The Seybold REPORT

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

ENHANCING MIGRAINE DIAGNOSIS: A SYMPTOM-BASED CLASSIFICATION APPROACH WITH SVM AND DECISION TREE

LAMA SADI AWAD [1], ENAS TAWFIQ NAFFAR [2] and SALEH AL-SHARAEH [3]

[1] King Abdullah II School of Information Technology, The University of Jordan Amma.
Email: lma9220481@ju.edu.jo
[2] Department of Software Engineering, Philadelphia University. Email: enaffar@philadelphia.edu.jo
[3] Professor, King Abdullah II School of Information Technology, The University of Jordan Amma.
Email: ssharaeh@ju.edu.jo

Abstract

Migraine, a widespread neurological condition with various symptoms, poses a significant challenge in achieving a precise diagnosis. The accurate identification of migraine can be a difficult task, considering the interaction of symptoms among different individuals. To address this diagnostic challenge, our research introduces an innovative Symptom-Based Classification Approach, utilizing advanced machine learning methods, particularly Support Vector Machines (SVM) and Decision Trees. These models are used to decipher the complexities within the wide range of migraine symptoms since they are skilled at identifying sophisticated patterns, combining the complexity of human physiology with the power of technology. To facilitate improved migraine diagnosis, we have carefully selected a large dataset that includes 400 patients' complete symptom histories, ranging in age from 15 to 77. This comprehensive collection documents the complex experiences that every person has during migraine attacks, providing the foundation for a strong and customized categorization model. This study compares SVM and Decision Trees to determine how each one contributes to the field of migraine diagnosis. The outcomes demonstrate the effectiveness of SVM over Deceions Trees and offer insightful information for future developments in accurate and customized migraine diagnosis.

Keywords: Migraine, Symptom-Based Classification, Support Vector Machines, Decision Tree, Machine Learning, Healthcare, Precision Diagnosis.

INTRODUCTION

Migraine is a complex neurological disorder that causes a severe headache, often accompanied by other symptoms like vomiting, nausea, and extreme sensitivity to light and sound. A migraine headache could last from hours to days, affecting the daily activities of the person who suffers from a migraine. Migraine comes with various subtypes, each requiring different treatment strategies. For some people, warning symptoms may occur before the headache, like having an aura, which is a visual disturbance such as having flashing lights or some blind spots in the eyes. Other warning symptoms might be having a tingling on one side of the face or in some parts of the body [1]. Nevertheless, not all migraine types come with those warning signs. Approximately, 75% of individuals diagnosed with migraine lack those signs before a migraine attack.

According to the World Health Organization, it is estimated that 14.7% of the world's population (almost 1 for 7 persons) suffers from migraine. Precise diagnosis of migraine in a timely manner helps control and manage the medical condition. Receiving a suitable medication can help reduce the pain, along with following proper treatment plans and

DOI: 10.5281/zenodo.10730830

strategies, the frequency and severity of migraines for many individuals can be reduced. In this study, we are driven by the following research question: *How to accurately classify migraines into subtypes based on patients' reported symptoms*? Considering the varied symptoms of migraine, identifying different migraine subtypes accurately and effectively can be a challenging task. However, it is crucial to correctly identify and diagnose migraine, to avoid medication overuse (MO). MO not only results in the body not responding to medication but may also lead to a rebound headache, making the severity of the pain worse than before [2].

Ongoing research works are being undertaken in the field of migraine prediction and classification. Most of the existing research works are concerned with the classification of migraine based on the analysis of EEG (Electroencephalogram) or MRI (Magnetic Resonance Imaging) images. *We aim to develop a reliable classification system that employs machine learning techniques, specifically Support Vector Machines (SVM) and Decision Trees, to automatically classify migraines into the relevant subtypes based on reported symptom data*. The goal is to increase the accuracy of migraine diagnosis using cost-effective techniques, which make use of simple data reported by individuals, and hence provide them with more individualized, efficient treatment plans.

We outline our objectives in this research as follows:

1. Develop a predictive model, that can classify migraine into subtypes based on patients' reported symptoms, using Support Vector Machines (SVM) and Decision Trees.
2. Evaluate and compare the performance of the SVM and Decision Trees models, to determine their effectiveness and accuracy, to enhance the precision of migraine diagnosis.
3. Offer insights into the use of SVM and Decision Trees for migraine classification, potentially paving the way for the development of novel diagnostic tools in the future. This will contribute to the expanding body of research in the field of medical data analysis and machine learning.

Background

In the following section, we demonstrate the current diagnostic landscape and explain the use of machine learning algorithms in this research: SVM and Decision Trees.

The Current Diagnostic Landscape

The current State of Diagnostics for migraines is based mostly on patient-reported symptoms, clinical examinations, and adherence to recognized criteria, such as the International Classification of Headache Disorders (ICHD). Although these techniques have been fundamental tools, their limits are evident when considering the diversity of migraineurs. Accurate and exact diagnosis is hampered by the one-size-fits-all approach's inability to recognize the unique character of symptom manifestation.

Machine Learning in Health Diagnostics

The application of machine learning in medical diagnostics represents a significant shift in the treatment of critical diseases. Methods like Support Vector Machines (SVM) and Decision Trees are remarkably effective at identifying patterns in vast and varied datasets. The use of machine learning has the potential to provide a more complex and individualized diagnostic framework for migraines, a condition whose symptoms can differ greatly from person to person.

