An Intelligent Systems Approach to Primary Headache Diagnosis

Robert Keight[1], Ahmed J. Aljaaf[1], Dhiya Al-Jumeily[1], Abir Jaafar Hussain[1], Aynur Özge[2], and Conor Mallucci[3]

[1] Applied Computing Research Group,
The Faculty of Engineering and Technology,
Liverpool John Moores University,
Byrom Street, Liverpool, L3 3AF, UK
R.Keight@2015.ljmu.ac.uk, A.J.Kaky@2013.ljmu.ac.uk,
{D.Aljumeily, A.Hussain}@ljmu.ac.uk
[2] Mersin University School of Medicine, Mersin, Turkey
aozge@mersin.edu.tr
[3] Department of Neurosurgery,
The Royal Liverpool Childrens NHS Trust,
Alder Hey, Liverpool
cmallucci@me.com

Abstract. In this study, the problem of primary headache diagnosis is considered, referring to multiple frames of reference, including the complexity characteristics of living systems, the limitation of human information processing, the enduring nature of headache throughout history, and the potential for intelligent systems paradigms to both broaden and deepen the scope of such diagnostic solutions. In particular, the use of machine learning is recruited for this study, for which a dataset of 836 primary headache cases was considered, originating from two medical centres located in Turkey. Five primary headache classes were derived from the data obtained, namely Tension Type Headache (TTH), Chronic Tension Type Headache (CTTH), Migraine With Aura (MwA), Migraine Without Aura (MwoA), followed by Trigeminal Autonomic Cephalalgia (TAC). A total of 9 machine learning based classifiers, ranging from linear to non-linear ensembles, in addition to 1 random baseline procedure, were evaluated within a supervised learning setting, yielding highest performance outcomes of AUC 0.985, sensitivity 1, and specificity 0.966. The study concludes that modern computing platforms represent a promising setting through which to realise intelligent solutions, which in turn support the space of analytical operations needed to drive forward diagnostic capability in the primary headache domain and beyond.

1 Introduction

Over the centuries, the phenomenon of headache, which is formally known as Cephalalgia, has remained an enduring burden for societies and a significant source of enigma within the domain of healthcare systems [1, 2]. Principally,

headache is defined as pain occurring anywhere in the region of the head or neck [3], though the boundary of such conditions remains incompletely defined. Today's accepted international standard is provided by the International Classification for Headache Disorders (ICHD) criteria [4], first introduced in 1988, which describes both headache entities and diagnostic criteria. The prompt and accurate diagnosis of headache conditions represents a critical step in the healthcare pathway, since effective treatment relies on identification of the type of headache process responsible for the patient's condition [5, 6]. Such conditions are currently dichotomised into two broad categories, primary (benign), and secondary (serious), where the former occurrences are relatively common and the latter constitute only around 10% of all known cases [7]. Although epidemiological and impact studies of headache remain incomplete [8], it is indicated that the combined prevalence of headache conditions is set at around 50% of the global population, rendering an economic impact estimated to be £500 million per annum in the UK and between $5.6 and $17.2 billion in the United States alone [9]. Correspondingly, headache conditions account for one of the most common complaints within primary and neurological care settings, with a significant proportion of sufferers choosing not to seek medical attention [10] and therefore remaining unreflected in the estimates.

In this study we focus on demonstrating a potential pathway of advancement for the diagnostic process for the primary headache conditions, through the use of an intelligent systems methodology. In particular, a consideration of machine learning based diagnostics is undertaken, examining the potential role for algorithmically composed services in unlocking the information utility of increasingly numerous medical data channels. Our approach is grounded within the supervised learning framework, seeking to evaluate a range of model architectures that are trained to perform classification mappings, where response classes are representative of entities within the headache domain and therefore correspond to specific headache definitions. A consideration of 5 primary headache types is undertaken, comprising Chronic Tension Type Headache (CTTH), Cluster Headache (CH), Migraine with aura (MwA), Migraine without aura (MwoA), Tension type headache (TTH), and Trigeminal Autonomic Cephalalgia (TAC). Accordingly, the rest of this paper is therefore organised as follows. In Section 2 a review of prior works is undertaken, Section 3 considers machine learning as a diagnostic solution; the approach and methodology for this study are covered in section 4, sections 5 and 6 present results of exploratory and classification analysis, discussion and conclusions are given finally in sections 7 and 8 respectively, completing the paper.

