defined as pain occurring anywhere in the region of the head or the neck [7], representing a range of conditions that cast a considerable burden on societies and individuals [8]–[11]. It is estimated that around 50% of the global population are directly affected by headache, with an economic impact of £600 million per annum in the UK and between $5.6 and $17.2 billion in the US [12]. A broad distinction is drawn between very common benign (primary) headaches and relatively rare, potentially serious (secondary) types of headache, resulting from underlying pathologies. A formidable and enduring challenge within today's healthcare setting is to determine which cases represent conditions that are painful and possibly debilitating, yet benign, and those which pose a serious threat to patient health due to an underlying pathology. Despite significant advances in both sciences relating to living systems and in medical technologies, established procedures for headache diagnosis remain complex and error prone, exhibiting significant inter-observer variability, and often leaving patients with considerable anxiety. Consequently, the risk of misdiagnosis within populations remains moderate, with potentially catastrophic consequences for individual patients. The head itself constitutes a particularly essential anatomical component of the human system, home to key functionality centers, such as the sensory modalities of sight, hearing, smell, and taste, and not least our centre of our consciousness, personhood, and vital regulatory function, the brain. Unsurprisingly, disruption to such an intricate and highly sensitive structure can lead to immediate and irreversible consequences to the health of individuals, especially where loss of brain function occurs [13]. However, formations of symptoms that may indicate the presence of serious problems are confounded by a space of very common benign headache types, known as primary headaches, for instance tension type headache, migraine, cluster headache, that exhibit a significant range of symptom expressions, overlapping in semblance

with the more serious conditions. The high prevalence of primary headache within populations renders it impractical within current healthcare frameworks to bring economically costly diagnostic practices to each individual headache patient; instead an emphasis is placed on clinical examination to select cases only where deviations from known benign headache prototypes are deemed evident. In this research, we report on the potential and challenges of transforming the healthcare posture for headache diagnostics through the use of intelligent systems pathways that may disrupt existing practices. In particular, we focus on the proposition that mass data from physiological signals can be harnessed through the use of artificial intelligence techniques to shift both the accuracy and availability of diagnostics means, in addition to furthering the capacity to observe and characterise headache causing conditions. The rest of this paper is therefore organised as follows. In Section II we review the background to headache classification, following by Section III in which the space of conflicting headache symptoms is considered. In Section IV we discuss the characterisation of headache discrimination, following which we report our progress towards empirical analysis, in the form of a case study focusing on the classification of eplileptiform events, located in Section V. Finally, we complete the paper in Section VI, in which conclusions and future work are presented.

II. BACKGROUND TO HEADACHE CLASSIFICATION

The International Classification of Headache Disorders (ICHD) specification represents the globally accepted standard in headache classification, regarded by both researchers and medical practitioners as the leading authority. The introduction of this model by the International Headache Society (IHS), beginning in 1998, has served to rationalise the diagnosis and treatment of headaches, providing a standard set of terminology and operational diagnostic criteria that has contributed significantly to the accuracy and consistency of headache diagnosis and treatment [14]. The World Health Organisation has recognised the ICHD headache classification, incorporating criteria into the International Classification of Diseases 10 (ICD 10) definition [15]. The progression of ICHD models, in order of succession, comprises editions ICHD-I (1988) [16], ICHD-II (2004) [17], still in wide use, and more recently ICHD-III-beta (2013) [18]. The ICHD-II model provides an evidence based update to the ICHD-I, which is based mainly on expert opinion. The ICHD model editions were preceded by earlier attempts to classify headaches, with works including Brown in 1951 [19], followed by an Ad-hoc model by introduced by the US organisation the National Institute of Health (NIH) in 1962 [20]. Early models, however, did not encompass the full range of headaches and did not gain wide acceptance. The modern era of headache classification is further preceded by historical interpretations [21], including work by Christian Baur in 1787, in which a total of 84 categories are described and broadly divided into "idiopathic" and "symptomatic", reflecting the notion of primary and secondary headaches, in addition to the work of Thomas Willis, recognised as a founder of modern neurological thought [22]. The depth of history surrounding headache conditions highlights both the enduring and elusive nature of the phenomenon, for which further review can be found in [23]–[25].

III. THE SPACE OF CONFLICTING SYMPTOMS

Owing to the complexity of the human biological system, within which a wide range of potential disorders may correspondingly (co)exist, manifestations of differing headache disorders have the flexibility to strongly resemble one another, especially at the level of clinical symptoms. A danger therefore exists that a serious underlying condition may be mistaken for a benign primary headache, with potentially critical consequences for the patient. This phenomenon of "false friends" in the presentation space has been acknowledged in the literature, where the constellation of reports collectively highlights the complex nature of accurately linking clinical symptoms and signs to true underlying conditions. Sentinel headache (SH), as reported in [26], [27], is an early warning sign of the more serious Subarachnoid Haemorrhage (SAH), of which milder presentations are possible that may resemble primary headache types in terms of clinical features. Transient Ischemic attacks are another significant form of intracranial aetiology, where a subspace of presentations that resemble migrainous features are possible [28]. Cardiac Cephalgia may also be misdiagosed as migraine, as reported in [29], in addition to cerebellar infarction [30].