Support Vector Machine (SVM)

An essential part of our study and a smart tool is Support Vector Machines (SVM), which are intended to help interpret complicated patterns perceived in data. SVM functions as a detective in the context of migraine diagnosis, assisting us in comprehending the distinctive characteristics of symptoms.

SVM does quite well in determining how to partition our migraine data into distinct groups. If migraine symptoms were shown on a graph, SVM would deftly create a line, or border, that optimizes the area between various symptom patterns. The optimal decision boundary allows SVM to classify and understand the diverse ways migraines can show up. [4]

SVM is skilled in managing the intricacy of migraines, which are similar to puzzles with complicated components. It examines the relationships between symptoms as well as the individual symptoms. SVM is therefore very useful for identifying minute differences and trends in migraine data.

Decision Trees

A decision tree classifies the population of a problem into segments (branches), which construct a tree with a root node, internal nodes, edges, and leaf nodes that correspond to final decisions. A decision tree can deal efficiently with large and complex datasets.

Recently, decision trees have become a popular machine-learning algorithm to be used in healthcare applications, especially for diagnosing medical conditions based on patients' reported symptoms. [7]

If the used dataset is prepared and tuned properly, decision trees can perform well in predicting medical conditions with a high accuracy. They can also identify and prioritize the most significant features contributing to a diagnosis.

LITERATURE REVIEW

In a seminal work conducted by Fu-Jung Hsiao, Wei-Ta Chen, and their colleagues (2022) [5], including *240 patients, Support* Vector Machine (SVM) methods were utilized to develop a novel classification model. Finding the ideal hyperplane that reduced hazards and generated a useful classification model was the main goal of their investigation. The SVM, a potent machine learning technique, was used in a supervised learning context to train classifiers for paired decoding of two distinct situations; in this example, patients with chronic migraine (CM)

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

were contrasted with healthy controls (HC). The level of categorization they attained was higher.

Francisco J. Pérez-Benito and colleagues (2020) *[6]* carried out a study in the field of migraine research with the goal of distinguishing patient subgroups according to attack frequency and intensity. Using state-of-the-art machine learning methods, such as robust random forest models and closest neighbors algorithms, their study aimed to find important characteristics in patient data for efficient subgroup classification. The research aligns with the growing trend of integrating machine learning into migraine studies, offering a nuanced exploration of the condition's multifaceted nature. By concentrating on distinguishing features associated with attack intensity and frequency, Pérez-Benito's study contributes valuable insights into understanding distinct characteristics underlying different migraine subgroups. As a result, machine-learning algorithms successfully pinpointed a specific subset of women experiencing migraine characterized by common intensity levels.

DTI (Diffusion Tensor Imaging) was used (2017) by Garcia-Chimeno and their team to obtain multiple measurements like fractional anisotropy (FA), which measures the completeness of the white matter. The DTI pictures and test results were used to streamline the features in the initial dataset utilizing feature selection methods (Gradient Tree Boosting, L1-based, Random Forest, and Univariate). Furthermore, to categorize the migraine group, classification algorithms (Support Vector Machine (SVM), Boosting (Adaboost), and Naive Bayes) were implemented. After testing the reliability of these algorithms, SVM yield an accuracy of 90%, followed by 93%, and 67% for boosting and naïve Bayes, respectively. As a result, the SVM classifier is an effective strategy for sample classification, profiting from solid statistical learning foundations and allowing optimization of the decision function throughout the training process. Furthermore, when performing feature selection, the initial dataset including 41 characteristics (questions and DTI pictures) was reduced to the 7 most relevant features using a combination of DTI images and questionnaires linked to emotion and cognition. This feature selection increased the classification ratio by 28% in the case of the Naive Bayes classification.

This suggests that the incorporation of these algorithms can be used to aid physicians in the classification of migraines and in giving more accurate diagnoses to patients.

Problem Statement

Migraine has varied symptoms and has many subtypes; identifying different migraine subtypes accurately and effectively can be a challenging task. Furthermore, an extensive proportion of individuals diagnosed with migraine lack warning signs, making timely and precise diagnosis essential to prevent medication overuse and accompanying complications. It is crucial to correctly identify and diagnose migraine, to help provide an individualized treatment plan that could reduce the severity and frequency of migraine.

Existing methods of migraine diagnosis primarily focus on classification using EEG or MRI, which may not be constantly accessible for all individuals. There is a need to employ machine-learning techniques to provide an automated migraine classification that is reliable and cost-effective, based on patients' reported symptoms.

METHODOLOGY

In the following section, we present our methodology to identify migraine and classify it into its subtypes based on patients' reported symptoms, using SVM and decision trees. The processes carried out in this work are depicted in Figure 1, which starts with the raw data and concludes with an assessment of the prediction model's performance. After discussing our methodology, we introduce our implemented migraine diagnostic tool.

Figure 1: The Study's Experimental Methodology

Data Collection: We obtained a dataset consisting of symptom data from 400 migraine patients, ranging in age from 15 to 77 years. This dataset includes detailed information on the symptoms experienced by each patient during migraine episodes. In the following table, we provide a detailed description and values of important variables in this study. The metadata description is displayed in Table 1.