2 Review of Prior Works

At the intersection of the domains of primary headache diagnostics and intelligent systems, a range of works are seen to exist in the literature. Intuitionistic fuzzy sets are considered in [11, 12]. The studies use an interval-valued intuitionistic fuzzy (IVIFS) weighted arithmetic average operator, with a min-max com-

position rule to determine the final diagnosis of disease. The authors concluded that the IVIFS method is viable for medical diagnosis. An Artificial Neural Network approach is considered in [13]. The study considers four types of primary headache disorder over a sample of 2,177 patients. Results reported a sensitivity and specificity of 0.93 and 0.91 for tension type headache, 0.99 and 0.94 for migraine without aura, 1.0 and 0.98 for migraine with aura, and 1.0 and 0.96 for medication overuse headache. A manually defined Decision Tree approach was considered in [14], in which the criteria published by the International Headache Society (IHS) are compared in terms of the performance of ad-hoc criteria. The research found a 10% error rate yielded by use of the IHS diagnostic criteria.

Celik et al address migraine, tension-type, cluster, and other primary headaches using an Artificial Immune System (AIS) algorithm in [15]. A sample of 850 patients was considered, reporting 94% classification accuracy, with the best result reaching 99.65%. In another study, the same authors investigate the use of the k-means algorithm in [16], restricting the scope of the study to migraines with and without aura. The work uses the data of 288 students for use with cluster based analysis, the authors conclude in closing discussion that the final system performance is not acceptable when compared with results obtained from a neurologist. Tezel and Kose apply an artificial immune approach, using a Clonal Selection Algorithm [17] to analyse a sample of 150 headache patients and 65 healthy subjects, reporting a highest classification accuracy of 92%.

3 Machine Learning as a Diagnostic Solution

Diagnostic tasks can be considered a form of information intensive problem, demanding processing pathways capable of maximising the utility of available input. The space of primary headaches, seated in biological substrates, is complex and demands a high-information factor. Conventionally, human actors have necessarily undertaken the role of information processor, acting out routines depending on a bounded numbers of variables, intuition, and commonly a procedural prototype or some form of regulation. However, given the availability of inexpensive computing power and the increasingly available sensor technologies [18], it is possible to provided concentrated analytical power from within a computational setting, shifting processing requirements into a form that is no longer directly bounded by human resource dynamics or biological cognitive limits. In fact the composition of intelligent systems, of which machine learning represents an important constituent, allows information workloads to be addressed autonomously, without the interference expected of thought-based process artefacts, namely biologically bounded attention, memory, and decision making [19, 20], in addition to intra and inter-observer variability [21], coupled with limited, localised availability. Consequently, the integration of computationally driven intelligent agency therefore promises to efficiently shift the boundaries in diagnostic accuracy in headache, allowing an arbitrary number of simultaneous patient features to be analysed without necessitating a team of human experts. Such properties offer a good response to the problem of headache, given both its

complexity and prevalence in populations. Moreover, the continued adoption of data oriented diagnostics promises to act as a feedback loop, allowing intelligent systems to sustainably self improve and adapt, in addition to furthering our understanding of the disease processes themselves.

4 Methodology

4.1 Dataset

The dataset used in this study represents 836 applicable primary headache cases, collected from two medical institutions located in Turkey, namely Cerrahpasa Medical Faculty, Istanbul, followed by Mersin University Medical Facility, Mersin. The ground truth for each case was determined by expert opinion, which is assumed in this study to represent a reasonable proxy for the true diagnosis. Diagnoses in the original dataset span a total of 8 primary headache types, which were subsequently reduced to 5 distinct classes through consolidation of the Trigeminal Autonomic Cephalalgias, which individually lacked sufficient case quantities to be useful. The final headache classes considered therefore comprise: Tension type headache (TTH), Chronic Tension Type Headache (CTTH), Migraine with aura (MwA), Migraine without aura (MwoA), followed by Trigeminal Autonomic Cephalalgia (TAC). The TAC category originally existed as 4 distinctly labelled subsets, namely Cluster Headache (CH), Paroxysmal Hemicrania (PHem), Hemicrania continua (HC), and Short-lasting unilateral neuralgiform headache with conjunctival injection and tearing (SUNCT).

