Migraine Classification Using Machine Learning and Deep Learning in Low-Resource Healthcare Settings

Anithamol Ashokan, Ikram Ur Rehman and Parisa Saadati

January 29, 2025

Migraine Classification Using Machine Learning and Deep Learning in Low-Resource Healthcare Settings

Anithamol Ashokan
School of Computing and Engineering
University of West London
London W5 5RF, UK
anithamolashokan.9495@gmail.com

Dr. Ikram Ur Rehman
School of Computing and Engineering
University of West London
London W5 5RF, UK
ikram.rehman@uwl.ac.uk

Dr. Parisa Saadati
School of Computing and Engineering
University of West London
London W5 5RF, UK
parisa.saadati@uwl.ac.uk

Abstract

Migraine is a neurological condition that impairs quality of life, with diagnostic challenges, especially in resource-limited settings lacking specialised tools and expertise. While AI models for migraine classification have been explored in standard healthcare, limited research focuses on low-resource environments. To address this, we evaluate the efficacy of Machine Learning and Deep Learning models (SVM, KNN, DT, RF, and TabNet) for migraine classification, with a focus on computational efficiency and interpretability. Among the models, RF emerged as the best model, achieving 95.8% accuracy, precision, recall, and F1 score, while TabNet achieved slightly lower performance 91.1%, 91.8%, 91.1%, 90.7% respectively. RF demonstrated enhanced computational efficiency, with a training time of 0.9s and memory usage of 0.14 MB, compared to TabNet's 10.8s and higher memory usage. Furthermore, SHAP analysis supported RF's interpretability, and we propose RF as a cost-effective, AI-driven diagnostic tool for migraine classification, improving access to healthcare in resource-limited regions.

Keywords: Machine Learning, Deep Learning, Migraine, Computational Efficiency, SHAP, SMOTE, Random Forest, TabNet.

1.0 Introduction

Migraine, a complex and intense neurological condition, affect approximately 1 billion individuals worldwide (Pradeep et al., 2020), ranking as the second leading cause of disability according to the World Health Organisation. Though non-life-threatening, migraines severely impact work productivity, physical health, and emotional well-being (Khan et al., 2024). Triggered by factors like sensitivity to light, irregular sleep, and skipped meals, migraines can be challenging to diagnose accurately due to their symptomatic overlap with other headache types (Migraine Trust, n.d.). According to the study conducted by Ge and Chang (2023), migraine poses significant and persistent concerns especially in non-high-income East and Southeast Asia. Traditional diagnostic tools such as Magnetic Resonance Imaging, Computed Tomography, and Positron Emission Tomography scans are often employed to diagnose migraines, but these methods are costly and typically require access to highly skilled neurologists (Khan et al., 2024). This presents a significant barrier in low-resource settings where access to imaging facilities and neurology professionals is limited (Mortel *et al.*, 2022).

Moreover, in developing nations, where personal health insurance is often lacking, especially for the poorest populations, access to costly neuroimaging testing might be difficult. For example, almost 40% of people in Cameroon live under the poverty line and 96–98% lack financial support for medical bills (Mortel *et al.*, 2022). Consequently, expenses out of pocket for CT imaging can adversely affect patients and their families financially. These circumstances emphasise the crucial need for early disease intervention to decrease the impact of chronic illnesses on patient’s lives and a country's socioeconomic conditions, especially in areas with limited access to neurologists.

The rapid advancement of Artificial Intelligence (AI), particularly in Machine Learning (ML) and Deep Learning (DL), offers promising solutions for healthcare diagnostics by extracting patterns from complex clinical data (Rathore and Mannepalli, 2021). ML/DL models have been effectively used in disease classification, including patient clustering and diagnostic support (Torrente et al., 2024), which could aid in diagnosing migraine subtypes more affordably and efficiently. However, ML/DL model adoption in clinical settings has been limited by challenges in interpretability and the high computational demands of some models, which may be impractical in low-resource environments (Habehh and Gohel, 2021). Thus, developing computationally efficient and interpretable models is crucial (Rundel et al., 2024) for enabling broader use of AI-driven diagnostic tools in underserved regions.

To the best of our knowledge, this study is the first to propose a computationally efficient and interpretable model for migraine classification specifically trained for low-resource healthcare settings, while assessing key performance metrics accuracy, precision, recall, F1 score, Area under the Receiver Operating Characteristic Curve (AUC-ROC) and Mathew's correlation coefficient (MCC). In this research we compared the performance of ML models—Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Decision Tree (DT), and Random Forest (RF) as well as the DL model TabNet. The models were evaluated on a secondary migraine dataset, with performance metrics assessed both prior to and following Hyper-Parameter Tuning (HPT) to identify the most effective approach for the classification of migraine subtype.

In addition, computational efficiency was assessed based on training and prediction time, memory usage during these process, and overall size of the model. Interpretability, a key requirement for clinical adoption was addressed by analysing the effective model output through SHapley Additive exPlanations (SHAP) summary and waterfall plots, which highlight each feature's contribution to predictions, helping healthcare professionals understand and trust the model decision. By focusing on performance, computational efficiency, and interpretability, we aim to establish an AI-driven diagnostic approach for migraine classification that is both practical and clinically acceptable, thereby reducing reliance on expensive diagnostics and facilitating early intervention in underserved areas.

This paper is structured into five sections, beginning with the introduction following the abstract. The second section presents a literature review, examining existing studies relevant to our research and identifying their limitations. The third section outlines the methodology applied in this research, while the fourth section discusses the results and analysis. The fifth and final section offers the conclusion.

2.0 Literature Review

AI poses several noteworthy applications in headache diagnosis. A Computer-based Diagnostic Engine is one of the tools designed to diagnose migraine. The engine employs a rule set that is derived from the International Classification of Headache Disorder-3 criteria for primary headaches, in addition to evaluating secondary headaches and medication overuse headaches (Cowan *et al.*, 2022).

Chiang et al. (2024) proposed a novel natural language processing technique which was recently posted on GitHub to reliably assess headache frequency from free-text clinical notes. This technology attempts to help individuals determine the proper degree of treatment depending on their headache frequency, demonstrating another unique application of AI in headache management.

Moreover, AI offers innovative methods for diagnosing and categorising migraines using a range of ML and DL algorithms. Many studies have been proposed ML/DL models that can be used for migraine diagnosis and classification.

SVM is one of the commonly utilised algorithms in the studies, especially for classification problems. For instance, the study conducted by Hsiao et al. (2023) utilised SVM with different types of kernels and discovered that a median Gaussian kernel yielded the highest level of accuracy when distinguishing chronic migraines from other types of headaches with accuracy 92.6%. SVM's adaptability, together with its efficacy in managing high-dimensional data, has made it a widely favoured option (Kazemi and Katibeh, 2018).

Another efficient algorithm identified is the RF model and its variations, such as Extremely Randomised Trees, were extensively employed, especially in ensemble techniques aimed at enhancing classification accuracy. Sasaki et al. (2023) utilised these algorithms alongside additional methods such as XGBoost, showcasing their effectiveness in accurately identifying migraine headaches with accuracy noted 94.50%. ORHANBULUCU and LATİFOĞLU (2024) employed the Rotation Forest ensemble method to increase the resilience of the model by creating a variety of decision trees, resulting in an improvement in the overall accuracy of diagnosis by 95.14%. Utilising the EEG signal data Subasi, Ahmed and Alickovic (2018) showed that RF achieved the performance of 85.18% using flash simulation.

Alternatively, DL models have had a substantial impact on studies, especially in dealing with intricate data like MRI and MEG scans. The study conducted by Khan et al. (2024) emphasised the effectiveness of Deep Neural Network (DNN) in accurately categorising seven types of migraines. The DNN attained an exceptional accuracy rate of 99.66% by utilising augmented tabular data. Moreover, the researchers Rahman Siddiquee et al. (2023) successfully employed the advanced deep learning model 3D ResNet-18 to accurately diagnose migraines and post-traumatic headaches. The model demonstrated its capability to analyse high-dimensional imaging data, achieving accuracies between 75% in differentiating migraineurs from healthy controls.

Another noteworthy DL model is TabNet employed in research conducted by (S. N. Mudassir and R. M, 2024). Due to its capacity to capture intricate patterns in tabular data while preserving interpretability, authors were able to effectively capture complex linkages within the data, leading to enhanced diagnostic outcomes noted as 98%.

Additionally, Artificial Neural Network (ANN) have been utilised in several studies to categorise different forms of headaches. Despite being less sophisticated than newer deep learning architectures, these models have demonstrated their effectiveness in research including structured or less complex data. A study showed the performance of ANN as 98% accuracy in classifying migraine subtypes (Sanchez-Sanchez, García-González and Ascar, 2020). The adaptability of ANN to different headache categorisation problems has made it a highly significant tool in this field of research (Taufique et al., 2021).

Despite being a relatively simple algorithm, KNN’s effectiveness in classifying migraine achieved 85% demonstrated its continued relevance in ML research (Romould et al., 2024).

In the research conducted by Mitrović et al. (2023), Logistic Regression and Linear Discriminant Analysis (LDA) models were used for classifying migraine categories and healthy controls, with LDA in particular achieved high accuracy of 98% when paired with feature selection methods.

Moreover, DT models were also examined, frequently as components of ensemble methods or in conjunction with other algorithms to enhance performance, to improve the accuracy and robustness of classification models which is analysed in a study utilising EEG data of patients, DT achieved performance level of 87.5% (Hsiao et al., 2023).

2.1 Limitation

While many studies report high predictive accuracy, particularly for DL models such as DNNs and 3D ResNet-18, few studies focus on the computational efficiency required for these models to run in resource-constrained environments. For instance, sophisticated DL models, although highly accurate, often demand substantial computational power and expertise, limiting their feasibility for use in low-resource clinical settings. For example, Study conducted by Rahman Siddiquee et al. (2023) highlighted the effectiveness of these models in handling data like MRI or MEG scans, but the practical applicability of these models in developing countries, where computational infrastructure is limited, is largely unexplored.