An abnormality of intracranial pressure is a further significant aetiology in patients, reported in [31] to mimic the features of cluster headache in the case of idiopathic intracranial hypertension. Furthermore, following mild head injury, acute post-traumatic headache may clinically resemble a range of primary headache types, including Migraine, Probable migraine, Tension type headache (TTH), Probable TTH, or unclassified primary headache [32]. Additional intracranial pathologies add to the list of primary headache mimics, including confounding aspecific headaches [33]. The epilepsies are also a group of disorders that may give rise to presentations that resemble primary headaches. Occipital lobe epilepsy for example is known in certain forms to exhibit symptoms that may be misdiagnosed as migraine [34]–[36]. Non-epileptic neurodisorders can also trigger a misdiagnosis of migraine due to coinciding patterns of symptoms, for instance Huete et al report on Hashimoto's encephalopathy in [37]. Secondary headaches driven by cervical and ocular pathologies give rise to additional symptom groups that reflect primary headache prototypes such as migrain and TTH, examples of such scenarios include cervical spine pathologies [38], [39] and subacute angle-closure glaucoma [40], [41]. Toxicological effects must also be considered as a source of various headache symptoms, for example carbon monoxide poisoning is known to resemble migrainous symptoms as reported in [42]. Further unspecified pathologies are likely to exist in addition to those covered here, continuing the spectrum of serious disorders that may appear benign.

IV. The Problem of Secondary Headache Discrimination

The problem of secondary headache diagnosis does not admit a well defined entry point or a single sufficient solution element. The compact term "secondary headache" in fact refers to any entity or subset within the space of all possible diseases and disorders from which headache may consequently result. The problem is therefore highly heterogeneous, rooted in a large body of pathologies, and is resistant to the idea of an overarching unifying principle. The symptoms of any one inclusive disorder are often themselves heterogeneous, sensitive to initial conditions, and may interrelate in complex ways with further coexisting disorders. A key objective for front line physicians, who represent the first point of contact with presenting patients, is to answer the question, "is there an underlying condition present in this patient?". Addressing this question is crucial, since more costly tests, for example neuroimaging, cannot be applied a priori without compromising the economic viability of the healthcare system as a whole [43], [44]. Such a question should, in light of emerging healthcare paradigms [45]–[47], be widened in scope to include additional concerns, in particular the detection of pre-symptomatic disorders and the extension of diagnostic capacity to operate as a non-disruptive agency within the course of patient lives.

V. Initial Research Progress: Epileptic Seizure Classification

To begin our investigation into secondary headache analysis, we present a case study through which we examine the potential of machine learning for the analysis of biosignals originating from the Electroencephalographam (EEG). In particular, we focus upon supervised classification, defined as the ability to learn a functional mapping between a feature space of arbitrary dimension and a set of discrete classes, giving the form $\hat{y} = \hat{f}(x)$, $x \in \mathbb{R}^n$, $\hat{y} \in \{0, 1\}$. We set as an initial objective the classification of seizure activity in Epilepsy, a well documented condition that is known to present in many cases with the symptoms of headache [48], [49]. The epilepsies are a group of chronic, heterogeneous neurological disorders, affecting an estimated 1% of the worlds population [50], characterised by recurrent, intermittent, involuntary, paroxysmal seizure episodes [51]. The purpose of this pilot study is to establish a proof of principle result, within the parameters of our intelligent systems approach, so as to demonstrate the viability to detect and potentially therefore predict a headache generating condition. In similar scenarios, it is widely documented that further headache producing conditions, such as brain tumour aetiologies, may also often lead to anomalies of neuronal activity that consequently affect the EEG and neuroimaging recordings [52]. A natural next phase in this research is therefore to apply the methodology developed for this case study to a more general selection of headache generating scenarios. As such, the methodology used in this case study is designed to provide an initial basis to suitably justify the continued exploration of the limits and generalisability of biosignal analysis for the secondary headache inquiry, using epilepsy as a first pilot example.

A. Scalp EEG Data Set

To empirically evaluate the effectiveness of our approach, we make use of an open dataset of scalp EEG recordings obtained from the Physionet project (listed as CHB-MIT) [53], [54]. The data consists of the digitised EEG segments collected from a total sample of 23 subjects, 22 of which were observed at the Childrens Hospital in Boston, with a further subject whose details are unknown. The known subjects range in age from 1.5 to 22 years. All of the subjects under study, who were admitted for refractory epilepsy, had anticonvulsant medication withdrawn prior to monitoring, to facilitate the observation of seizure events. During the recording process, 181 events occurred that were identified as clinical seizures, as determined by expert opinion. The experts provided annotation of the earliest and latest EEG state transitions associated with each seizure event. The resulting annotations are adopted as ground truth labels within this study. The output from the EEG activity records is organised according to 24 cases (1 subject appears in two cases, recorded 1.5 years apart), where each case reflects the observation of a single subject over time. Within each case, the EEG output is distributed over a number of consecutive files, each generally containing an hour in duration of uninterrupted data. Time gaps exist at each file boundary, most reportedly enduring for around 10 seconds. It is stated that such interruptions were introduced at the time of recording due to hardware limitations of the equipment in use.