Data Preprocessing and Exploratory Data Analysis: Performing data cleaning and validation to ensure data quality and consistency, including handling missing values, and any potential errors in the dataset to make it suitable for machine learning algorithms. To gain a deeper comprehension and understanding of the dataset, data is analyzed. The distribution of each form of migraine along with its frequency is displayed in Figure 2. Age-based patient classification is popular in medical research since some medical disorders or treatment outcomes may change with age. Following data analysis, one important finding became apparent: those under 50 years of age reported a much greater prevalence of migraines (89.0%), compared to those 50 years of age and above who reported a significantly lower prevalence rate of 10.75%. This difference emphasizes how the incidence of migraines may vary with age in the community under study. Fig.3 displays the distribution of migraine subtypes according to age.

Table 1: Metadata of Migraine Dataset

Age	**Patient's age**	**Duration**	**duration of symptoms in the last episode in days**
Frequency	Frequency of episodes per month	Location	Unilateral or bilateral pain location (None - 0, Unilateral - 1, Bilateral - 2)
Character	Character: Throbbing or constant pain (None - 0, Throbbing - 1, Constant - 2)	Intensity	Pain intensity, i.e., mild, medium, or severe (None - 0, Mild - 1, Medium - 2, Severe - 3)
Nausea	Nauseous feeling (Not - 0, Yes - 1)	Vomit	Vomiting (Not - 0, Yes - 1)
Phonophobia	Noise sensitivity (Not - 0, Yes - 1)	Photophobia	Light sensitivity (Not - 0, Yes - 1)
Visual	Number of reversible visual symptoms	Sensory	Number of reversible sensory symptoms
Dysphasia	Lack of speech coordination (Not - 0, Yes - 1)	Dysarthria	Disarticulated sounds and words (Not - 0, Yes - 1)

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

Vertigo	Dizziness (Not - 0, Yes - 1)	Tinnitus	Ringing in the ears (Not - 0, Yes - 1)
Hypoacusis	Hearing loss (Not - 0, Yes - 1)	Diplopia	Double vision (Not - 0, Yes - 1)
Visual defect	Simultaneous frontal eye field and nasal field defect in both eyes (Not - 0, Yes - 1)	Ataxia	Lack of muscle control (Not - 0, Yes - 1)
Conscience	Jeopardized conscience (Not - 0, Yes - 1)	Paresthesia	Simultaneous bilateral paresthesia (Not - 0, Yes - 1)
PDF	Family background (Not - 0, Yes - 1)	Type	Diagnosis of migraine type (Typical aura with migraine, Migraine without aura, Typical aura without migraine, Familial hemiplegic migraine, Sporadic hemiplegic migraine, Basilar-type aura, Other)

Feature Selection: Determining the most essential symptom features for the classification of migraine types, applying feature selection techniques. This stage is crucial for optimizing the performance of the model. In the context of feature selection for our migraine classification investigation, mutual information—known as information gain and information gain ratio in some analytical contexts—is a crucial statistical parameter. This measure allows us to measure the degree of dependency between symptoms and the types of migraines they are intended to describe. Mutual information essentially measures the amount of knowledge gained about a certain type of migraine by witnessing the emergence of unique symptoms. This method works very well for identifying the symptoms that have a significant impact on our classification model's ability to predict outcomes. A high mutual information score indicates a strong correlation, meaning that the symptom in question provides useful information for differentiating between different forms of migraine. Using this method enables our study to *determine which symptoms are more effective. Table 2 shows the result of the information gain and information gain ratio.*