Furthermore, model interpretability is another critical factor often overlooked. Many of the algorithms, especially DL models, operate as "black box" systems, making it difficult for healthcare professionals to trust or interpret the outputs without a clear understanding of the underlying processes (Miotto *et al.,* 2018). While models like TabNet offer some degree of interpretability, the balance between performance, interpretability, and computational demands has not been comprehensively evaluated. Existing research primarily focuses on limited performance evaluation metrics rather than comparing a range of metrics also with computational efficiency and interpretability analysis (Rundel et al., 2024).

We aim to address these gaps by conducting a comprehensive comparison of ML/DL algorithms, focusing not only on predictive accuracy but also on their computational efficiency and interpretability, which are critical for practical application in resource-constrained environments. By working with tabular data instead of imaging data, we provide an alternative approach that minimises costs while ensuring that the algorithms are accessible and usable in low-resource settings. In addition, we will contribute to the literature by evaluating models using a wider range of metrics in a single study, providing valuable insights into which models are most suitable for classifying migraine subtypes, particularly in underdeveloped regions where computational and financial resources are limited.

3.0 Methodology

This section describes the methods used to identify the effective algorithm in migraine classification. Firstly, the secondary dataset was obtained followed by Data preprocessing, Exploratory Data Analysis (EDA), Feature selection, and Data augmentation technique. Secondly, the model selection process for migraine classification was conducted, followed by model training and testing both prior to and post HPT. Various evaluation metrics, including accuracy, precision, recall, F1 score, AUC-ROC, MCC, and K-fold cross-validation scores, were applied to assess performance. Additionally, computational resource requirements were analysed, and the interpretability of the effective algorithm was thoroughly examined to ensure practical applicability in a clinical setting. The detailed workflow is shown in the Figure 1 and 2.

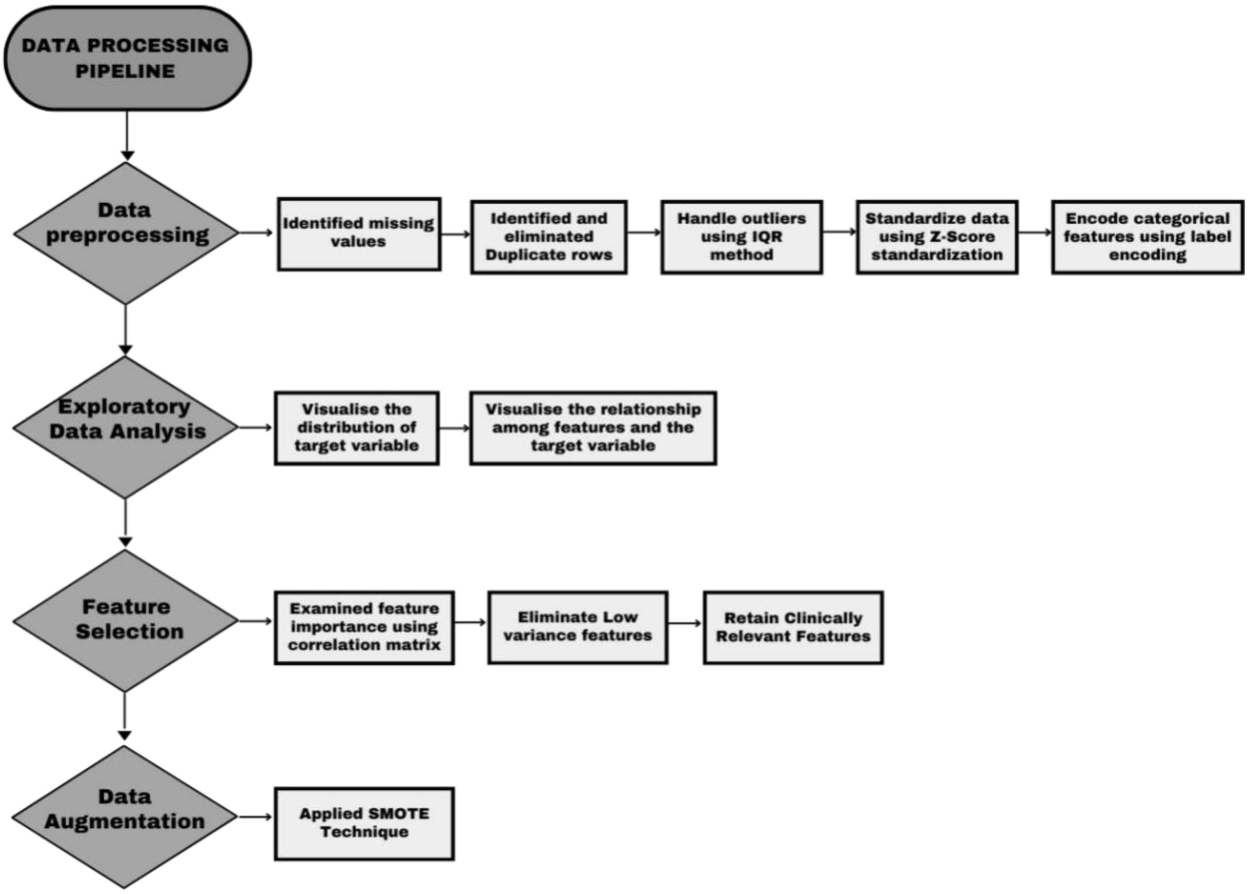


Figure 1. **Data Processing Pipeline.**

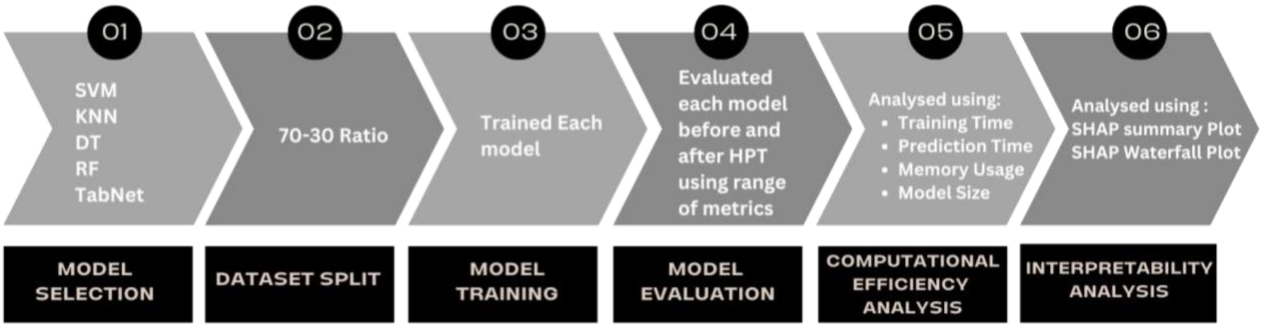


Figure 2. Model Development and Evaluation Workflow.

3.1 Dataset Collection

The dataset was acquired via Kaggle, an online community platform tailored for machine learning enthusiasts (www.kaggle.com, n.d.).

Moreover, the research conducted by (S. N. Mudassir and R. M, 2024) discovered that the dataset consists of medical records documenting patients diagnosed with various migraine-related disorders and it was carefully gathered by qualified medical experts at the Centro Materno Infantil de Soledad in the first quarter of 2013.

Furthermore, during the literature investigation, it was discovered that this migraine dataset has been extensively utilised in several studies on migraine analysis, including those conducted by (N. N. Aung and W. Srimaharaj, 2023), (Romould et al., 2024), and (Sanchez-Sanchez, García-González, and Ascar, 2020).

3.2 Dataset Description

The obtained data consists of 400 records with 24 features including both numerical and categorical variables. The attribute of the dataset is described in the Table 1. The “Type” column holds the different types of migraine diagnosed and it is the target variable in the study that need to be classified by the selected model.

SL. NO	ATTRIBUTES	DESCRIPTION	TYPE OF DATA	RANGE OF VALUE
1	Age	Age Of the Patient Reported	Continuous	15 To 77
2	Duration	Length Of Symptoms During the Most Recent Episode, Measured in Days.	Discrete	1 To 3
3	Frequency	Monthly episode frequency.	Discrete	1 To 8
4	Location	Either unilateral or bilateral pain location (none - 0, unilateral - 1, bilateral - 2)	Discrete	0 To 2
5	Character	Character depicts the throbbing or persistent sensation of pain. (none - 0, throbbing - 1, Persistant - 2)	Discrete	0 To 2
6	Intensity	The level of pain, categorised as mild, moderate, or severe (none - 0, mild - 1, moderate - 2, severe - 3)	Discrete	0 To 3
7	Nausea	Patient's sensation of nausea (not - 0, yes - 1)	Discrete	0 To 1
8	Vomit	Patient's sensation of vomiting (not - 0, yes - 1)	Discrete	0 To 1
9	Phonophobia	Sensitivity to noise (not - 0, yes - 1)	Discrete	0 To 1
10	Photophobia	Sensitivity to light (not - 0, yes - 1)	Discrete	0 To 1
11	Visual	The count of reversible visual symptoms	Discrete	0 To 4
12	Sensory	The count of reversible sensory symptoms	Discrete	0 To 2
13	Dysphasia	Impaired speech coordination (not - 0, yes - 1)	Discrete	0 To 1
14	Dysarthria	Disarticulated noises and words (not - 0, yes - 1)	Discrete	0 To 1
15	Vertigo	Symptom of dizziness (not - 0, yes - 1)	Discrete	0 To 1
16	Tinnitus	Patient's ability to hear things (not - 0, yes - 1)	Discrete	0 To 1
17	Hypoacusis	Deafness (not - 0, yes - 1)	Discrete	0 To 1
18	Diplopia	Dual vision (not - 0, yes- 1)	Discrete	0 To 1
19	Defect	Simultaneous frontal eye field and nasal field defect in both eyes (not - 0, yes - 1)	Discrete	0 To 1
20	Ataxia	Lack of muscular control (not - 0, yes - 1)	Discrete	0 To 1
21	Conscience	Compromised moral awareness (not - 0, Yes - 1)	Discrete	0 To 1
22	Paresthesia	Bilateral paraesthesia at the same time (not - 0, yes - 1)	Discrete	0 To 1
23	DPF	Family history (not - 0, yes - 1)	Discrete	0 To 1
24	Type	Type Of Migraine Diagnosed (Typical Aura With Migraine, Migraine Without Aura, Typical Aura Without Migraine, Familial Hemiplegic Migraine, Sporadic Hemiplegic Migraine, Basilar-Type Aura, Other)	Nominal	Nill

Table 1. **Dataset Description.**

3.3 Hardware and Software Specification

The analysis of the optimal algorithm for migraine classification was performed on Google Colaboratory or Colab, a cloud-based platform that offers access to robust computational resources (Carneiro *et al.,* 2018). The investigation was conducted using a MacBook Pro as the local interface to interact with the cloud environment.