4.2 Features

A total of 65 features are considered in this study, which can be summarised according to 9 broad categories. Firstly, the patient's sociodemographics were extracted, revealing 4 features, namely age, gender, smoking status, and smoking duration. Subsequently, Headache characteristics themselves are considered using 6 distinct features, comprising Headache Onset (months), Headache Frequency (days/months), Headache Characteristic (throbbing, pressing, stabbing, dull, lightening), Headache duration (hours), Headache localisation (Bilateral generalised, Bi-temporal, Calvarial, Facial Pain, Frontal, Occipital, Periocular, Secondary generalised, and Unilateral), followed by Headache intensity (Visual Analogue Scale, VAS). In conjunction with the former attributes, further features comprised of Precipitant factors (8 features), Accompanying symptoms (11 features), Psychological condition (5 features), in addition to examinations, namely Fundoscopy (1 feature) and Neurological (1 feature). The patient's medical history was also included (21 features), followed by their family medical history (8 features).

4.3 Preprocessing

Prior to the application of analytical operations, the raw dataset, originating as a result of human input, was scrutinised to ensure consistency with the expected

domain of each feature, 2 cases were dropped due to the presence of malformed information. Subsequently, all human readable character representations were subject to mapping to discrete integer enumerations, yielding a numerical encoding over all features, as is required to provide suitable input for the machine learning phase of analysis. Missing values at this point were substituted for dummy values, such that the number of cases could be maintained; appropriately values were chosen for each feature domain so as not to introduce conflicts. It was later found that this approach did not unduly impact classifier performances. The numerical representation of the problem space was then subjected to zscore normalisation, yielding a consistent scale and location of values over all features. As a result, the original dataset used for this study was applied to the analytical phase in a unified numerically encoded form.

4.4   Supervised Learning Formulation

A numerically encoded representation derived from the original domain data is used in this study as an operational substrate upon which a machine learning problem may be established. In particular, a supervised learning problem is formulated, where data may be interpreted as a series of numerically encoded examples, of the form: $x_{i,1}, ..., x_{i,p}, y_i$, where x represents a feature vector with parameters 1 through p, and $y \in 1, ..., 5$ a ground truth value. Each value of y may be considered a index pointing to one of the 5 classes of headache considered in this study. Classifier models may be formed through repeated exposure to such x, y pairs, following which the fitted model may be used to produce estimates of y, given x alone, to produce $\hat{y}$. Such a functional mapping may be given formally as $f : \bar{x} \to y \in \{1, ..., 5\}$, where the learning element of the algorithm searches by induction a hypothesis space F for a suitable $\hat{f} \in$ F, with the aim of optimal generalisation performance. In effect, once trained, the knowledge synthesised by a model may be used to classify future inputs for which outputs are not known.

4.5   Model Architectures

The series of models considered in this study are listed in Table 1, accompanied by respective architecture and hyperparameter values. In order to investigate which model types may operate effectively with the dataset under consideration, a variety of architectures have been posed, spanning multiple learning paradigms and complexity classes. The use of a diverse model space ensures that the potential information utility of the primary headache data may be profiled while highlighting the strengths and weakness of the member algorithms within this domain. Each model may be thought of as a hypothesis concerning which structural elements of the dataset may be conducive to learning a classification rule. An uninformed model, ROM, is used as a random baseline, while the LNN, LDA, and SVM models represent linear model classes, followed by KNN and TREEC which serve as weak non-linear learners, with the remaining models acting as experimental models, capable of powerful non-linear approximation.

In particular, two hybrid ensemble models are introduced, H1 and H2, such that weaknesses of individual models may be overcome while maintaining their strengths [22].

Table 1: 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.	LDA	Linear Discriminant Analysis	Linear Combination	Between class maximisation via closed form equation [23]	LB
5.	SVM	Support Vector Machine	Matrix Kernel	Quadratic Optimisation	LB
3.	KNN	K-Nearest Neighbour	5 Neighbours	Instance Induction	WNLB
3.	TREEC	Decision Tree Classifier	Decision Tree	Tree induction using information gain criterion	TC
6.	RFC	Random Forest Classifier	Decision Tree Ensemble with 200 trees	Random feature subset bagging	TC
8.	LEVNN	Levenberg Neural Network	Units: 68-55-5	Levenberg Marquardt	TC
9.	H1	Hybrid Stacked Generaliser 1	Learners: LEVNN 68-55-5, SVM, KNN 25,10,3; Combiner: RFC 200 trees	Hybrid	TC
10.	H2	Hybrid Stacked Generaliser 2	Learners: KNN 25,20,15,12,10,3; Combiner: RFC 200 trees	Hybrid	TC