The EEG data were recorded according to the 10-20 international standard, with the exception of 17 files with different montages. Consequently, the sensor space consists of 19 electrodes and 1 ground attached to the surface of the scalp, resulting in simultaneous output from 23 spatially associated channels via differential amplification, namely FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FZ-CZ, CZ-PZ, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8, P8-O2, P7-T7, T7FT9, FT9-FT10, FT10-T8, and T8-P8. Recording equipment was used to sample channel output at a rate of 256 Hz, with 16 bit resolution, for all observations. Uninterrupted file durations are generally 1 hour, with the exception of those present in case 10 which are two hours in length, and those of cases 4, 6, 7, 9, and 23, which endure for four hours each. Present in the sample of subjects observed were 5 males, aged between 3 and 22 years, and 17 females, aged between 1.5 and 19 years, in addition to an unknown subject assigned to a single case. The captured data presents 141 record files that contain at least one seizure and 545 record files with no seizure activity, referred to as seizure and non-seizure files respectively. Overall, a total of 982 hours of EEG recording are represented across 686 edf record files, encompassing a total of 181 clinical seizure events.

B. Feature Definitions

We use six feature in our study, as formally defined in Table I. Each distinct feature serves to isolate a signal characteristic of interest by substituting the full granularity of the raw signal representation with a constrained computational product. Such computed features represent an estimation of the true properties of the underlying dynamical system, in this case neuroelectric brain activity, where the closeness of correspondence to the original system is bounded by the duration, resolution, and quality of the time series measured for such a system. Our feature set covers spectral characteristic estimators, signal strength estimators, and statistical moments, all of which are univariate measures. Peak frequency (PF) is used to compute the frequency component of each signal that satisfies the global maxima when viewed from the perspective of the signals power spectral density (PSD). A related location measure is the median frequency (MF), providing an estimate of the typical signal frequency, while showing robustness against outlying frequency components. Root Mean Square (RMS) is also used in this study as a means to estimate the strength of the signal over a given time domain interval. The RMS provides a summary of the magnitude of the signal. We consider next a signal strength estimator, the Signal Energy (SE). SE measures the sum of the squared magnitude of the samples taken over a given signal interval. Finally, we consider the statistical moment estimators the Skew and Kurtosis, defined respectively as measuring the symmetry and peakedness of the varying amplitude of signal in the time domain.

TABLE I: Feature Definitions

Index	Feature	Name	Definition
1.	PF	Peak Frequency	$f_{max} = \arg(\frac{f_s}{n} max_{i=0}^{N-1} p(i))$
2.	MF	Median Frequency	$f_{med} = i_m \frac{f_s}{N}, \sum_{i=0}^{i=i_m} P(i) = \sum_{i=i_m}^{i=N-1}$
3.	RMS	Root Mean Square	$RMS = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} x(i)^2}$
4.	E	Signal Energy	$E = \sum_{k=1}^{N} X_k^2$
5.	SKW	Skew	$s = \frac{E(X-\mu)^3}{\sigma^3}$
6.	KRT	Kurtosis	$k = \frac{E(X-\mu)^4}{\sigma^4}$

C. Performance Measurement

In order to evaluate the capability of our classifiers, we introduce a framework of performance measures, which are evaluated for each of the models tested. We used models purposely chosen from different theoretical classes, to provide points of reference in performance. We reason that to be justified, our test classifiers should significantly outperform both linear classes of models and also simplistic models such as K-nearest neighbour (weak learners). Additionally, to demonstrate that the data we used as input contains true dependency, we compared the performance of all of the models presented with a random guessing baseline, serving as a calibration point, to show that uninformed decisions are insufficient to produce significant results. Consequently, the informed models are shown to synthesise patternful information from the data. Additionally, since generalisation is the goal of our classifier, we apply the described performance framework to both the training (in-sample) and test (out-of-sample) results. Such a comparison of results between the training and test procedures provides an indication of any overfitting that may have occurred during model parameter optimisation, in addition to demonstrating the generalisation capability of each classifier, through the correspondence of both sets of results. To furnish the classifier responses with objective measures of performance, we utilise both graphical methods and scalar summaries, as listed in Table II.

TABLE II: Performance Metrics

Metric	Abbr.	Computation	Range
Area Under Curve	AUC	$0 \leqslant area(ROC) \leqslant 1$	[0,1]
Sensitivity	SEN	TP/(TP+FN)	[0,1]
Specificity	SPEC	TN/(TN+FP)	[0,1]
Precision	PRE	TP/(TP+FP)	[0,1]
F1 Score	F1	2·(PRE·RC)/(PRE+RC)	[0,1]
Youden's J Statistic	J	Sensitivity + Specificity - 1	[-1,1]
Accuracy	ACC	(TP+TN)/(TP+FN+TN+FP)	[0,1]

TP = True Positive Count

TN = True Negative Count

PRE = Precision

RC = Recall

ROC = Receiver Operating Characteristic

D. Exploratory Analysis

Prior to the development and training of classifier models, we considered visualisations of the dataset, such that the structure could be approximated via a mapping into two dimensional graphical form. Results are obtained for both Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbourhood Embedding (t-SNE) [55] procedures, as shown in Figures 2 and 3, respectively. Additionally, we viewed the EEG montages to gain visual insight into the nature of the signals, an EEG montage exemplar is shown in Figure 1. The exploratory results indicate that the signals are complex and exhibit no clear discrimination in the two dimensional setting, especially in the PCA domain, justifying the classification approach via machine learning procedures.