3.3.1 Hardware Specification

- Local Device: MacBook Pro 13-inch (2022 model).
- Processor: Apple M2 chip that features an 8-core CPU and a 10-core GPU.
- Memory: 16 GB RAM facilitated efficient multitasking, enabling effortless transitions between several applications and browser tabs during project work.
- Operating System: macOS Sonoma, version 14.6.1 offers a reliable and protected platform for accessing cloud-based services.

3.3.2 Software Specification

- Programming Language: Python.
- Processor: Intel(R) Xeon(R) CPU @ 2.20GHz.
- Memory: 12.67 GB RAM.
- Disk: 107.72 GB total, 74.88 GB free.

3.4 Data Preprocessing

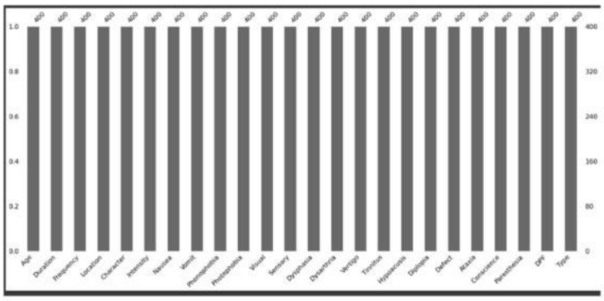
The data preprocessing involved multiple essential steps to ensure the dataset quality for analysis. Firstly, we checked for the missing values using the missingno library, and it was ensured that there were no missing values, with each column containing 400 non-missing entries. Secondly, duplicate rows were identified, revealing 6 duplicates, which were then removed, leaving 394 unique records.

Following this, outliers were identified using the Interquartile Range (IQR) method (Ur Rehman and Belhaouari, 2021), visualized through box plots. Instead of removing the outliers, they were capped to preserve the data integrity while controlling for extreme values. After addressing outliers, standardisation was applied using Z-score transformation via the StandardScaler library, ensuring each numerical feature had a mean of 0 and a standard deviation of 1 (Gao *et al.,* 2019), critical for algorithms sensitive to scale. This process confirmed that several features lacked variability, as indicated by a standard deviation of 0, suggesting minimal influence on classification outcomes.

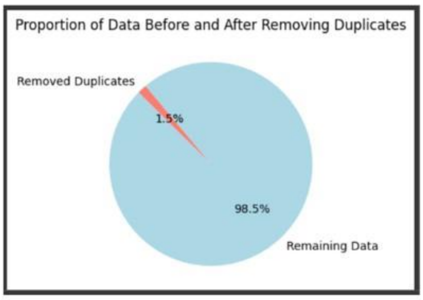
Finally, categorical data in the "Type" column was encoded using label encoding method (Qiu and Liu, 2023), transforming the 7 unique values into numerical representations suitable for model training. The encoded value corresponding to the categorical value are represented below:

- Basilliar-Type aura: 0
- Familial hemiplegic migraine: 1
- Migraine without aura: 2
- Other: 3
- Sporadic hemiplegic migraine: 4
- Typical aura with migraine: 5
- Typical aura without migraine: 6

Together, these steps created a well-prepared dataset for reliable and interpretable analysis. Figure 3 represents the data preprocessing steps done.



Bar Plot representation of analysing missing values.



Pie chart representing proportion of data before and after removing duplicates.

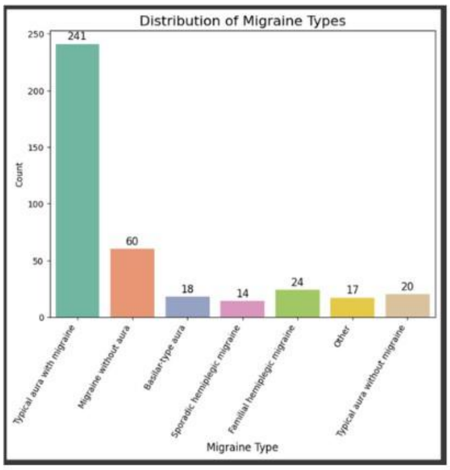
Figure 3: Data Preprocessing Steps.

3.5 Exploratory Data Analysis

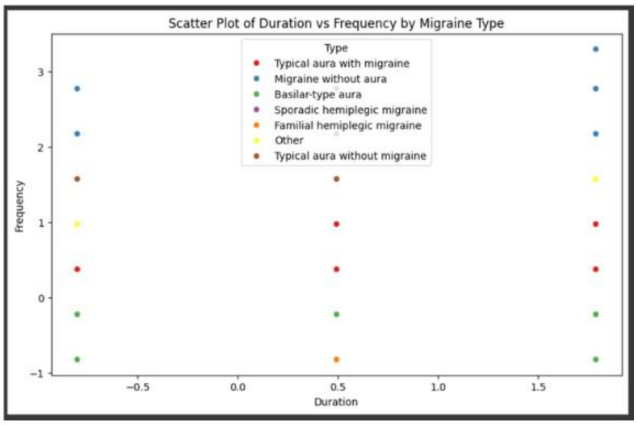
The EDA aimed to uncover key insights and patterns (DSouza, 2020) within the migraine dataset through visual analysis. Initially, the distribution of the target variable was analysed using a bar plot, which revealed an imbalance, particularly favouring the "Typical aura with migraine" class. To address this, the Synthetic Minority Oversampling Technique (SMOTE) was applied later to balance the dataset.

Secondly, a boxplot was used to examine the age distribution across migraine types, showing that certain types (e.g., "Basilar-type aura" and "Sporadic hemiplegic migraine") are associated with specific age ranges, suggesting age as a valuable classification feature. Following this, a violin plot highlighted the distribution of intensity scores among migraine types, indicating that intensity variations might enhance classification accuracy.

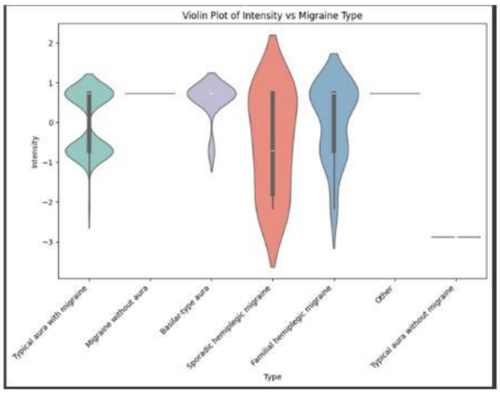
Finally, a scatter plot was used to explore the relationship between two continuous variables Duration and Frequency relative to the target variable. The plot indicated some clustering by migraine type, suggesting that these variables, when combined, could aid in differentiating between classes. These visual analyses provide foundational insights for selecting key features in the migraine classification process. Figure 4 depicts the EDA process.



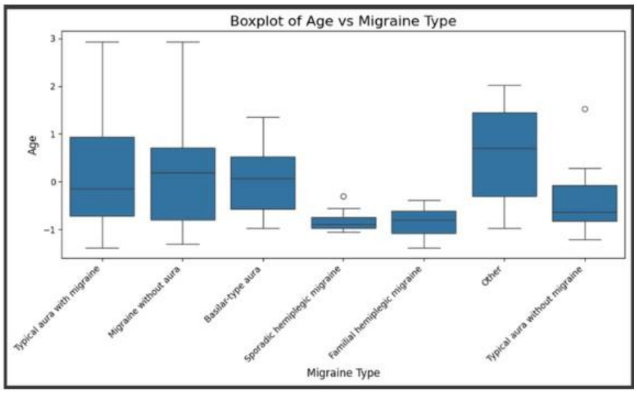
Distribution of Target Variable.



Distribution of two variables Vs Target Variable.



Distribution of Intensity Vs Target variable.



Distribution of Age Vs Target Variable.

Figure 4. **EDA processes.**

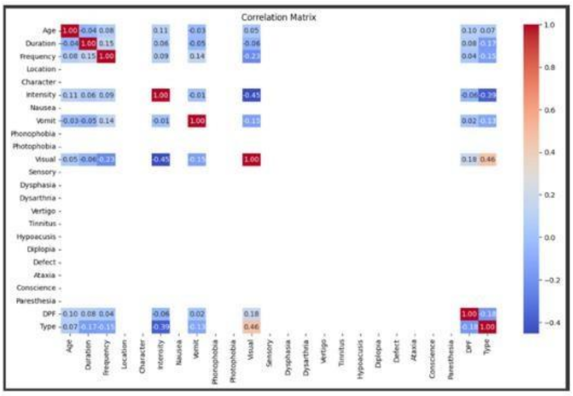
3.6 Feature Selection Process

To identify features most relevant to migraine classification, a correlation matrix analysis was conducted supplemented by insights derived from relevant domain literature. This approach enabled a nuanced selection process, balancing statistical relevance with clinical insights.