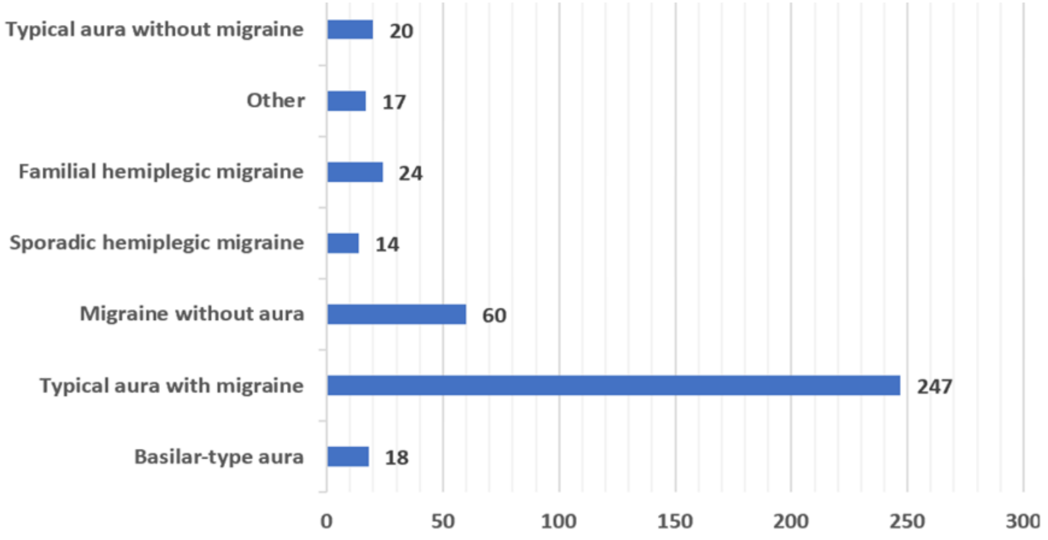


Figure 2: The distribution of migraine types with their frequencies

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

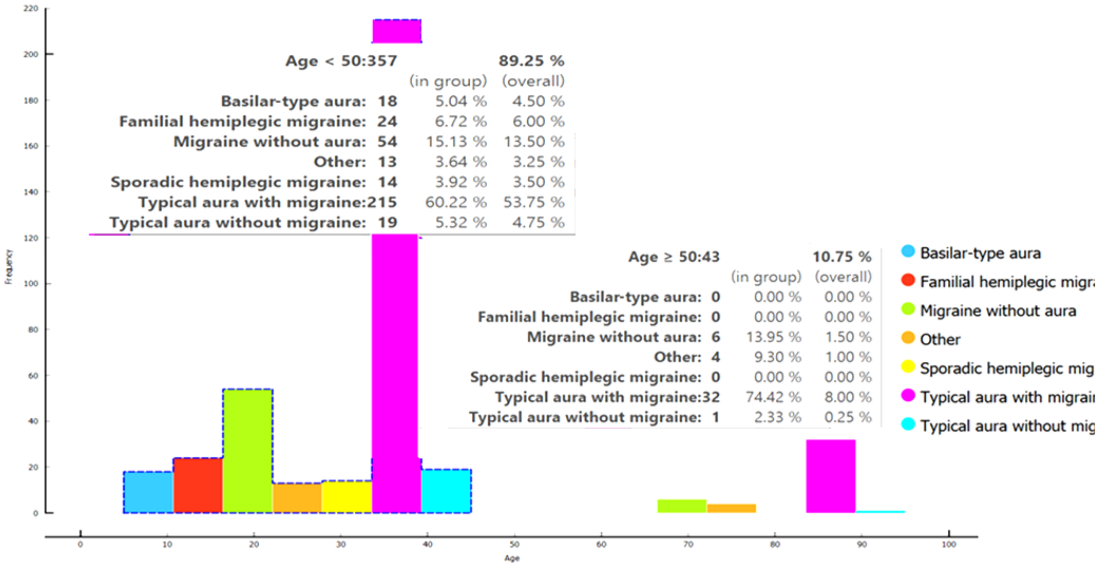


Figure 3: The Distribution of Migraine Subtypes According to Age

In our migraine classification investigation, the variables ‘Character’ and ‘Location’ demonstrate significant efficacy, as indicated by their gain ratio scores of 0.915 and 0.904, respectively, while their information gain scores are 0.426 and 0.398 respectively.

The top two variables that have the highest information gain scores are ‘Visual’ and ‘Intensity’ with information gain scores of 0.617 and 0.499 respectively.

These results highlight the critical role these characteristics play in differentiating between various types of migraines and show a high degree of knowledge gain.

In addition, within our migraine, classification study the variables ‘Phonophobia’ and ‘Photophobia’ stand out as the second most influential, with significant gain ratio scores of 0.727 and 0.700, respectively.

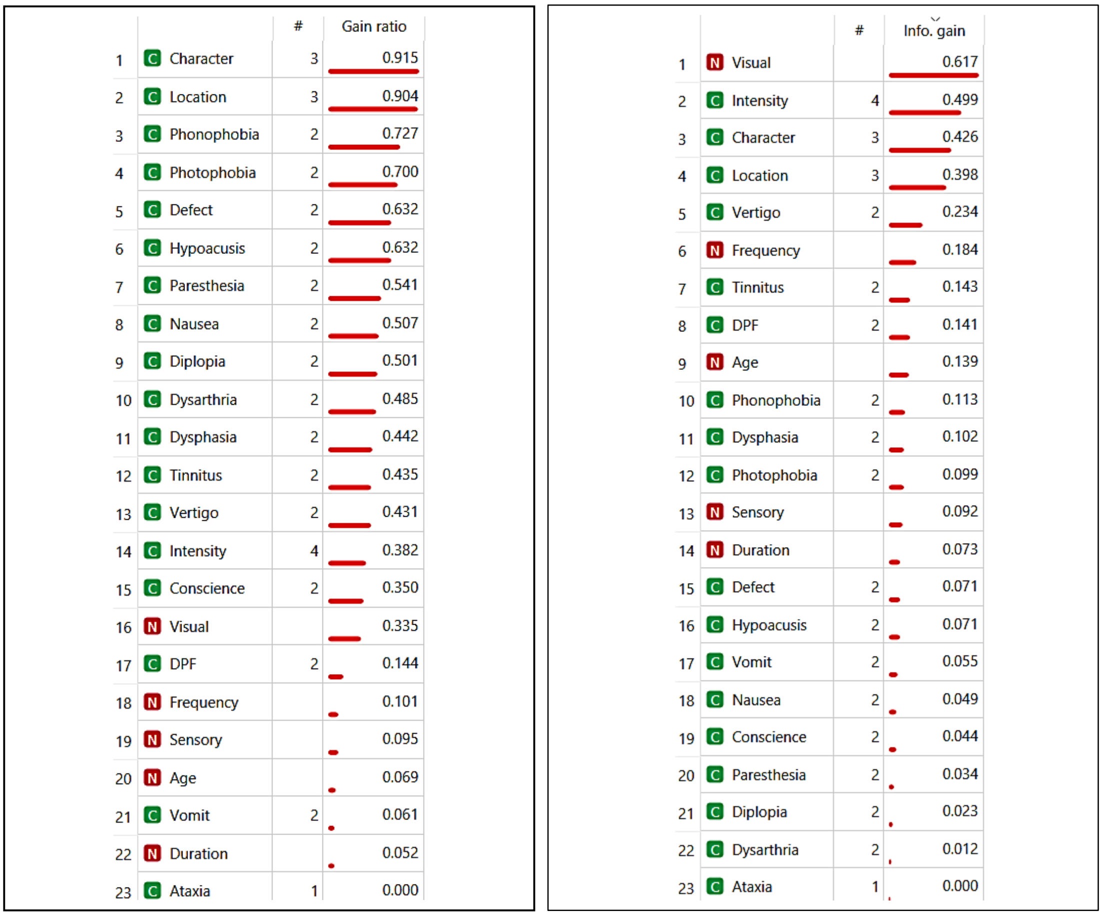
These scores highlight the important informative gain linked to these characteristics, emphasizing their importance as crucial components in distinguishing between various migraine types.