1) Overview of Main Experimental Setup: To evaluate the capability of our proposed classifiers, we conducted a series of empirical simulations, using the CHB-MIT EEG signal
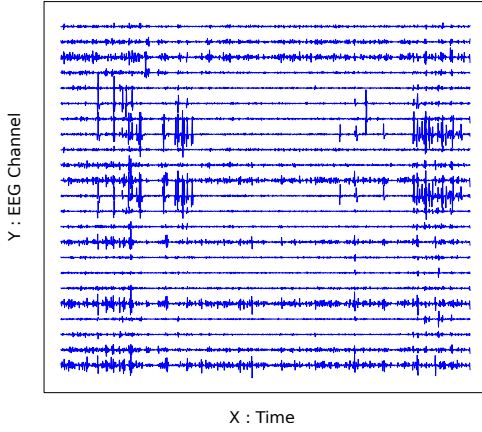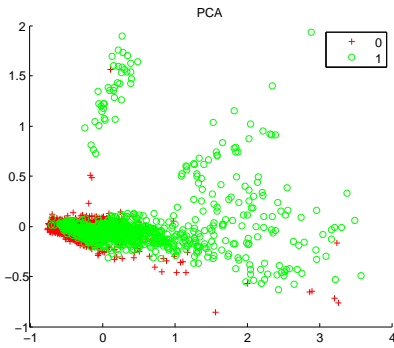
Fig. 1: EEG Montage Exemplar

Fig. 2: Principal Component Analysis Plot
0 = no seizure, 1 = seizure

Fig. 3: tSNE Plot
0 = no seizure, 1 = seizure

database as our sample. The models used in our experiment are listed in Table III. We provide details of our analytical parameters in the sections immediately following, where aspects of the procedural setup are organised into loosely orthogonal concerns.

2) Feature Origination: Six features were extracted from 10 second discretised quasi-stationary segments of the original time domain signals, using calculations as defined in Section V-B. The feature calculations were applied to all 23 EEG time domain channels, resulting in a total of 138 features for each unit in the discretised time domain (later reduced using PCA to 35 mapped features).

3) Data Holdout Procedure: The trials presented in this work were conducted using a data set partitioning scheme, reserving 70% of the data for model training and the remaining 30% as hold out for testing.

4) Simulation Setup: For the simulation of our models we undertook a data loading phase, followed by a primary model execution phase. Firstly, the prepared features matrix was loaded and PCA applied to map the original feature set to a lower dimensional space, by capturing the first 35 principal components. The purpose of this initial mapping operation is
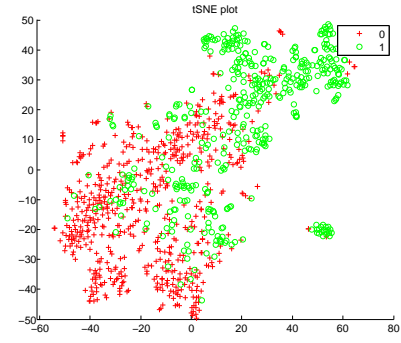
to improve the training example to feature ratio and also to discard noise components, to promote better generalisation. The models to be evaluated were simulated 30 times, following which the means of results from individual trials for each model were taken and used as a basis to derive performance metrics.

5) Models and Trial Definitions: As shown in Table III, we pose 11 models for evaluation. Each model was presented with the set of features derived from EEG signal data, as described in previous sections. Each trial in our experiment corresponds to an entry listed in the table.

E. Results

The results from our experimental procedure are presented and organised in the following listing, where classifier performance is evaluated over the test holdout set from our data. Table IV lists the results for each respective classifier, according to the performance metrics we introduced previously. We then proceed to present our evaluation of the classifiers according to ROC curve analysis, where AUC and ROC curve plots are considered. Figure 4 shows a bar plot for the AUC values for each classifier, while Figure 5 presents the corresponding ROC curves.
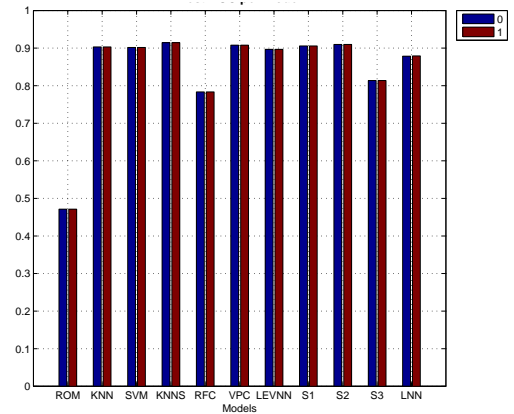

Fig. 4: AUC Plots for the Test Set
0 = no seizure, 1 = seizure

TABLE III: Models

No.	Model	Description	Architecture	Training Algorithm	Role
1.	ROM	Random Oracle Model	Pseudo-random Number Generator	NA	RAND
2.	LNN	Linear Neural Network	35 Units, Linear Activations	Batch training with weight and bias learning rules	LB
3.	KNN	K-Nearest Neighbour	3 Neighbours	Instance Induction	WNLB
4.	KNNS	K-Nearest Neighbour Stacked	Stacked KNN	Stacked Instance Induction	TC
5.	SVM	Support Vector Machine	Matrix Kernel	Quadratic Optimisation	TC
6.	RFC	Random Forest Classifier	Decision Tree Ensemble	Random feature subset bagging	TC
7.	VPC	Voted Perceptron Classifier	2054 Perceptron Ensembles	Delta Rule	TC
8.	LEVNN	Levenberg Neural Network	Units: 35-28-2	Levenberg-Marquardt	TC
9.	S1	Stacked Generaliser 1	4 LEVNN, with hidden layer neurons: 2,5,10,20	Hybrid	TC
10.	S2	Stacked Generaliser 2	LEVNN 35-20-2, LEVNN 35-10-2, VPC 35-20-2, VPC 35-30-2	Hybrid	TC
11.	S3	Stacked Generaliser 3	LEVNN 35-20-2, VPC 35-20-2, RFC	Hybrid	TC