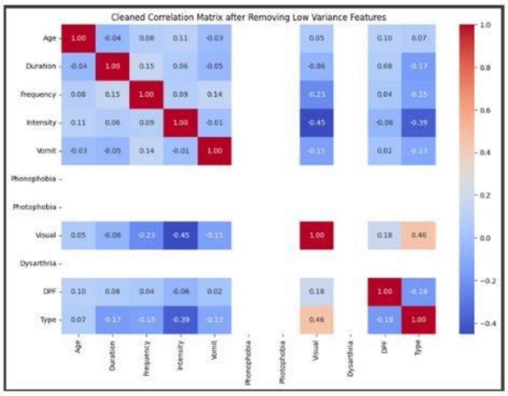
While evaluating the linear relationships among all numerical variables and their association with the target variable. Strong correlations, such as those between intensity, visual symptoms, and the target variable, were retained due to their potential predictive power. Conversely, features with minimal correlations, such as “location” and “character,” were flagged as less relevant for classification.

Features exhibiting low variance across the dataset, such as "Photophobia," "Phonophobia," and "Dysarthria," were reviewed for possible exclusion. Although these features displayed limited statistical contribution, they hold clinical significance in distinguishing migraine types (Pescador Ruschel & De Jesus, 2023; Demarquay et al., 2018). Given their medical relevance, these features were preserved to maintain a clinically comprehensive dataset, ensuring meaningful contributions to migraine classification.

This balanced approach allowed for a streamlined feature set that leverages statistically robust predictors while preserving essential clinical attributes, improving the model’s accuracy and applicability in migraine diagnosis. Figure 5 illustrates the correlation matrix prior to and after feature selection process.



Correlation matrix before feature selection.



Correlation matrix of selected feature.

Figure 5. **Correlation Matrix.**

3.7 Data Augmentation Process

To address the small sample size and imbalance in migraine types identified in preprocessing stage, SMOTE technique was applied. This data augmentation approach generates synthetic samples by interpolating between existing minority class instances and their nearest neighbours, enhancing class balance without duplication (Temraz & Keane, 2022). Following SMOTE, the dataset expanded from 394 to 1,687 records, achieving a balanced distribution across all seven migraine classes (see Figure 6). This balanced dataset better supports effective model training and classification across all migraine types.

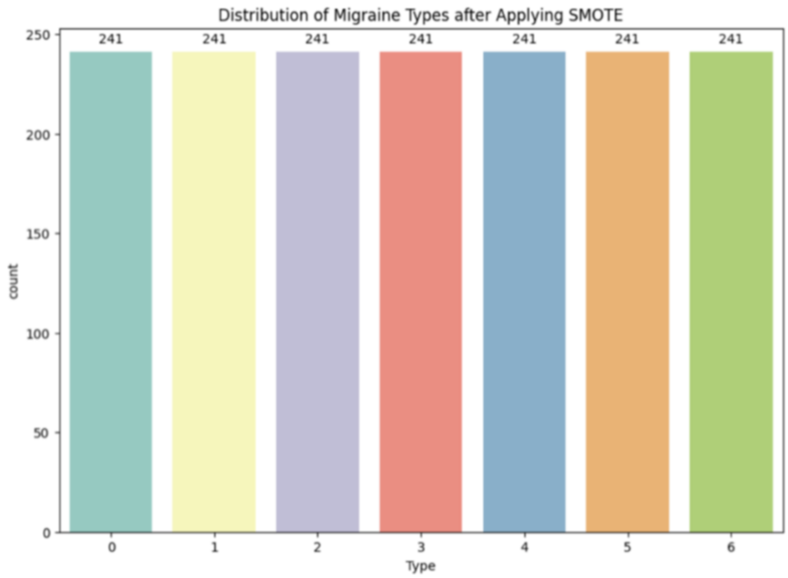


Figure 6. **Distribution of Target variable following SMOTE.**

3.8 Model Selection

From the literature analysis done, we chose some of the effective ML/DL models for migraine classification on tabular data. SVM, KNN, DT, RF, and TabNet. These models were chosen for their efficiency, interpretability, and proven performance on similar classification tasks.

3.9 Data Partitioning

The dataset was divided into training and testing subsets using 70:30 ratio, with 70% allocated for training and 30% reserved for testing to evaluate the model's predictive performance. This ratio was chosen to balance model learning and testing, providing a substantial amount of data for training while ensuring a sufficient test set to validate the model's generalisability to new, unseen data. This approach is particularly crucial in medical research, where data is limited, and models must undergo rigorous testing due to their potential impact on patient outcomes

(Gunawan Kurnia, 2024). After splitting, the training subset contained 1180 records, and the testing subset contained 507 records (see Figure 7), each with 10 features.

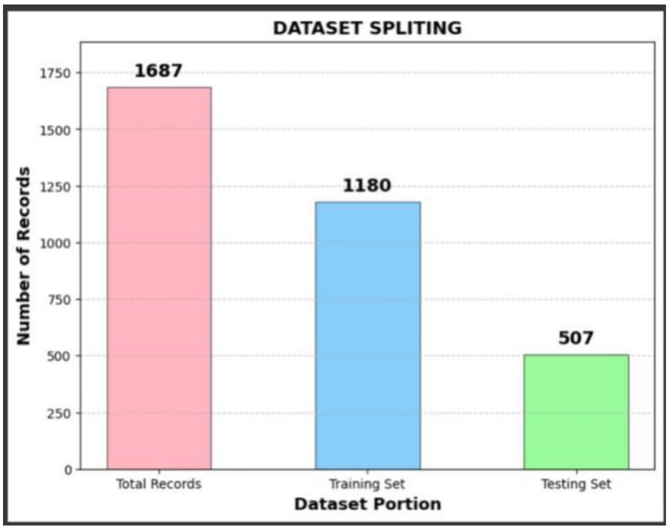


Figure 7. Number of records after data partitioning.

3.10 Model Evaluation

To classify migraine types effectively, five algorithms were evaluated prior to and after HPT. The performance was assessed using accuracy, precision, recall, F1 score, AUC-ROC, MCC and confusion matrix.

To prevent overfitting, we used stratified K-fold cross-validation with n_splits= 10, shuffle=True, and random_state=42. We employed GridSearchCV for hyperparameter optimisation, using n_jobs= -1 for parallelization and verbose=1 for monitoring the search progress. The Table 2 provides a comparison of the default and tuned parameters for each model, showing the parameter values pre and post HPT. The comparison illustrates how the tuning process influenced the model's performance.

Furthermore, the confusion matrix and AUC-ROC curve for each model, both before and after HPT, are illustrated in Figure 8 to 11. The performance comparison of the models are shown in Tables 3 and 4, respectively.

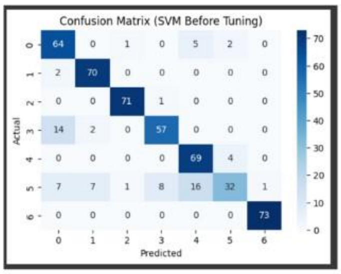
The effectiveness of the ML/DL models pre and post HPT reveals significant differences in their underlying mechanisms and their interactions with the data. Among all the classification methods considered, RF, the ML algorithm stands out as the most promising model, as it consistently outperforms others in terms of all the metrics with accuracy 95.8%. This is because of the RF's ensemble learning approach, which consists of multiple DTs to reduce overfitting and improved generalization. In contrast, DT showed underfitting due to its single-tree structure, limiting capacity to generalize well on unseen data which is reflected in its relatively lower CV score of 89.3% even with HPT.

On the other hand, SVM and KNN had significant improvement after HPT, with both algorithm's accuracy reached up to 93% and 92.8% respectively. These models exhibited substantial improvements in F1-score and MCC, suggesting improved balance and reliability in their predictions. The improvement in SVM can be attributed to the fine-tuning of its regularization parameter C where others remain default. The enhancement in KNN is likely due to the optimisation of the number of neighbours and distance metrics, which results in more precise decision boundaries. Initially, KNN's performance was limited by its sensitivity to noise and local patterns. However, HPT mitigated these issues, enabling it to capture more global structures in the data.

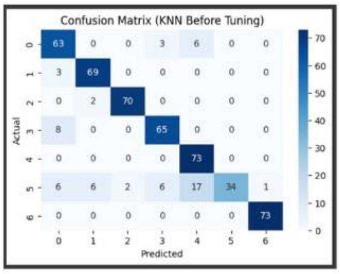
TabNet, despite being a DL model, did not exhibit as such an improvement as SVM and KNN following HPT. This could be attributed to the inbuilt design of TabNet, which already includes mechanisms such as sequential attention and feature selection, thereby optimising its ability to learn from tabular data efficiently. This model was analysed with maximum epochs 100 with early stopping parameter 10 and the batch size was set to 64 since the dataset size is comparatively small. The optimizer algorithm used was adam. Like RF model, its architecture may already be well-suited to the task, as evidenced by its stable performance both pre and post HPT. Nevertheless, it was unable to surpass RF's performance, despite achieving comparable CV score of 91.6% post HPT. This may be due to RF's bagging approach that provides greater robustness against overfitting (IBM, 2023a), a well-known challenge in deep learning models such as TabNet.