These noteworthy ratings attest to their efficacy, which further solidifies their central position in our categorization model.

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

Table 2: Information Gain and Information Gain Ratio for the 23 features



		#	Gain ratio
1	C Character	3	0.915
2	C Location	3	0.904
3	C Phonophobia	2	0.727
4	C Photophobia	2	0.700
5	C Defect	2	0.632
6	C Hypoacusis	2	0.632
7	C Paresthesia	2	0.541
8	C Nausea	2	0.507
9	C Diplopia	2	0.501
10	C Dysarthria	2	0.485
11	C Dysphasia	2	0.442
12	C Tinnitus	2	0.435
13	C Vertigo	2	0.431
14	C Intensity	4	0.382
15	C Conscience	2	0.350
16	N Visual		0.335
17	C DPF	2	0.144
18	N Frequency		0.101
19	N Sensory		0.095
20	N Age		0.069
21	C Vomit	2	0.061
22	N Duration		0.052
23	C Ataxia	1	0.000

		#	Info. gain
1	N Visual		0.617
2	C Intensity	4	0.499
3	C Character	3	0.426
4	C Location	3	0.398
5	C Vertigo	2	0.234
6	N Frequency		0.184
7	C Tinnitus	2	0.143
8	C DPF	2	0.141
9	N Age		0.139
10	C Phonophobia	2	0.113
11	C Dysphasia	2	0.102
12	C Photophobia	2	0.099
13	N Sensory		0.092
14	N Duration		0.073
15	C Defect	2	0.071
16	C Hypoacusis	2	0.071
17	C Vomit	2	0.055
18	C Nausea	2	0.049
19	C Conscience	2	0.044
20	C Paresthesia	2	0.034
21	C Diplopia	2	0.023
22	C Dysarthria	2	0.012
23	C Ataxia	1	0.000

Whereas the characteristics ‘Defect’ and ‘Hyperacusis’ show similar influence in our migraine classification effort, with both features receiving a gain ratio score of 0.632. Their comparable efficacy highlights how well they can both offer insightful information about differentiating between different forms of migraines. With these notable results, ‘Defect’ and ‘Hyperacusis’ stand out as important characteristics that greatly enhance our classification model's ability to discriminate.

Our migraine classification investigation shows that the characteristics ‘Paresthesia’, ‘Nausea’, and ‘Diplopia’ are highly useful as selection factors. With gain ratio scores of 0.541, 0.570, and 0.501, respectively, each characteristic contributes uniquely to the classification of migraine. These scores demonstrate the useful information that ‘Paresthesia’, ‘Nausea’, and ‘Diplopia’ contribute, underscoring their significance in distinguishing between various forms of migraine. Their distinct efficacy highlights their importance in molding our classification model's capacity for discrimination.

Variables that do not have high mutual information scores are not as successful as the previously mentioned characteristics. They could not have the same discriminating power. A

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

thorough evaluation of these variables' relative contributions to the overall performance of our classification model requires an understanding of their differing degrees of effectiveness. The characteristic 'Ataxia' has a gain ratio score and an information gain score of 0.00, which indicates that it has no obvious influence on the overall categorization of migraines. The results indicate that 'Ataxia' does not provide any unique information that helps distinguish between different forms of migraines. Based on our investigation, it seems that 'Ataxia' has no bearing on the classification model since its mutual information score suggests that it has no real relationship with the target variable. It is crucial to identify elements that have less influence to improve our model and concentrate on those that have a greater influence on the classification process.

To measure the strength and direction of potential linear correlations between the variables under inquiry, we used the Pearson correlation coefficient in this study to evaluate the strength of the relationship between features in our dataset. Using the Pearson correlation coefficient offers a comprehensive understanding of the architecture of the dataset, revealing subtle relationships that can indicate hidden patterns or dependencies in the context of our research into the categorization and diagnosis of migraines [8]. This statistical metric helps identify subtle associations that help us comprehend the intricate interactions between the many features in our dataset on a deeper level. *The* Pearson correlation coefficient is seen in Table 3.