Key:

RAND = Random Baseline

LB = Linear Baseline

WNLB = Weak Non-linear Baseline

TC = Test Classifier

TABLE IV: Results (test holdout, seizure class)

Model	Sensitivity	Specificity	Precision	F1	J	Accuracy	AUC
ROM	0.548	0.467	0.412	0.47	0.0154	0.5	0.471
LNN	0.819	0.832	0.768	0.793	0.651	0.827	0.879
KNN	0.789	**0.889**	**0.829**	0.809	0.679	0.849	0.903
KNNS	0.837	0.828	0.768	0.801	0.665	0.832	**0.915**
SVM	0.843	0.861	0.805	0.824	0.704	0.854	0.902
RFC	0.705	0.803	0.709	0.707	0.508	0.763	0.783
VPC	**0.861**	0.82	0.765	0.81	0.681	0.837	0.908
LEVNN	0.849	0.828	0.77	0.808	0.677	0.837	0.897
S1	0.825	0.881	0.825	**0.825**	**0.706**	**0.859**	0.906
S2	0.843	0.836	0.778	0.809	0.679	0.839	0.91
S3	0.807	0.746	0.684	0.74	0.553	0.771	0.814

F. Discussion

Prior to consideration of the test classifiers, we first observe through the results that the uninformed guessing model (ROM) is unable to discriminate better than chance between the two classes, providing a suitable benchmark against which the significance of classifier performances can be gauged. Furthermore, the establishment of the non-discriminative random performance provides an indication that no systematic bias is
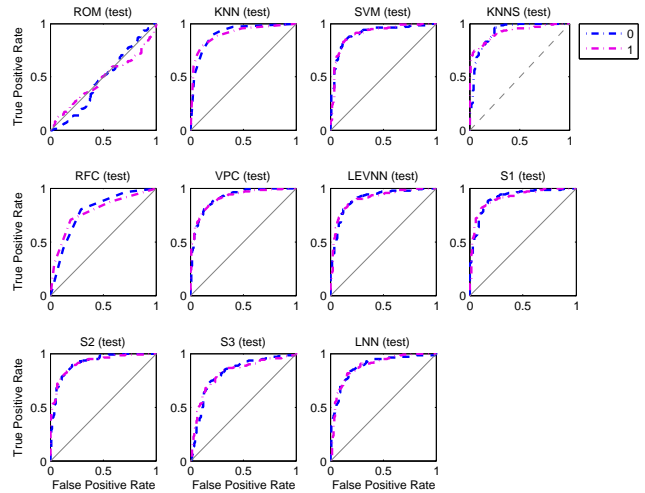

Fig. 5: ROC Plots for the Test Set
0 = no seizure, 1 = seizure

present as a function of the data itself. Subsequently, evident in the classifier results is the capability of the linear baseline, the LNN, to achieve performance levels close to those of the other informed models of non-linear expressive capacities. For example, the LNN achieved an AUC of 0.902 for the training set, and 0.879 for the testing set respectively, in comparison to 0.911 and 0.897 for the LEVNN, a non-linear model with universal approximation capability [56]. Nearly all of the informed classifiers obtained AUC values near to 0.9 for the test set evaluation, with the exceptions of the RFC and S3 models, which appear to have overfit the training set. The KNNS classifier is also indicated for overfitting, having achieved ideal results during training and dropping in performance upon exposure to the test set. The increase in expressive power added to the KNN base learners through stacking has not improved the models capacity to synthesise useful information. Overall, the S1 model appears to have performed marginally better than the other models, though performance over the informed models, with the exception of RFC and S3, appears similar. The KNN, although having achieved an AUC of 0.903 for the test set, has suffered a lower sensitivity of 0.789, due to the geometry of the ROC curve; it can be seen in Figure 5 that for the seizure class the curve falls away from the maxma of the sensitivity axis (Y) to a greater extent than from the minima of the false positive rate axis (X), resulting in the observed bias towards specificity.

When considering the results from the perspective of relative standing, a lack of a clear separation in discriminative performance between the non-linear classes of model versus that of the LNN baseline is evident, extending also to less complex non-linear models such as the KNN. The lack of distinction between the non-linear versus the linear results suggests that the current set of features may have failed to adequately capture relevant non-linear components from

the original signal, since non-linear capacity appears to offer very little advantage for the experimental trials conducted. There are clearly, however, informative linear components within the data, since all informed classifiers were able to perform significantly beyond that of the ROM baseline. It is possible that excessive granularity has been discarded form the original EEG signals, indicating that further investigation of the original signals may be of benefit to improve classification performance outcomes. Additionally, it is also a possibility that the current feature set may be closely coupled to variance of the signal, rather than acquiring a deeper structure in the signal, since seizure events tend towards greater signal volatility. A further source of performance impairment may be a remainder of artefacts and noise sources embedded within the data, disrupting the discriminative capability of the classifiers. We therefore recommend for future work relating to this experiment that the space of features is further interrogated and expanded, including examination of the discrete time intervals over which the features are extracted, that a surrogate approach is included to account for the proportion of excess performance due to signal variance, and that further preprocessing techniques and their implications are assessed to improve the signal to noise ratio prior to the extraction of features.