Model	Parameter	Default Value	Tuned Value
SVM	C	1.0	10
	SVM Kernel	Radial Basis Function (RBF)	Radial Basis Function (RBF)
	SVM Gamma	Scale	Scale
KNN	n_neighbors	5	3
	KNN Weights	Uniform	Distance
	KNN p	2 (Euclidean distance)	1 (Manhattan distance)
DT	Criterion	Gini	Gini
	DT max_depth	None	None
	DT min_samples_leaf	1	1
	DT min_samples_split	2	2
	DT ccp_alpha	0.0	0.001
RF	n_estimators	100	100
	RF Criterion	Gini	Entropy
	RF max_depth	None	20
	RF min_samples_split	2	5
	RF min_samples_leaf	1	1
	RF max_features	sqrt	log2
TabNet	n_d	8	24
	TabNet n_a	8	24
	TabNet n_steps	3	5
	TabNet gamma	1.3	1.0
	TabNet lambda_sparse	0.04	0.001
	TabNet optimizer_fn	torch.optim.Adam	torch.optim.Adam
	TabNet optimizer_params	0.02	0.02
	TabNet mask_type	sparsemax	sparsemax
	TabNet max_epochs	100	100
	TabNet patience	10	10
	TabNet batch_size	64	64
	TabNet virtual_batch_size	128	128

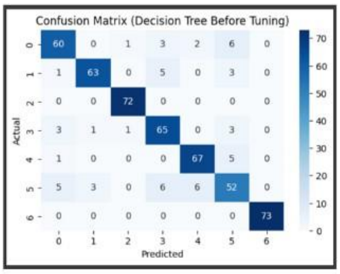
Table 2. Comparison of parameters value Before and After Optimisation.



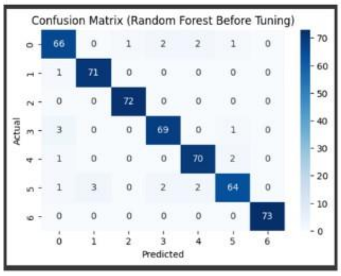
Support Vector Machine



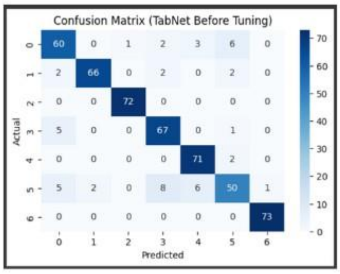
K-Nearest Neighbour



Decision Tree

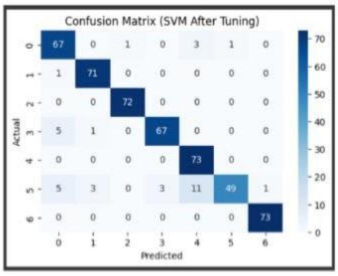


Random Forest

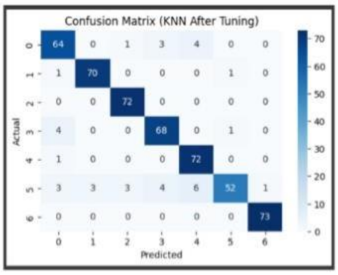


TabNet

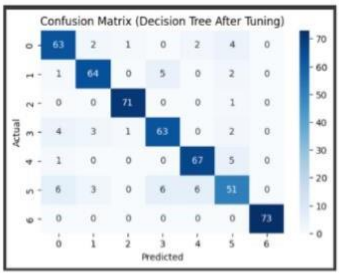
Figure 8. Confusion matrix Prior HPT.



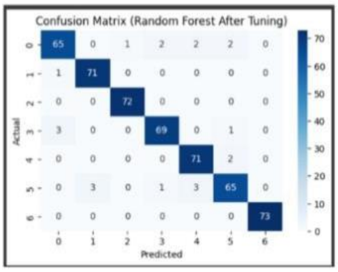
Support Vector Machine



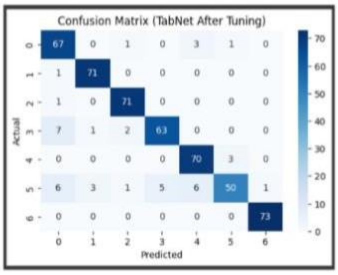
K-Nearest Neighbour



Decision Tree

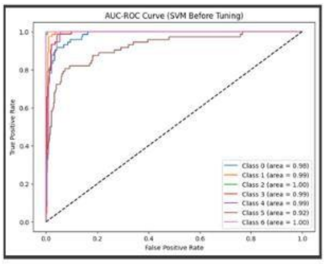


Random Forest

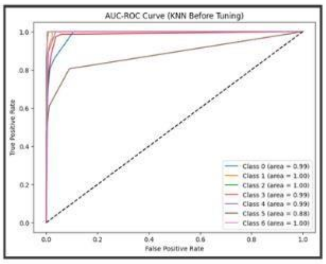


TabNet

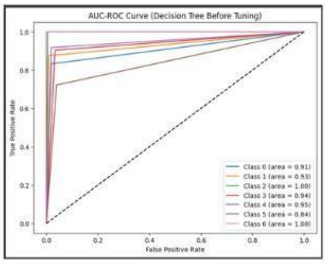
Figure 9. Confusion matrix Post HPT



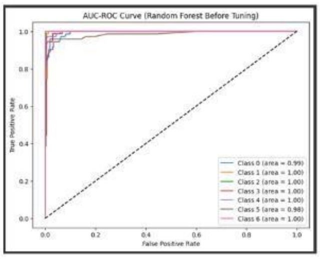
Support Vector Machine



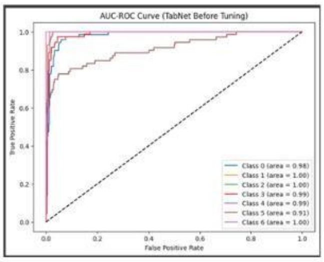
K-Nearest Neighbour



Decision Tree

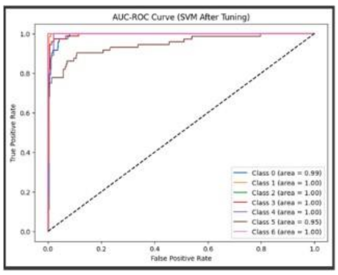


Random Forest

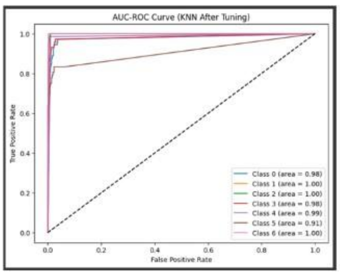


TabNet

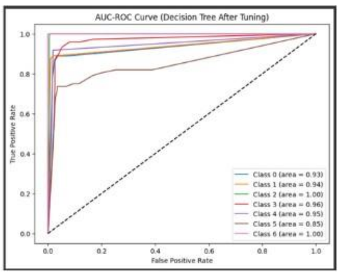
Figure 10. **AUC-ROC Curve Prior HPT.**



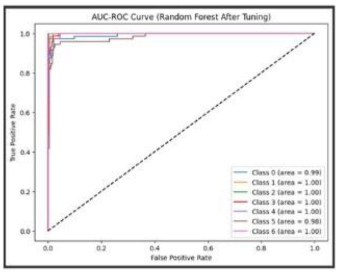
Support Vector Machine



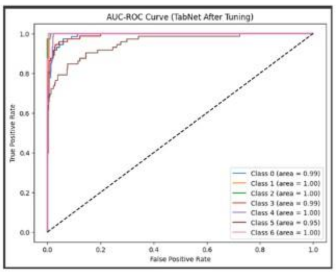
K-Nearest Neighbour



Decision Tree



Random Forest



TabNet

Figure 11. **AUC-ROC Curve Post HPT.**

MODEL	PRECISION	RECALL	F1 SCORE	AUC-ROC	MCC	ACCURACY (%)	10-FOLD CV-ACCURACY
SVM	86.4	85.9	85.0	98.2	0.83	85.9	86.6
KNN	89.7	88.1	87.3	97.7	0.86	88.1	88.1
DT	89.1	89.1	89.1	93.7	0.87	89.1	88.0
RF	95.6	95.6	95.6	99.5	0.94	95.6	92.4
TABNET	90.4	90.5	90.3	98.2	0.88	90.5	88.1

Table 3. Classification report of all algorithms before HPT.

MODEL	PRECISION	RECALL	F1 SCORE	AUC-ROC	MCC	ACCURACY (%)	10-FOLD CV-ACCURACY
SVM	93.6	93.0	92.8	98.9	0.92	93.0	91.6
KNN	93.0	92.8	92.6	97.9	0.91	92.8	91.9
DT	89.0	89.1	89.0	94.7	0.87	89.1	89.3
RF	95.8	95.8	95.8	99.6	0.95	95.8	92.6
TABNET	91.8	91.1	90.7	98.6	0.89	91.1	91.6

Table 4. Classification report of all algorithms after HPT.

3.11 Computational Efficiency Analysis

We assessed the computational efficiency of each model based on the factors:

- Training time of the model
- Prediction time of the model.
- Memory usage during training.
- Memory usage during prediction.
- Model Size.

Where the time was calculated in Seconds and memory usage in Megabytes.

Model Size is the quantity of storage capacity needed to store the trained model (Saeed, 2023). This encompasses the overall count of parameters, in addition to any supplementary data structures, such as trees in decision trees or layers in neural networks, that are essential for the model's functioning. A lower model size is beneficial in situations when there are limited resources for deploying models. The computational efficiency analysed for each model is depicted in the Figure 12 and the comparison is shown in the table 5.

The computational efficiency of the models varied significantly across different metrics. KNN and DT exhibited the fastest training times at 0.1 and 0.11 seconds, respectively, followed closely by SVM at 0.26 seconds. RF took longer to train, requiring 0.9 seconds, while TabNet had the longest training time at 10.8 seconds. In terms of prediction time, TabNet was the fastest, predicting in just 0.04 seconds. The remaining models SVM, RF, and DT had similar prediction times, ranging between 0.10 and 0.13 seconds, while KNN was the slowest, with a prediction time of 0.19 seconds.