Key:
RAND = Random Baseline
LB = Linear Baseline
WNLB = Weak Non-linear Baseline
TC = Test Classifier

4.6 Simulation Procedure

In order to obtain empirical results for the proposed case study, the dataset previously discussed was used for the purposes of both training and testing in

relation to the range of classifiers under consideration. The procedures for training and testing were applied to all models, producing a set of class probabilities, one for each of the 5 classes for each model. Importantly, such procedures were subjected to 50 repeated trials, such that average responses could be obtained, to attenuate the effect of outlying realisations. Additionally, the prepared dataset was partitioned according to a 70% to 30% ratio to allow systematic isolation between model fitting results and the generalisation performance obtained within the testing procedure. Feature space mapping via Principal Component Analysis (PCA) was subsequently estimated over the training set and the mapping applied to both the training and test feature vectors, such that noise components and the dimensionality were reduced. The first 55 principal components were used.

4.7 Performance Evaluation

To objectively measure the capability of the classifiers under study, a framework of performance metrics is introduced and applied to the responses obtained from the simulation procedure. A total of 7 scalar metric calculations are listed in Table 2, including sensitivity and specificity, which are applied in conjunction with graphical domain Receiver Operating Curve (ROC) analysis. ROC analysis comprises a parametric curve in two dimensional ROC space, formed from true and false positive rate measurements. Such graphical encoding provides detail of the classification performance pertaining to a given model when accounting for the complete space of decision thresholds. It is from this curve that an optimal decision threshold may be chosen. In the case of this study, the minimum Euclidean distance from the optimal location, the top North West corner, is used. Conversely, ideal random guessing performance, given a dataset of infinite size, falls along the diagonal running from the South West to the North East corners respectively. Of additional importance is that ROC analysis, in addition to being widely accepted in the medical domain [24], is agnostic to imbalances in class representation, in contrast to a simple accuracy calculation. The scalar summary metrics with the exception of Area Under Curve (AUC) are therefore dependent on a distinct decision point.

Table 2: Performance Metrics

Metric	Abbr.	Computation	Range
Area Under Curve	AUC	$0 \leqslant \mathrm{area(ROC)} \leqslant 1$	[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	$2 \cdot (PRE \cdot RC)/(PRE+RC)$	[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

5   Exploratory Analysis

Following the derivation of a feature matrix and prior to primary classification analysis, graphical domain representations of the data were constructed, encoding the problem in multiple perceptually accessible forms, each potentially revealing structural facets of interest. A single perceptually complete view of the data is not possible due to the high dimensional nature of the problem, though the combination of views may used to offer a nonetheless useful representation. In this study, both Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbourhood Embedding (tSNE) [25] techniques were used for this purpose, providing graphical encodings of the primary headache feature space as two dimensional views. Results from this procedure are listed in Figures 1 and 2. Jointly, the two graphical views reveal distinctive areas of clustering for TTH, MwoA, and CTTH, with TAC and MwA appearing more diffuse and less coherent. The tSNE domain plot appears to capture more sharply the essence of such clustering, showing both areas of separability and conflation between headache types within the two dimensional summary space. The visual domain suggests that the data does in fact carry distinct structures and that classification has a justifiable basis.
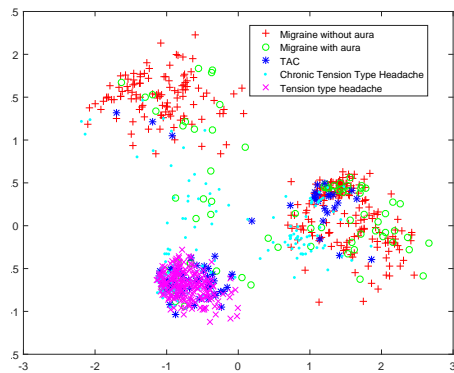