VI. CONCLUSION AND FUTURE WORK

In this work we have considered the enduring medical problem of diagnostic differentiation between primary and secondary forms of headache, the misdiagnosis of which may result in immediate and irreversible consequences for the patient. Both primary and secondary headache terms are seen not to comprise a single entity, but in fact encompass a highly complex, heterogeneous space of conditions, within which symptoms relating to causes of markedly differing origins may exhibit significant overlap. Costly specialist procedures such as CT and MRI neuroimaging, while providing rich internal views of the human biological system, do not represent a sustainable, scalable means of diagnosis, since they can be applied only in cases where a priori indication of a problem exists. We therefore presented a strategy grounded in the hypothesis that data intensive biosignals analysis, originating from increasingly available sensor technologies, may be enabled through an intelligent systems methodology, giving rise to a deepened scope of analysis with the necessary operational features for use over arbitrary diagnostic junctures. Moreover, we recognise that the complexity inherent in bio-domain analysis in fact necessitates the extrication of low level human input, since considerable complexity must be handled within a restrictive time frame. The use of intelligent systems therefore permits a level of utility normally provided through the work of multidisciplinary teams, to be combined into a single unit of operation, driven by a marginal resource footprint. To demonstrate the potential of our proposed strategy, we reported a preliminary experiment, in which the Epilepsies, a group of paroxysmal neurological disorders, known to result in headaches, were considered within the framework of a classi-

fication problem. Using supervised machine learning, the time series of EEG signals were segmented and analysed, with the resulting outcomes classified as either seizure or no-seizure. It was shown that a maximum AUC of 0.915 was achieved by a stacked KNN model (KNNS). Further results were consistent with this finding, with models SVM, KNN, VPC, S1, and S2 achieving close to an AUC of 0.9. Generalising this approach to further pathologies, we propose that the dichotomisation of primary and headaches can be achieved through continued development of the strategy presented.

In future work we intend to expand on our current progress to identify and operationalise further points of difference between primary headache scenarios and secondary headache causing conditions, exploring also an expanded range of both biosignals and analytical techniques with which to exploit such discriminative opportunities. In particular, we aim to test key pathological conditions such as subarachnoid haemorrhage, forms of brain tumour, infection, traumatic brain injury conditions, raised intracranial pressure, in addition to further seizure conditions. In terms of biosignals, we consider the use of emerging intracranial biosignal modalities that hold promise for portable solutions, such as microwave imaging [57], magneto-acoustic imaging and detection (MAI) [58], ultrasound methods such as Transcranial Doppler ultrasonography (TCD) [59], and near infrared imaging (NIR) [60], in addition to alternatives to EEG such as Magnetoencephalography (MEG). Moreover, we consider the use of deep learning algorithms to represent a key direction in analysis, allowing arbitrarily complex signal combinations to be evaluated without the bottleneck of feature hand engineering [61], [62].

REFERENCES

[1] R. Assina, C. E. Sarris, and A. Mammis, "The history of craniotomy for headache treatment," *Neurosurgical focus*, vol. 36, no. 4, p. E9, 2014.

[2] A. P. Friedman, "The headache in history, literature, and legend." *Bulletin of the New York Academy of Medicine*, vol. 48, no. 4, p. 661, 1972.

[3] L. Bayan, S. M. Modarres Mousavi, and A. Gorji, "History of neurological disorders in persian medicine," *Journal of Research on History of Medicine*, vol. 2, no. 4 Nov, 2013.

[4] A. Gorji and M. K. Ghadiri, "History of headache in medieval persian medicine," *The Lancet Neurology*, vol. 1, no. 8, pp. 510–515, 2002.

[5] M. J. Eadie, *Headache: through the centuries*. OUP USA, 2012.

[6] S. Collado-Vázquez and J. Carrillo, "Cranial trepanation in the egyptian," *Neurología (English Edition)*, vol. 29, no. 7, pp. 433–440, 2014.

[7] J. La Jolla Donald *et al.*, *Wolff's headache and other head pain*. Oxford University Press, 2001.

[8] P. Pop, C. Gierveld, H. Karis, and H. Tiedink, "Epidemiological aspects of headache in a workplace setting and the impact on the economic loss," *European Journal of Neurology*, vol. 9, no. 2, pp. 171–174, 2002.

[9] B. K. Rasmussen, R. Jensen, M. Schroll, and J. Olesen, "Epidemiology of headache in a general populationa prevalence study," *Journal of clinical epidemiology*, vol. 44, no. 11, pp. 1147–1157, 1991.

[10] L. J. Stovner, J.-A. Zwart, K. Hagen, G. Terwindt, and J. Pascual, "Epidemiology of headache in europe," *European journal of neurology*, vol. 13, no. 4, pp. 333–345, 2006.