Memory usage also differed substantially across models. During training, TabNet consumed high memory at 1.13 MB, far exceeding the other models, which used between 0.02 MB (DT) and 0.14 MB (KNN and RF). Similarly, during prediction, TabNet continued to have the highest memory consumption (0.11 MB), while SVM and DT used the least 0.01 MB. Furthermore, the model sizes for all models were very small, with SVM, KNN, DT, and RF all at 0.046 MB, and TabNet being slightly smaller at 0.035 MB.

The computational efficiency analysis shows that KNN and DT are the fastest to train, making them suitable for environments requiring frequent updates. However, TabNet excels in prediction speed, despite its longer training time, which makes it ideal for real-time applications. KNN's slower prediction time limits its use for immediate decision-making, while SVM and DT's low memory usage makes them better suited for resource-constrained settings. Even though, TabNet’s small model size and fast predictions make it best suitable, it failed in the training time taken whereas, RF balances memory, speed, and size positioning it as an adaptive option.

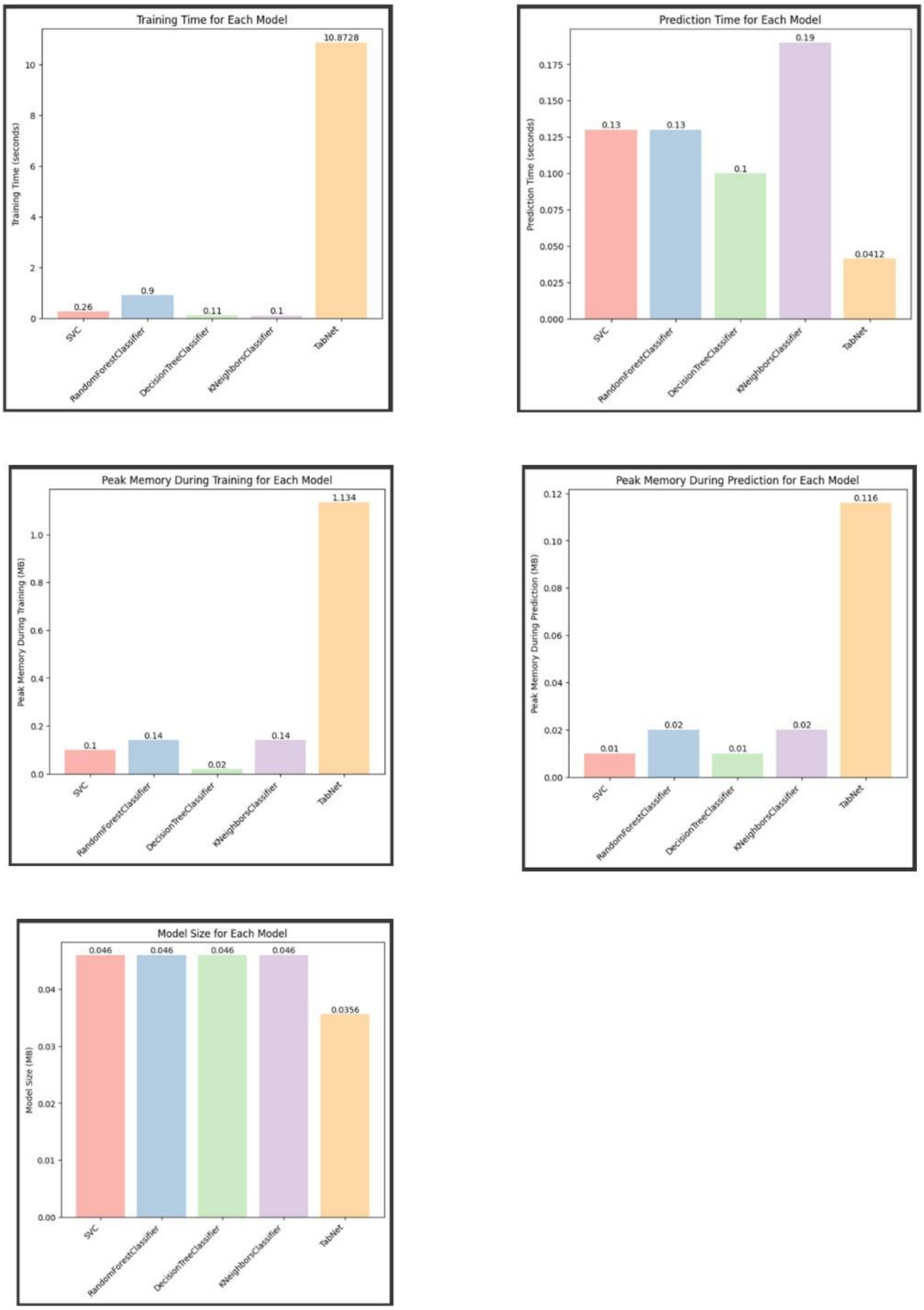


Figure 12. Computational Analysis of the model.

MODEL	Training Time (s)	Prediction Time (s)	Memory usage (Training) (MB)	Memory Usage (Prediction) (MB)	Model Size (MB)
SVM	0.26	0.13	0.1	0.01	0.046
KNN	0.1	0.19	0.14	0.02	0.046
DT	0.11	0.1	0.02	0.01	0.046
RF	0.9	0.13	0.14	0.02	0.046
TABNET	10.8	0.04	1.13	0.11	0.035

Table 5. Comparison of Computational Efficiency of all Algorithms.

3.12 Interpretability Analysis Using SHAP

To make model decisions clear and actionable for healthcare professionals interpretability analysis was performed for the effective algorithm identified. The technique used to identify the interpretability of the model was SHAP (Nguyen *et al.,* 2021). SHAP is used to better understand the model's prediction for a particular input, the contribution of each feature to this prediction is calculated.

The feature that holds the most importance is the one that exhibits the most extensive range of SHAP values acquired for that feature. The relative contribution is determined in relation to the base value. Each dot represents the contribution of a specific feature, where blue denotes lower values and red suggests higher values in a SHAP plot.

The SHAP summary plot (see Figure 13) provides an overview of the way in which various features interact and contribute to the model's predictions throughout the dataset. It was noted that features such as Age, Intensity, and Vomit have substantial interactions with other features, such as Frequency and Duration, which underscores their significant impact on the model's predictions. In contrast, features such as Phonophobia and Photophobia exhibit minimal interaction with other features as they exhibit narrower spreads of SHAP values. This suggests that these features have fewer complex relations with other features in the dataset, resulting in more direct and isolated contributions to the model's predictions.

At the same time, the SHAP waterfall plot (see Figure 14) provides a detailed explanation of how individual features contribute to the prediction for class 1 for the first instance in the test set. Starting from a baseline value of 0.143, reflective of the average probability for class 1 throughout the dataset, the figure 78 illustrates how each feature either increases or decreases the predicted chance for this specific case. Phonophobia has the most significant positive impact, increasing the predicted probability by 0.11 units, pushing the prediction toward class 1. Conversely, Duration and Photophobia have the largest negative impacts, decreasing the prediction by 0.07 and 0.06, respectively. Other features such as Intensity and Vomit contribute smaller positive and negative effects, respectively. The final prediction of 0.143 reflects the cumulative influence of these features on the model's decision, showing that while some features increase the probability, others tend to reduce the probability of this instance being classified as class 1.