The strength of the correlation between 'Character' and 'Location' is +0.934, indicating a significant positive link. Likewise, a strong positive connection is seen at +0.659 in the correlation between 'Intensity' and 'Location'. These results highlight the qualities' interdependence and point to possible co-occurrences or common patterns in the dataset. In addition, there is a negative correlation of -0.466 between intensity and visual, a negative correlation of -0.342 between character and visual, and a negative correlation of -0.327 between location and visual. These results reveal that there is an inverse relationship between the qualities listed, meaning that modifications to one property may have the opposite effect on modifications to the visual attribute in the dataset.

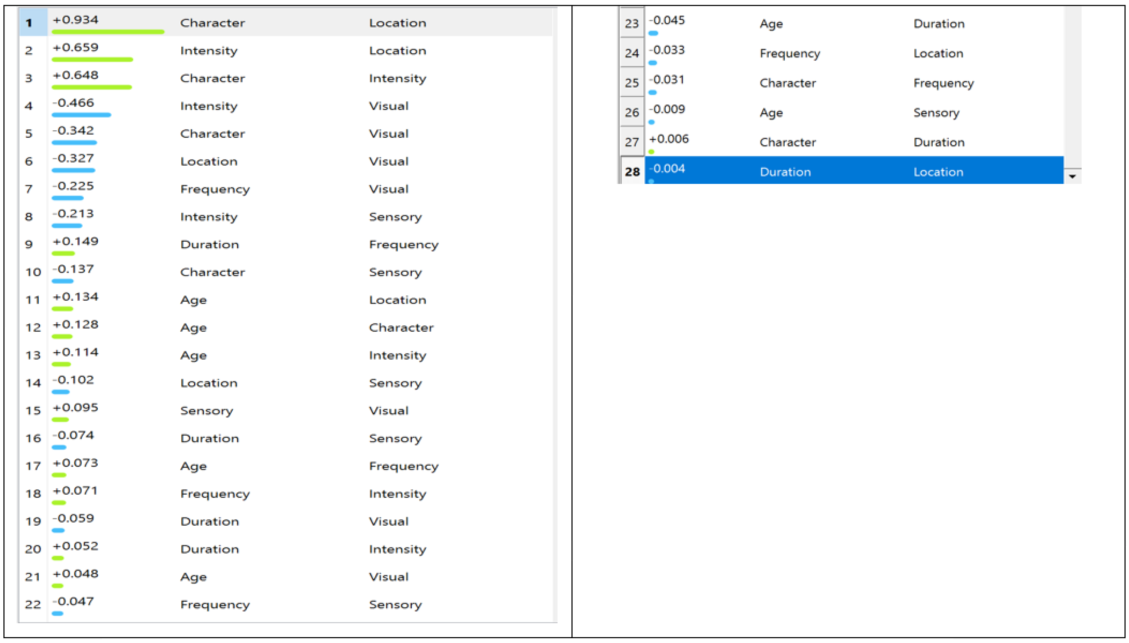
Data Splitting: Our migraine dataset must be carefully split into training and testing groups for the Support Vector Machine (SVM) model and Decision Tree model to function properly. We carefully divided the data into two sets: the training set and the testing set, since we recognized how crucial it was to evaluate the model accurately. The maintenance of a realistic distribution of migraine subtypes in both sets was carefully considered during the partitioning process, which is essential for training and evaluating the model's capacity for generalization. The training set received 70% of the data, with the remaining 30% being reserved for testing, according to the 70:30 splitting ratio.

Symptom -based Classification: When choosing a model for our migraine classification study, one of the most important steps was to figure out which machine learning algorithms would work best to capture the complex patterns that come with different migraine subtypes. Decision trees and Support Vector Machines (SVM), two well-known algorithms, emerged as the top contenders for several convincing reasons.

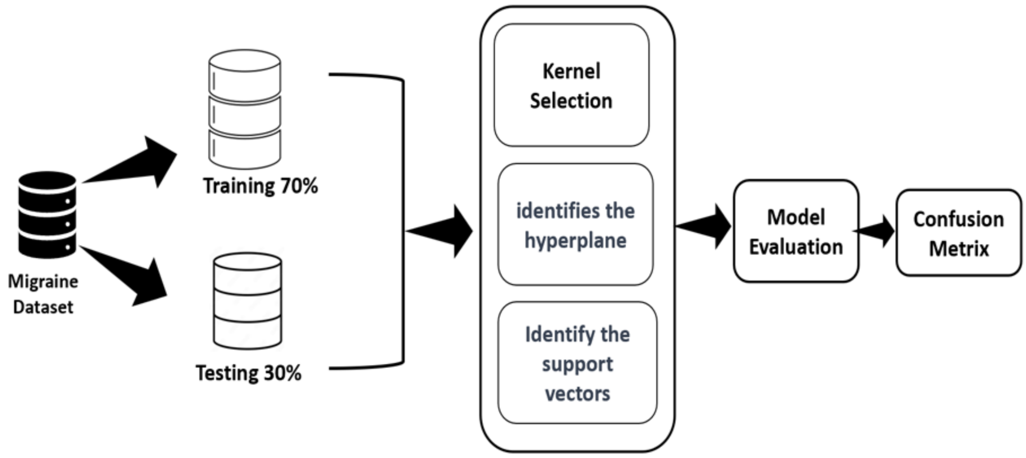
DOI: 10.5281/zenodo.10730830

The selection of SVM was based on its capacity to manage intricate, multidimensional data and its efficacy in distinguishing linear and non-linear decision boundaries. The ability of SVM to identify the best hyperplanes for class separation provided a reliable solution for our classification challenge, especially considering the multidimensional nature of migraine-related variables. Figure 4 shows the phases of Support Vector Machines for Classifying Migraines.