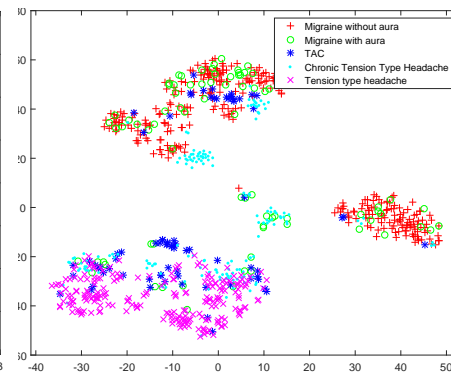Fig. 1: Principal Component Analysis Plot   Fig. 2: tSNE Plot

6   Results

The test results from the experimental procedure undertaken in this study are presented as follows. Table 3 lists the results from each classifier in terms of the 7 scalar performance metrics considered; Figure 3 shows plots resulting from ROC domain analysis, allowing visual inspection of the classifier responses over each model and class, Figure 4 provides a visual summary of the AUC values obtained from the former.
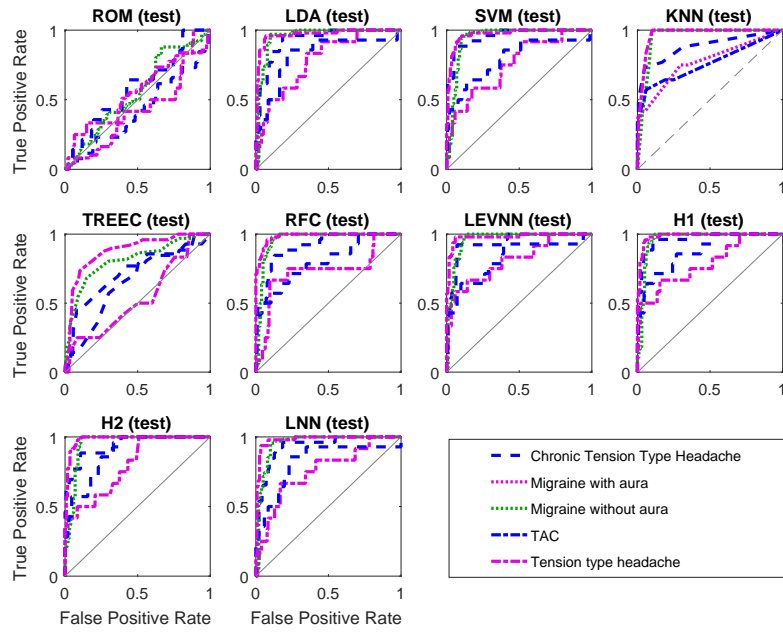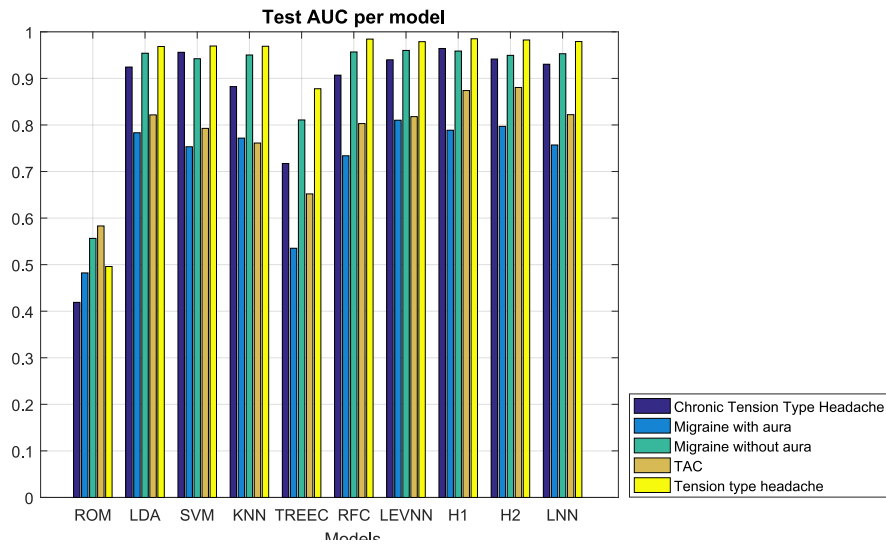
Fig. 3: ROC Plots for the Test Set

Fig. 4: AUC Plots for the Test Set

Table 3: Results (test holdout)