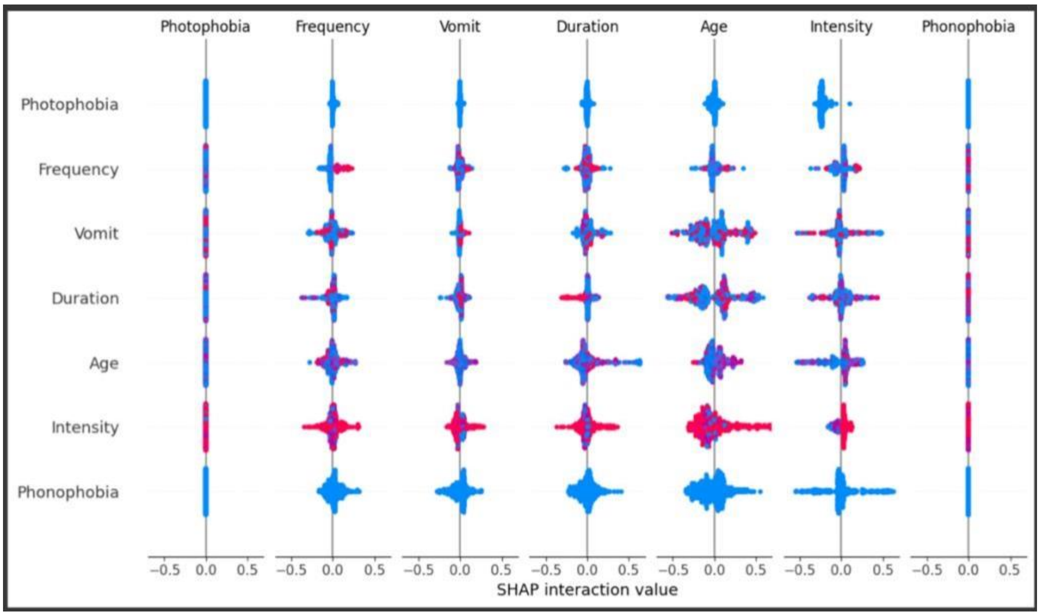


Figure 13. **SHAP summary plot for RF model Interpretability Analysis.**

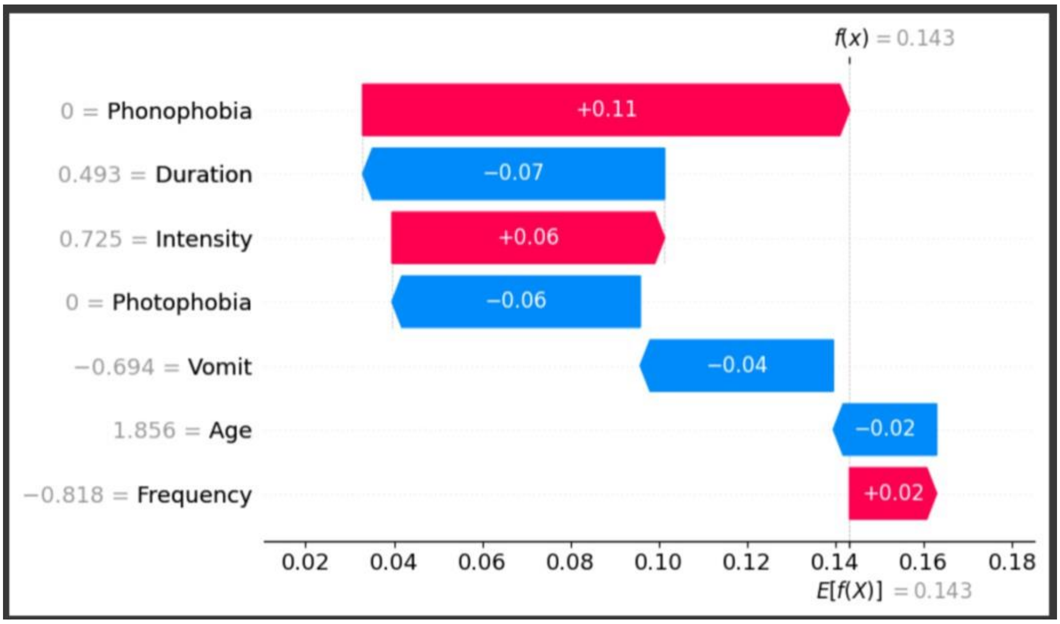


Figure 14. **Waterfall plot Analysis of RF model.**

4.0 Result and Analysis

In this study, we conducted a comprehensive evaluation of ML/DL models for classifying migraines in low-resource healthcare settings, assessing key performance metrics like accuracy, precision, recall, F1 score, AUC-ROC, and MCC, as well as computational efficiency and interpretability.

Our best-performing model, RF, demonstrated high accuracy, with an accuracy score of 95.8% and a 10-fold cross-validation accuracy of 92.6%. RF also achieved superior metrics across the board, including an F1 score of 95.8, an MCC of 0.95, and an AUC-ROC of 99.6, indicating robust performance and high reliability. The RF model's high AUC-ROC score illustrates its strong discriminatory ability between classes, further affirming its suitability for migraine classification.

Comparatively, models such as SVM, KNN, DT, and TabNet achieved notable results, yet were less optimal than RF. For instance, SVM reached an accuracy of 93.0%, with an F1 score of 92.8, an MCC of 0.92, and an AUC-ROC of 98.9. While these metrics are strong, RF outperformed SVM, particularly in AUC-ROC and cross-validation accuracy, suggesting better generalizability. KNN, with an accuracy of 92.8% and an MCC of 0.91, and TabNet, with an accuracy of 91.1% and an MCC of 0.89, both performed comparably but were less consistent and computationally efficient for resource-constrained settings. DT, on the other hand, yielded an accuracy of 89.1% and an MCC of 0.87, further reinforcing RF as the top choice in terms of balanced performance metrics.

Beyond these performance metrics, our study emphasized computational efficiency and interpretability to align with the needs of low-resource environments. RF’s computational requirements were efficiently met through cloud-based Google Colab, which provided a cost-effective and accessible environment. This setup allowed for HPT and model evaluation without demanding advanced local hardware, highlighting RF's suitability for economically challenged regions. TabNet, despite its rapid prediction capabilities, required intensive computational resources during training, presenting challenges in low-resource settings.

Interpretability was another critical focus, given the need for healthcare practitioners to understand the model’s decision-making process. SHAP interpretability analysis showed that RF relied heavily on clinically significant features, such as photophobia and intensity, validating its decision-making in a clinically relevant manner. This transparency enhances trust in the model, especially in clinical settings where interpretability supports diagnostic accuracy.

In comparison to other studies in the literature (see table 6), which report accuracy levels of RF as 78–99% (Dhiyaussalam et al., 2020; S. S. Esfahan et al., 2023; David et al., 2023) but may overlook constraints faced in resource-limited environments, our results are intentionally contextualized. By evaluating models not only on accuracy but also on computational efficiency and interpretability, our work aligns closely with the practical requirements of low-resource healthcare settings. This approach ensures that the model not only performs effectively but also generalizes well across various migraine subtypes, establishing it as a viable and reliable solution for healthcare providers in such contexts.

Sl. no	Study	Outcome by the model	Dataset type	Effective algorithm identified	Accuracy of the effective model
1	(Romould et al., 2024)	Classification of 6 types of migraine	Migraine Tabular dataset	KNN	85%
2	(S. S. Esfahan et al., 2023)	Classification of migraine vs. tension-type headaches (TTH)	Psychological and demographic data	RF	97.92%
3	(David et al., 2023)	Classification of 7 migraine types	Clinical dataset with 24 features	RF with Scatter Search	98.26%
4	(S. N. Mudassir and R. M, 2024)	Classification of various migraine subtypes	Clinical and patient-reported data	TabNet	98% overall; 99% for specific migraine subtypes
5	(Fu et al., 2024)	Classification of migraine with aura vs. migraine without aura	Structural and functional MRI data	RF	78.1%
6	(Dhiyaussalam et al., 2020)	Classification of migraine, tension-type headache (TTH), and cluster headaches	Migbase dataset with 39 features	RF	99.56%
7	(Tahhan et al., 2024)	Risk assessment and classification of migraine among university students	Survey data on lifestyle, dietary habits, and behavioral factors from university students	Linear SVM	92.7%
8	(Chen et al., 2024)	Classification of Migraine Without Aura	Resting-state fMRI data	SVM with MVPA	81.54% (First cohort); 76.47% (Second cohort)

		vs. Healthy Controls			
9	(Subasi, Ahmed and Alickovic, 2018)	Classification of migraine vs. healthy controls	EEG data collected during flash stimulation (4 Hz)	RF	85.18%
10	(Fu et al., 2022)	Classification of migraine without aura (MwA) vs. healthy controls and prediction of tVNS treatment efficacy	fMRI data from 70 MwA patients and 70 healthy controls	SVM for classification, SVR for prediction	79.3%
11	(Qawasmeh et al., 2020).	Classification of headache types: Migraine, Cluster Headache, Tension-Type Headache, and Secondary Headache	Data collected from patients	RF	99.1% (Migraine), 93% overall accuracy
12	(Hsiao et al., 2023)	Classification of chronic migraine vs. healthy controls	EEG data from	DT	87.5%
13	(Nie et al., 2023).	Classification of migraine patients vs. healthy controls	Resting-state fMRI data	SVM with combined features	96.81%
14	(Subasi et al., 2019)	Classification of migraine patients vs. healthy controls	EEG data with and without photic stimulation	RF	85.95%
15	(Kazemi and Katibeh, 2018)	Classification of pediatric migraine without aura patients vs. healthy controls	EEG data	SVM	93%
16	Proposed Study	Classification of migraine subtypes	Tabular Data	RF - ML Model TabNet - DL model	95.8% 91.1%

Table 6. **Comparative Analysis of Effective Model Performance with Related Studies.**

5.0 Conclusion

In this study, we evaluated various ML/DL algorithms for classifying migraine subtypes, focusing on their applicability in low-resource healthcare settings. Our analysis found that RF, a ML model, was the most effective, achieving 95.8% accuracy and outperforming other models in terms of additional metrics. TabNet, the DL model, achieved slightly lower performance with 91.1% accuracy. RF also demonstrated better computational efficiency compared to TabNet, making it more suitable for resource-constrained environments. Furthermore, RF's interpretability was enhanced using SHAP summary and waterfall plots, ensuring transparency for clinical use. These findings suggest that AI-driven diagnostic tools, such as RF, could reduce reliance on costly medical imaging and specialized neurologists, offering a more accessible solution for migraine classification in low-resource healthcare settings.

References

Awad, M., Khanna, R., Awad, M. and Khanna, R. (2015) 'Support vector machines for classification', *Efficient learning machines: Theories, concepts, and applications for engineers and system designers,* , pp. 39–66.

Carneiro, T., Da Nóbrega, R.V.M., Nepomuceno, T., Bian, G., De Albuquerque, V.H.C. and Reboucas Filho, P.P. (2018) 'Performance analysis of google colaboratory as a tool for accelerating deep learning applications', *Ieee Access,* 6, pp. 61677–61685.

Chen, Y., Xu, J., Wu, J., Chen, H., Kang, Y., Yang, Y., Gong, Z., Huang, Y., Wang, H. and Wang, B. (2024) 'Aberrant concordance among dynamics of spontaneous brain activity in patients with migraine without aura: A multivariate pattern analysis study', *Heliyon*, 10(9).

Chiang, C., Luo, M., Dumkrieger, G., Trivedi, S., Chen, Y., Chao, C., Schwedt, T.J., Sarker, A. and Banerjee, I. (2024) 'A large language model-based generative natural language processing framework fine-tuned on clinical notes accurately extracts headache frequency from electronic health records', *Headache: The Journal of Head and Face Pain,* 64(4), pp. 400–409.

Cowan, R.P., Rapoport, A.M., Blythe, J., Rothrock, J., Knievel, K., Peretz, A.M., Ekpo, E., Sanjanwala, B.M. and Woldeamanuel, Y.W. (2022) 'Diagnostic accuracy of an artificial intelligence online engine in migraine: a multi-center study', *Headache: The Journal of Head and Face Pain,* 62(7), pp. 870–882.

David, K.M.L., Sharmili, V.V.S., Babu, T. and Nair, R.R. (2023) *Migraine Categorization using the Scatter Search and Random Forest Classifier.* IEEE, pp. 1.

Dhiyaussalam, A. Wibowo, F. A. Nugroho, E. A. Sarwoko and I. M. A. Setiawan (2020) *Classification of Headache Disorder Using Random Forest Algorithm.* pp. 1.