Table 3: Pearson Correlation Coefficient



1	+0.934	Character	Location
2	+0.659	Intensity	Location
3	+0.648	Character	Intensity
4	-0.466	Intensity	Visual
5	-0.342	Character	Visual
6	-0.327	Location	Visual
7	-0.225	Frequency	Visual
8	-0.213	Intensity	Sensory
9	+0.149	Duration	Frequency
10	-0.137	Character	Sensory
11	+0.134	Age	Location
12	+0.128	Age	Character
13	+0.114	Age	Intensity
14	-0.102	Location	Sensory
15	+0.095	Sensory	Visual
16	-0.074	Duration	Sensory
17	+0.073	Age	Frequency
18	+0.071	Frequency	Intensity
19	-0.059	Duration	Visual
20	+0.052	Duration	Intensity
21	+0.048	Age	Visual
22	-0.047	Frequency	Sensory
23	-0.045	Age	Duration
24	-0.033	Frequency	Location
25	-0.031	Character	Frequency
26	-0.009	Age	Sensory
27	+0.006	Character	Duration
28	-0.004	Duration	Location



Below, we briefly explain the main three components of the SVM model:

Figure 4: Support Vector Machines (SVM) for Migraine Classification

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

Kernel Selection

We decided to use a Linear Kernel in conjunction with the Support Vector Machine (SVM) technique. The use of a Linear Kernel is consistent with the natural linearity shown in the correlations between patent symptoms. The interpretability of a linear decision boundary, which makes it possible to clearly explain the elements influencing the categorization of migraine subtypes, is what drove this choice. As a practical first step, the linear SVM provides simplicity and computing efficiency consistent with the linear patterns that predominate in the data. We can calculate the Linear Kernel using the equation: **K(xi, xj) =xi. X j**

Identify the Hyperplane

Using the training dataset to train the SVM model. The algorithm finds, during training, the hyperplane that maximizes the margin between support vectors and effectively divides classes.

Identify the Support Vectors

Determining which data, points are closest to the decision boundary, or the support vectors. These points are essential for figuring out the ideal distance between classes and for creating the hyperplane.

The selection of Decision Trees was based on their interpretability and simplicity. Decision trees offer a transparent depiction of the decision-making process with elements that contribute to the classification process, enabling insights into the factors driving migraine subtype categorization. Figure 5 demonstrates the phases of Decision trees for Classifying Migraines.

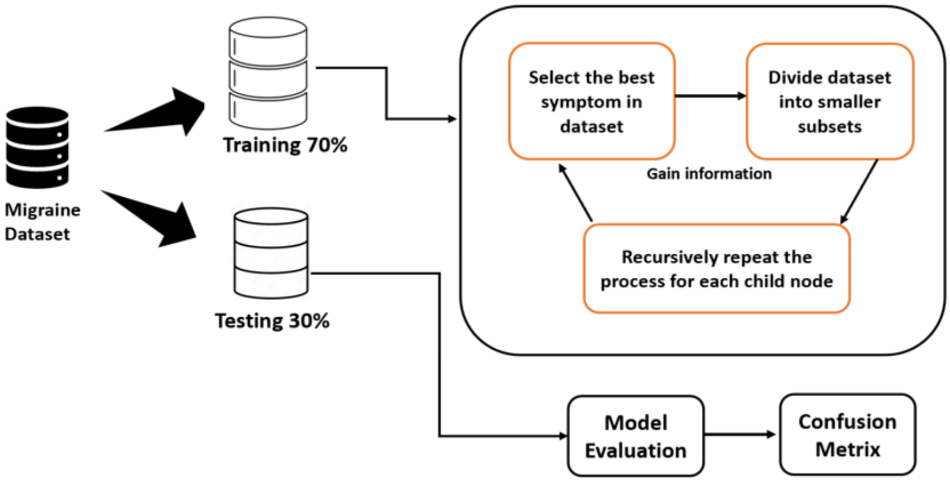


Figure 5: Decision Tree for Migraine Classification

The main components of the Decision Tree model are:

Select the Best Symptom in the Dataset

Selecting the most informative features that help make decisions at each node. Decision Tree algorithm evaluates and selects the best feature that separates data into a set of different classes