Model	Class	Sensitivity	Specificity	Precision	F1	J	Accuracy	AUC
ROM	TTH	0.51	0.61	0.352	0.417	0.12	0.581	0.496
	CTTH	0.615	0.362	0.151	0.242	-0.0229	0.401	0.419
	MwA	0.333	0.852	0.148	0.205	0.185	0.814	0.482
	MwoA	0.803	0.376	0.457	0.582	0.179	0.545	0.556
	TAC	0.643	0.575	0.122	0.205	0.218	0.581	0.583
LNN	TTH	0.938	0.966	0.918	0.928	0.904	0.958	0.979
	CTTH	0.923	0.865	0.558	0.696	0.788	0.874	0.93
	MwA	0.667	0.832	0.235	0.348	0.499	0.82	0.757
	MwoA	0.955	0.9	0.865	0.908	0.855	0.922	0.953
	TAC	0.857	0.771	0.255	0.393	0.628	0.778	0.822
LDA	TTH	0.959	0.932	0.855	0.904	0.891	0.94	0.969
	CTTH	0.923	0.83	0.5	0.649	0.753	0.844	0.924
	MwA	0.833	0.652	0.156	0.263	0.485	0.665	0.783
	MwoA	0.97	0.901	0.865	0.914	0.871	0.928	0.954
	TAC	0.857	0.784	0.267	0.407	0.641	0.79	0.822
SVM	TTH	0.939	0.915	0.821	0.876	0.854	0.922	0.97
	CTTH	0.885	0.936	0.719	0.793	0.821	0.928	0.956
	MwA	0.75	0.626	0.134	0.228	0.376	0.635	0.753
	MwoA	0.939	0.901	0.861	0.899	0.84	0.916	0.942
	TAC	0.643	0.863	0.3	0.409	0.506	0.844	0.793
KNN	TTH	1	0.898	0.803	0.891	0.898	0.928	0.969
	CTTH	0.731	0.95	0.731	0.731	0.681	0.916	0.882
	MwA	0.75	0.703	0.164	0.269	0.453	0.707	0.772
	MwoA	0.985	0.901	0.867	0.922	0.886	0.934	0.95
	TAC	0.643	0.81	0.237	0.346	0.453	0.796	0.761
TREEC	TTH	0.735	0.898	0.75	0.742	0.633	0.85	0.878
	CTTH	0.615	0.73	0.296	0.4	0.346	0.713	0.717
	MwA	0.417	0.6	0.0746	0.127	0.0167	0.587	0.535
	MwoA	0.682	0.851	0.75	0.714	0.533	0.784	0.811
	TAC	0.571	0.699	0.148	0.235	0.271	0.689	0.652
RFC	TTH	0.959	0.924	0.839	0.895	0.883	0.934	0.984
	CTTH	0.846	0.872	0.55	0.667	0.718	0.868	0.907
	MwA	0.75	0.781	0.209	0.327	0.531	0.778	0.734
	MwoA	0.955	0.891	0.851	0.9	0.846	0.916	0.957
	TAC	0.714	0.771	0.222	0.339	0.486	0.766	0.803
LEVNN	TTH	0.98	0.941	0.873	0.923	0.92	0.952	0.979
	CTTH	0.923	0.929	0.706	0.8	0.852	0.928	0.94
	MwA	0.667	0.858	0.267	0.381	0.525	0.844	0.81
	MwoA	0.985	0.881	0.844	0.909	0.866	0.922	0.96
	TAC	0.786	0.712	0.2	0.319	0.498	0.719	0.818
H1	TTH	0.959	0.949	0.887	0.922	0.908	0.952	0.985
	CTTH	0.962	0.894	0.625	0.758	0.855	0.904	0.964
	MwA	0.667	0.845	0.25	0.364	0.512	0.832	0.789
	MwoA	0.939	0.921	0.886	0.912	0.86	0.928	0.959
	TAC	0.857	0.732	0.226	0.358	0.589	0.743	0.874
H2	TTH	0.959	0.924	0.839	0.895	0.883	0.934	0.983
	CTTH	0.885	0.894	0.605	0.719	0.778	0.892	0.942
	MwA	0.75	0.626	0.134	0.228	0.376	0.635	0.797
	MwoA	1	0.891	0.857	0.923	0.891	0.934	0.95
	TAC	0.857	0.752	0.24	0.375	0.609	0.76	0.88