DSouza, J. (2020) *Using exploratory data analysis for generating inferences on the correlation of COVID-19 cases.* IEEE, pp. 1.

Fu, C., Zhang, Y., Ye, Y., Hou, X., Wen, Z., Yan, Z., Luo, W., Feng, M. and Liu, B. (2022) 'Predicting response to tVNS in patients with migraine using functional MRI: A voxels-based machine learning analysis', *Frontiers in Neuroscience*, 16, pp. 937453.

Fu, T., Gao, Y., Huang, X., Zhang, D., Liu, L., Wang, P., Yin, X., Lin, H., Yuan, J. and Ai, S. (2024) 'Brain connectome-based imaging markers for identifiable signature of migraine with and without aura', *Quantitative Imaging in Medicine and Surgery,* 14(1), pp. 194.

Gao, Z., Ding, L., Xiong, Q., Gong, Z. and Xiong, C. (2019) 'Image compressive sensing reconstruction based on z-score standardized group sparse representation', *IEEE access,* 7, pp. 90640–90651.

Ge, R. and Chang, J. (2023) 'Disease burden of migraine and tension-type headache in non-high-income East and Southeast Asia from 1990 to 2019', *The journal of headache and pain,* 24(1), pp. 32.

Gulati, S., Guleria, K. and Goyal, N. (2022) *Classification of migraine disease using supervised machine learning.* IEEE, pp. 1.

Gunawan Kurnia (2024). *Choosing the Optimal Data Split for Machine Learning: 80/20 vs 70/30*. [online] Medium. Available at: https://medium.com/@gunkurnia/choosing-the-optimal-data-split-for-machine-learning-80-20-vs-70-30-14ed37f3c686#:~:text=The%2070%2F30%20Split&text=More%20Testing%20Data%3A%20With%2030 [Accessed 31 Aug. 2024].

Habehh, H. and Gohel, S. (2021) 'Machine learning in healthcare', *Current Genomics,* 22(4), pp. 291.

Hsiao, F., Chen, W., Wang, Y., Chen, S., Lai, K., Coppola, G. and Wang, S. (2023) 'Identification of patients with chronic migraine by using sensory-evoked oscillations from the electroencephalogram classifier', *Cephalalgia,* 43(5), pp. 03331024231176074.

Hsiao, F., Chen, W., Wang, Y., Chen, S., Lai, K., Coppola, G. and Wang, S. (2023) 'Identification of patients with chronic migraine by using sensory-evoked oscillations from the electroencephalogram classifier', *Cephalalgia,* 43(5), pp. 03331024231176074.

IBM (2023a). *What Is Random Forest? | IBM*. [online] www.ibm.com. Available at: https://www.ibm.com/topics/random-forest [Accessed 4 Sep. 2024].

Kazemi, S. and Katibeh, P. (2018) 'Comparison of parametric and non-parametric EEG feature extraction methods in detection of pediatric migraine without aura', *Journal of biomedical physics & engineering,* 8(3), pp. 305.

Khan, L., Shahreen, M., Qazi, A., Jamil Ahmed Shah, S., Hussain, S. and Chang, H. (2024) 'Migraine headache (MH) classification using machine learning methods with data augmentation', *Scientific Reports,* 14(1), pp. 5180.

Miotto, R., Wang, F., Wang, S., Jiang, X. and Dudley, J.T. (2018) 'Deep learning for healthcare: review, opportunities and challenges', *Briefings in bioinformatics,* 19(6), pp. 1236–1246.

Mitrović, K., Petrušić, I., Radojičić, A., Daković, M. and Savić, A. (2023) 'Migraine with aura detection and subtype classification using machine learning algorithms and morphometric magnetic resonance imaging data', *Frontiers in neurology,* 14, pp. 1106612.

Mortel, D., Kawatu, N., Steiner, T.J. and Saylor, D. (2022) 'Barriers to headache care in low-and middle-income countries', *Eneurologicalsci,* 29, pp. 100427.

N. N. Aung and W. Srimaharaj (2023) *Migraine Categorization based on the Integration of EMD and Naive Bayes Classification.* pp. 438.

Nguyen, H.T.T., Cao, H.Q., Nguyen, K.V.T. and Pham, N.D.K. (2021) *Evaluation of explainable artificial intelligence: Shap, lime, and cam.* pp. 1.

Nie, W., Zeng, W., Yang, J., Zhao, L. and Shi, Y. (2023) 'Classification of migraine using static functional connectivity strength and dynamic functional connectome patterns: A resting-state fmri study', *Brain Sciences,* 13(4), pp. 596.

ORHANBULUCU, F. and LATİFOĞLU, F. (2024) 'Development of a Machine Learning Based Clinical Decision Support System for Classification of Migraine Types: A Preliminary Study', .

Pradeep, R., Nemichandra, S.C., Harsha, S. and Radhika, K. (2020) 'Migraine disability, quality of life, and its predictors', *Annals of neurosciences,* 27(1), pp. 18.

Qawasmeh, A., Alhusan, N., Hanandeh, F. and Al-Atiyat, M. (2020) 'A high performance system for the diagnosis of headache via hybrid machine learning model', *International Journal of Advanced Computer Science and Applications,* 11(5).

Qiu, Q. and Liu, H. (2023) *Numerical Embedding of Categorical Features in Tabular Data: A Survey.* IEEE, pp. 446.

Rahman Siddiquee, M.M., Shah, J., Chong, C., Nikolova, S., Dumkrieger, G., Li, B., Wu, T. and Schwedt, T.J. (2023) 'Headache classification and automatic biomarker extraction from structural MRIs using deep learning', *Brain Communications,* 5(1), pp. fcac311.

Rathore, D.K. and Mannepalli, P.K. (2021) *A Review of Machine Learning Techniques and Applications for Health Care.* IEEE, pp. 4.

Romould, R.V., Singh, V., Gourisaria, M.K., Das, H. and Dash, B.B. (2024) *Deciphering Migraine Types: A Machine Learning Odyssey for Precision Prediction.* IEEE, pp. 1610.

Rundel, D., Kobialka, J., von Crailsheim, C., Feurer, M., Nagler, T. and Rügamer, D. (2024) *Interpretable machine learning for TabPFN.* Springer, pp. 465.

Rundel, D., Kobialka, J., von Crailsheim, C., Feurer, M., Nagler, T. and Rügamer, D. (2024) *Interpretable machine learning for TabPFN.* Springer, pp. 465.

S. N. Mudassir and R. M (2024) *Enhancing Migraine Diagnosis and Classification with TabNet: A Data-Driven Approach.* pp. 679.

S. S. Esfahan, A. Haratian, A. Haratian, F. Shayegh and S. Kiani (2023) *Automatic classification of migraine and tension-type headaches using machine learning methods.* pp. 220.

Saeed, F. (2023). *Model Efficiency (The Unsung Hero of Deep Learning): Achieving Top-Notch Performance with Minimal Resource Utilization*. [online] www.linkedin.com. Available at: https://www.linkedin.com/pulse/model-efficiency-unsung-hero-deep-learning-achieving-top-notch-saeed-gkrke/ [Accessed 27 Aug. 2024].

Sanchez-Sanchez, P.A., García-González, J.R. and Ascar, J.M.R. (2020) 'Automatic migraine classification using artificial neural networks', *F1000Research,* 9.

Sasaki, S., Katsuki, M., Kawahara, J., Yamagishi, C., Koh, A., Kawamura, S., Kashiwagi, K., Ikeda, T., Goto, T. and Kaneko, K. (2023) 'Developing an artificial intelligence-based pediatric and adolescent migraine diagnostic model', *Cureus,* 15(8).

Subasi, A., Ahmed, A. and Alickovic, E. (2018) 'Effect of Flash Stimulation for Migraine Detection Using Decision Tree Classifiers', *Procedia Computer Science,* 140, pp. 223–229 Available at: 10.1016/j.procs.2018.10.332.

Subasi, A., Ahmed, A., Aličković, E. and Hassan, A.R. (2019) 'Effect of photic stimulation for migraine detection using random forest and discrete wavelet transform', *Biomedical signal processing and control,* 49, pp. 231–239.

Tahhan, Z., Hatem, G., Abouelmaty, A.M., Rafei, Z. and Awada, S. (2024) 'Design and validation of an artificial intelligence-powered instrument for the assessment of migraine risk in university students in Lebanon', *Computers in Human Behavior Reports*, 15, pp. 100453.

Taufique, Z., Zhu, B., Coppola, G., Shoaran, M. and Altaf, M.A.B. (2021) 'A low power multi-class migraine detection processor based on somatosensory evoked potentials', *IEEE Transactions on Circuits and Systems II: Express Briefs,* 68(5), pp. 1720–1724.

Temraz, M. and Keane, M.T. (2022) 'Solving the class imbalance problem using a counterfactual method for data augmentation', *Machine Learning with Applications,* 9, pp. 100375.

The Migraine Trust. (n.d.). *What is migraine?* [online] Available at: https://migrainetrust.org/understand-migraine/what-is-migraine/#page-section-6 [Accessed 2 Jun. 2024].

Torrente, A., Maccora, S., Prinzi, F., Alonge, P., Pilati, L., Lupica, A., Di Stefano, V., Camarda, C., Vitabile, S. and Brighina, F. (2024) 'The clinical relevance of artificial intelligence in migraine', *Brain Sciences,* 14(1), pp. 85.

Ur Rehman, A. and Belhaouari, S.B. (2021) 'Unsupervised outlier detection in multidimensional data', *Journal of Big Data,* 8(1), pp. 80.

www.kaggle.com. (n.d.). *Migraine Dataset*. [online] Available at: https://www.kaggle.com/datasets/ranzeet013/migraine-dataset [Accessed 15 Oct. 2024].