[11] L. Stovner, K. Hagen, R. Jensen, Z. Katsarava, R. Lipton, A. Scher, T. Steiner, and J.-A. Zwart, "The global burden of headache: a documentation of headache prevalence and disability worldwide," *Cephalalgia*, vol. 27, no. 3, pp. 193–210, 2007.

[12] C. Clarke, L. MacMillan, S. Sondhi, and N. Wells, "Economic and social impact of migraine." *Qjm*, vol. 89, no. 1, pp. 77–84, 1996.

[13] T. Wieloch and K. Nikolich, "Mechanisms of neural plasticity following brain injury," *Current opinion in neurobiology*, vol. 16, no. 3, pp. 258–264, 2006.

[14] K. Ravishankar, "The ihsclassification (1988, 2004)–contributions, limitations and suggestions," *J Assoc Physicians India*, vol. 58, pp. 7–9, 2010.

[15] I. H. C. Committee *et al.*, "Icd-10 guide for headaches," *Cephalalgia*, vol. 17, no. suppl 19, p. 32, 1997.

[16] H. C. C. of the International Headache Society *et al.*, "Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain," *Cephalalgia*, vol. 8, pp. 1–96, 1988.

[17] J. Olesen and T. Steiner, "The international classification of headache disorders, 2nd edn (icdh-ii)," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 75, no. 6, pp. 808–811, 2004.

[18] H. C. C. of the International Headache Society (IHS *et al.*, "The international classification of headache disorders, (beta version)," *Cephalalgia*, vol. 33, no. 9, pp. 629–808, 2013.

[19] M. R. Brown, "The classification and treatment of headache." *The Medical clinics of North America*, vol. 35, no. 5, pp. 1485–1493, 1951.

[20] A. P. Friedman, K. H. Finley, J. R. Graham, E. C. Kunkle, A. M. Ostfeld, and H. G. Wolff, "A classification of headache ad hoc committee on classification of headache," *Neurology*, vol. 12, no. 5, pp. 378–378, 1962.

[21] M. Levin, *Comprehensive review of headache medicine*. Oxford University Press, 2008.

[22] H. Isler, "Thomas willis', two chapters on headache of 1672: A first attempt to apply the new science to this topic," *Headache: The Journal of Head and Face Pain*, vol. 26, no. 2, pp. 95–98, 1986.

[23] ——, "Headache classification prior to the ad hoc criteria," *Cephalalgia*, vol. 13, no. 12, pp. 9–10, 1993.

[24] Á. L. Guerrero-Peral, V. de Frutos González, and M. I. Pedraza-Hueso, "Galeata: chronic migraine independently considered in a medieval headache classification," *The journal of headache and pain*, vol. 15, no. 1, p. 1, 2014.

[25] E. Magiorkinis, A. Diamantis, D.-D. Mitsikostas, and G. Androutsos, "Headaches in antiquity and during the early scientific era," *Journal of neurology*, vol. 256, no. 8, pp. 1215–1220, 2009.

[26] F. De Falco, "Sentinel headache," *Neurological Sciences*, vol. 25, no. 3, pp. s215–s217, 2004.

[27] J. L. B. Pereira, L. A. F. de Albuquerque, M. Dellaretti, G. T. C. de Carvalho, G. V. Junior, M. I. Rocha, L. L. A. Loures, P. P. Christo, and A. A. de Sousa, "Importance of recognizing sentinel headache," *Surgical neurology international*, vol. 3, 2012.

[28] S. Prabhakaran, A. J. Silver, L. Warrior, B. McClenathan, and V. H. Lee, "Misdiagnosis of transient ischemic attacks in the emergency room," *Cerebrovascular Diseases*, vol. 26, no. 6, pp. 630–635, 2008.

[29] Y. Torres-Yaghi, J. Salerian, and C. Dougherty, "Cardiac cephalgia," *Current pain and headache reports*, vol. 19, no. 4, pp. 1–5, 2015.

[30] S. I. Savitz, L. R. Caplan, and J. A. Edlow, "Pitfalls in the diagnosis of cerebellar infarction," *Academic emergency medicine*, vol. 14, no. 1, pp. 63–68, 2007.

[31] M. Volcy and S. Tepper, "Cluster-like headache secondary to idiopathic intracranial hypertension," *Cephalalgia*, vol. 26, no. 7, pp. 883–886, 2006.

[32] D. Lieba-Samal, P. Platzer, S. Seidel, P. Klaschterka, A. Knopf, and C. Wöber, "Characteristics of acute posttraumatic headache following mild head injury," *Cephalalgia*, vol. 31, no. 16, pp. 1618–1626, 2011.

[33] E. Wouda and J. Vanneste, "Aspecific headache during 13 years as the only symptom of idiopathic hypertrophic pachymeningitis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 64, no. 3, pp. 408–409, 1998.

[34] M. Walker, S. Smith, S. Sisodiya, and S. Shorvon, "Case of simple partial status epilepticus in occipital lobe epilepsy misdiagnosed as migraine: clinical, electrophysiological, and magnetic resonance imaging characteristics," *Epilepsia*, vol. 36, no. 12, pp. 1233–1236, 1995.

[35] B. Menon, "Symptomatic occipital epilepsy misdiagnosed as migraine," *Headache: The Journal of Head and Face Pain*, vol. 47, no. 2, pp. 287–289, 2007.