7 Discussion

From examination of the 5 primary headache classes, evaluated over the 9 models and 1 random baseline, it is apparent that TTH, CTTH, and MwoA have yielded viable results over nearly all of the informed classifiers, with exception of the decision tree learner, achieving highest AUC values of 0.985, 0.96, 0.964, respectively. Additionally, it can be observed that all three of these classes consistently yield AUC values greater than 0.9, with the exception of evaluation over the decision tree and KNN models. The classes TAC and MwA have yielded a significantly lower band of performance, with AUCs of 0.88 and 0.81 respectively, while following the same relative cross-model performance pattern as the other classes. On observation of the ROC plots, the similarity in the performance profile, with a few exceptions, can be confirmed. All models appear to exhibit a similar behaviour over the classes, with the exception of the decision tree, the KNN model, and to a lesser extent the RFC model. The highest sensitivities were achieved by the KNN model with respect to the TTH class, followed by the H2 ensemble model for the MwoA class, both of which achieved a perfect result in terms of this metric. Sensitivities for the same were 0.898 and 0.891 respectively, indicating that such model-class combinations are a viable prospect. Such a result can be filtered by the ROC geometry, showing that the KNN model, although equally sensitive when compared to the H2 model, shows a less stable performance profile overall in relation to the decision threshold. The highest overall accuracy outcome was obtained by the LNN mode over the TTH class, yielding a value of 0.958. Moreover, the other performance outcomes obtained by this model were also comparable to the more powerful non-linear classifiers, demonstrating the adequacy of a linear model complexity within the considered data scenario. Finally, it is observed that the non-linear models offered little advantage over the linear model classes, though all informed models significantly outperformed the ROM random baseline.

8 Conclusion and Future Work

To investigate primary headache diagnosis within the intelligent systems framework, an experimental procedure was undertaken in this study, considering a dataset originating from two medical institutions in Turkey, containing over 800 cases. Five primary headache classes were used as targets in the study, over which 9 machine learning based classifiers were simulated. Results showed that a highest AUC of 0.985 could be obtained, followed by a highest sensitivity of 1 and a best specificity of 0.958. The most clearly discriminated classes were found to be the TTH, MwoA, and CTTH respectively, with both linear and non-linear model classes demonstrating significant capability. Less accurate results were obtained for classes TAC and MwA, though it was found that data availability for these headache types comprised significantly fewer cases than the other three targets, leading to a hypothesis that an increase in data may equalise the result distribution, which is to be considered in future work. It was also found that

there was little indication of the presence of non-linear components within the dataset, as is demonstrated by the close similarity in performance between the non-linear and linear model types.

In aggregate, the results demonstrate that intelligent systems represent a promising approach for primary headache diagnosis, which in conjunction with wider technological ecosystems likely hold significant potential to disrupt conventional models of diagnostic delivery, positioned as a component of personalised medicine. In future work, we aim to overcome some of the key limitations inherent of the former case study. In particular, the assumption of a single headache diagnosis should be extended to allow for multiple simultaneous diagnoses, since multiple headache types are known to coexist in patients [26, 27]. Additionally, it is understood that the ground truth labels used, the diagnosis provided by expert opinion, has the potential itself to be inaccurate; an investigation of this issue is warranted. The use of deep learning algorithms is a further avenue of interest, since it was apparent that a change in feature representation via PCA caused a significant shift in results, where deep network architectures are known to internally facilitate such representational mappings [28]. Finally, the scope of the study should be expanded to included physiological signals such as EEG as an additional channel of information, since it may be possible that hidden neurophysiological markers may lie present within the dynamics of such signals, providing a more intimate vantage point of the living system state and therefore leading to increased diagnostic performance.

References

1. Magiorkinis, E., Diamantis, A., Mitsikostas, D.D., Androutsos, G.: Headaches in antiquity and during the early scientific era. Journal of neurology **256**(8) (2009) 1215–1220
2. Diamond, S., Franklin, M.A.: Headache through the ages. Professional Communications (2005)
3. La Jolla Donald, J., et al.: Wolff's headache and other head pain. Oxford University Press (2001)
4. of the International Headache Society (IHS, H.C.C., et al.: The international classification of headache disorders, (beta version). Cephalalgia **33**(9) (2013) 629–808
5. Friedman, B.W., Grosberg, B.M.: Diagnosis and management of the primary headache disorders in the emergency department setting. Emergency medicine clinics of North America **27**(1) (2009) 71–87
6. Sobri, M., Lamont, A., Alias, N., Win, M.: Red flags in patients presenting with headache: clinical indications for neuroimaging. The British journal of radiology **76**(908) (2003)
7. Martin, V.T.: The diagnostic evaluation of secondary headache disorders. Headache: The Journal of Head and Face Pain **51**(2) (2011) 346–352
8. Steiner, T.J., Stovner, L.J., Al Jumah, M., Birbeck, G.L., Gururaj, G., Jensen, R., Katsarava, Z., Queiroz, L.P., Scher, A.I., Tekle-Haimanot, R., et al.: Improving quality in population surveys of headache prevalence, burden and cost: key methodological considerations. The journal of headache and pain **14**(1) (2013) 87