[36] C. Panayiotopoulos, I. A. Sharoqi, and A. Agathonikou, "Occipital seizures imitating migraine aura," *Journal of the Royal Society of Medicine*, vol. 90, no. 5, pp. 255–257, 1997.

[37] A. J. Huete, M. Sanchez-del Rio, and O. Franch, "Hashimoto's encephalopathy mimicking migraine with aura," *Headache: The Journal of Head and Face Pain*, vol. 47, no. 1, pp. 130–131, 2007.

[38] W. Pöllmann, M. Keidel, and V. Pfaffenrath, "Headache and the cervical spine: a critical review," *Cephalalgia*, vol. 17, no. 8, pp. 801–816, 1997.

[39] X. Yi, A. J. Cook, R. J. Hamill-Ruth, and J. C. Rowlingson, "Cervicogenic headache in patients with presumed migraine: missed diagnosis or misdiagnosis?" *The Journal of Pain*, vol. 6, no. 10, pp. 700–703, 2005.

[40] K. S. Shindler, P. S. Sankar, N. J. Volpe, and J. R. Piltz-Seymour, "Intermittent headaches as the presenting sign of subacute angle-closure glaucoma," *Neurology*, vol. 65, no. 5, pp. 757–758, 2005.

[41] D. I. Friedman, L. K. Gordon, and P. A. Quiros, "Headache attributable to disorders of the eye," *Current pain and headache reports*, vol. 14, no. 1, pp. 62–72, 2010.

[42] M. Eberhardt, A. Powell, G. Bonfante, V. Rupp, J. R. Guarnaccia, M. Heller, and J. Reed, "Noninvasive measurement of carbon monoxide levels in ed patients with headache," *Journal of Medical Toxicology*, vol. 2, no. 3, pp. 89–92, 2006.

[43] P. J. Goadsby, "To scan or not to scan in headache: Some patients with primary headaches may need imaging," *BMJ: British Medical Journal*, vol. 329, no. 7464, p. 469, 2004.

[44] B. M. Frishberg, "The utility of neuroimaging in the evaluation of headache in patients with normal neurologic examinations," *Neurology*, vol. 44, no. 7, pp. 1191–1191, 1994.

[45] P. Cuijpers *et al.*, "Prevention: an achievable goal in personalized medicine." *Dialogues in Clinical Neuroscience*, vol. 11, no. 4, pp. 447–454, 2009.

[46] L. Hood and S. H. Friend, "Predictive, personalized, preventive, participatory (p4) cancer medicine," *Nature Reviews Clinical Oncology*, vol. 8, no. 3, pp. 184–187, 2011.

[47] W.-L. Liao and F.-J. Tsai, "Personalized medicine: a paradigm shift in healthcare," *BioMedicine*, vol. 3, no. 2, pp. 66–72, 2013.

[48] R. Ottman and R. B. Lipton, "Comorbidity of migraine and epilepsy," *Neurology*, vol. 44, no. 11, pp. 2105–2105, 1994.

[49] W. R. Gowers, "Clinical lectures on the borderland of epilepsy. iii.migraine," *British medical journal*, vol. 2, no. 2397, p. 1617, 1906.

[50] J. Sander and S. Shorvon, "Epidemiology of the epilepsies," *Journal of Neurology Neurosurgery and Psychiatry*, vol. 61, no. 5, pp. 433–443, 1996.

[51] E. Trinka, H. Cock, D. Hesdorffer, A. O. Rossetti, I. E. Scheffer, S. Shinnar, S. Shorvon, and D. H. Lowenstein, "A definition and classification of status epilepticus–report of the ilae task force on classification of status epilepticus," *Epilepsia*, vol. 56, no. 10, pp. 1515–1523, 2015.

[52] L. M. DeAngelis, "Brain tumors," *New England Journal of Medicine*, vol. 344, no. 2, pp. 114–123, 2001.

[53] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[54] G. B. Moody, R. G. Mark, and A. L. Goldberger, "Physionet: a web-based resource for the study of physiologic signals," *IEEE Eng Med Biol Mag*, vol. 20, no. 3, pp. 70–75, 2001.

[55] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.

[56] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[57] H. Zhang, B. Flynn, A. T. Erdogan, and T. Arslan, "Microwave imaging for brain tumour detection using an uwb vivaldi antenna array," in *Antennas and Propagation Conference (LAPC), 2012 Loughborough*. IEEE, 2012, pp. 1–4.

[58] M. Aliroteh, G. Scott, and A. Arbabian, "Frequency-modulated magneto-acoustic detection and imaging," *Electronics letters*, vol. 50, no. 11, pp. 790–792, 2014.

[59] T. Holscher, "Prehospital use of portable ultrasound for stroke diagnosis and treatment initiation," *Air Rescue*, vol. 5, no. 2, pp. 64–67, 2012.

[60] G. Strangman, D. A. Boas, and J. P. Sutton, "Non-invasive neuroimaging using near-infrared light," *Biological psychiatry*, vol. 52, no. 7, pp. 679–693, 2002.

[61] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[62] S. Min, B. Lee, and S. Yoon, "Deep learning in bioinformatics," *CoRR*, vol. abs/1603.06430, 2016. [Online]. Available: http://arxiv.org/abs/1603.06430