9. Clarke, C., MacMillan, L., Sondhi, S., Wells, N.: Economic and social impact of migraine. Qjm **89**(1) (1996) 77–84

10. Lipton, R.B., Stewart, W.F., Simon, D.: Medical consultation for migraine: results from the american migraine study. Headache: The Journal of Head and Face Pain **38**(2) (1998) 87–96

11. Ahn, J.Y., Han, K.S., Oh, S.Y., Lee, C.D.: An application of interval-valued intuitionistic fuzzy sets for medical diagnosis of headache. International Journal of Innovative Computing, Information and Control **7**(5) (2011) 2755–2762

12. Ahn, J.Y., Park, J.H.: Headache diagnosis method using an aggregate operator. CSAM (Communications for Statistical Applications and Methods) **19**(3) (2012) 359–365

13. Mendes, K.B., Fiuza, R.M., Teresinha, M., Steiner, A.: Diagnosis of headache using artificial neural networks. J. Comput. Sci **10**(7) (2010) 172–178

14. Andrew, M.E., Penzien, D.B., Rains, J.C., Knowlton, G.E., McAnulty, R.D.: Development of a computer application for headache diagnosis: The headache diagnostic system. International journal of bio-medical computing **31**(1) (1992) 17–24

15. Çelik, U., Yurtay, N., Koç, E.R., Tepe, N., Güllüoğlu, H., Ertaş, M.: Diagnostic accuracy comparison of artificial immune algorithms for primary headaches. Computational and Mathematical Methods in Medicine **2015** (2015)

16. Celik, U., Yurtay, N., Yurtay, Y.: Headache diagnosis with k-means algorithm. Global Journal on Technology **1** (2012) 1074–1081

17. Tezel, G., Köse, U.: Headache disease diagnosis by using the clonal selection algorithm. 6th International Advanced Technologies Symposium (2011) 144–148

18. Holzinger, A., Röcker, C., Ziefle, M.: From smart health to smart hospitals. In: Smart health. Volume 8700. Springer (2015) 1–20

19. Klingberg, T.: Limitations in information processing in the human brain: neuroimaging of dual task performance and working memory tasks. Progress in brain research **126** (2000) 95–102

20. Simon, H.A.: Models of man; social and rational. John Wiley (1957)

21. Crombie, D., Cross, K., Fleming, D.: The problem of diagnostic variability in general practice. Journal of epidemiology and community health **46**(4) (1992) 447–454

22. Dietterich, T.G.: Ensemble learning. The handbook of brain theory and neural networks **2** (2002) 110–125

23. Balakrishnama, S., Ganapathiraju, A.: Linear discriminant analysis-a brief tutorial. Institute for Signal and information Processing **18** (1998)

24. Zweig, M.H., Campbell, G.: Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine. Clinical chemistry **39**(4) (1993) 561–577

25. Van der Maaten, L., Hinton, G.: Visualizing data using t-sne. Journal of Machine Learning Research **9** (2008) 2579–2605

26. Lipton, R.B., Diamond, S., Reed, M., Diamond, M.L., Stewart, W.F.: Migraine diagnosis and treatment: results from the american migraine study ii. Headache: The Journal of Head and Face Pain **41**(7) (2001) 638–645

27. Stang, P.E., Korff, M.V.: The diagnosis of headache in primary care: factors in the agreement of clinical and standardized diagnoses. Headache: The Journal of Head and Face Pain **34**(3) (1994) 138–142

28. Bengio, Y., Courville, A., Vincent, P.: Representation learning: A review and new perspectives. Pattern Analysis and Machine Intelligence, IEEE Transactions on **35**(8) (2013) 1798–1828