The Seybold Report

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

Divide the Dataset into Smaller Subsets

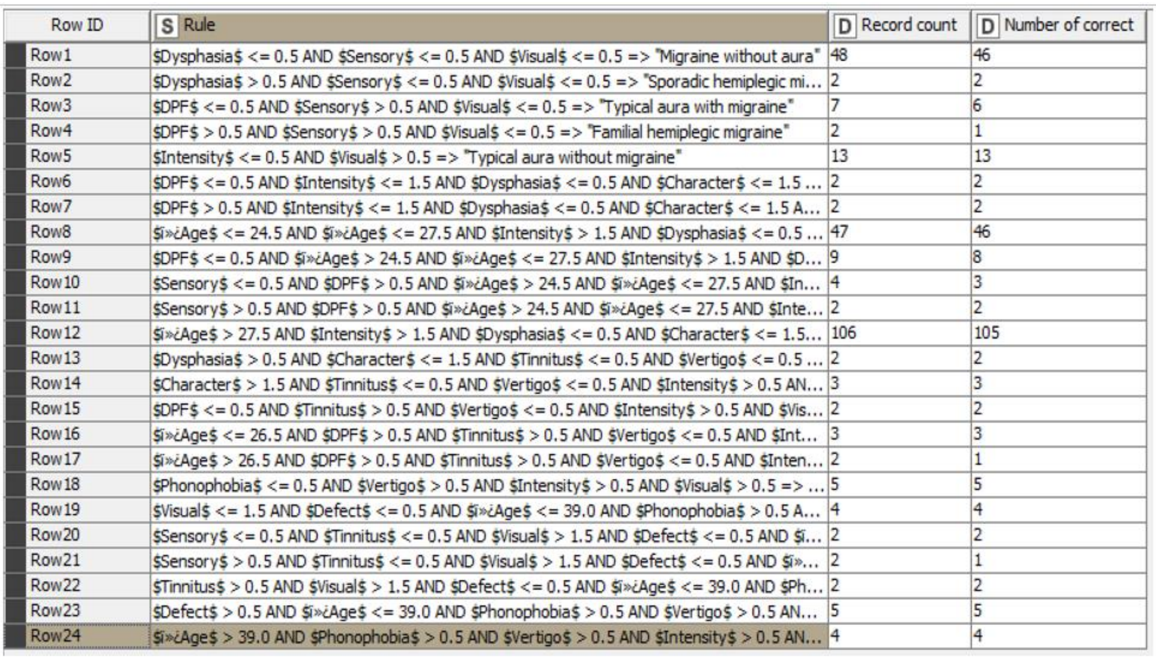
Based on the best-selected feature (symptom), the data is divided into smaller groups at each node. Hence, a subset of data is created at each branch of the node.

Recursively Repeat the Process for Each Child Node

The two steps of selecting the best symptom in the dataset, as well as dividing the dataset into smaller subsets, are repeated at each child node until a stopping condition is met.

It is worth mentioning that after a decision tree is constructed, each path of the tree is traversed until reaching a leaf node, and a rule table, as can be seen in Table 4, is constructed. Each path of the tree represents a condition that leads to a specific type of migraine.

Table 4: Rules table for Decision Tree



Row ID	S Rule	D Record count	D Number of correct
Row1	$Dysphasia$ <= 0.5 AND $Sensory$ <= 0.5 AND $Visual$ <= 0.5 => "Migraine without aura"	48	46
Row2	$Dysphasia$ > 0.5 AND $Sensory$ <= 0.5 AND $Visual$ <= 0.5 => "Sporadic hemiplegic mi...	2	2
Row3	DPF <= 0.5 AND $Sensory$ > 0.5 AND $Visual$ <= 0.5 => "Typical aura with migraine"	7	6
Row4	DPF > 0.5 AND $Sensory$ > 0.5 AND $Visual$ <= 0.5 => "Familial hemiplegic migraine"	2	1
Row5	$Intensity$ <= 0.5 AND $Visual$ > 0.5 => "Typical aura without migraine"	13	13
Row6	DPF <= 0.5 AND $Intensity$ <= 1.5 AND $Dysphasia$ <= 0.5 AND $Character$ <= 1.5 ...	2	2
Row7	DPF > 0.5 AND $Intensity$ <= 1.5 AND $Dysphasia$ <= 0.5 AND $Character$ <= 1.5 A...	2	2
Row8	$ï»¿Age$ <= 24.5 AND $ï»¿Age$ <= 27.5 AND $Intensity$ > 1.5 AND $Dysphasia$ <= 0.5 ...	47	46
Row9	DPF <= 0.5 AND $ï»¿Age$ > 24.5 AND $ï»¿Age$ <= 27.5 AND $Intensity$ > 1.5 AND $D...	9	8
Row10	$Sensory$ <= 0.5 AND DPF > 0.5 AND $ï»¿Age$ > 24.5 AND $ï»¿Age$ <= 27.5 AND $In...	4	3
Row11	$Sensory$ > 0.5 AND DPF > 0.5 AND $ï»¿Age$ > 24.5 AND $ï»¿Age$ <= 27.5 AND $Inte...	2	2
Row12	$ï»¿Age$ > 27.5 AND $Intensity$ > 1.5 AND $Dysphasia$ <= 0.5 AND $Character$ <= 1.5...	106	105
Row13	$Dysphasia$ > 0.5 AND $Character$ <= 1.5 AND $Tinnitus$ <= 0.5 AND $Vertigo$ <= 0.5 ...	2	2
Row14	$Character$ > 1.5 AND $Tinnitus$ <= 0.5 AND $Vertigo$ <= 0.5 AND $Intensity$ > 0.5 AN...	3	3
Row15	DPF <= 0.5 AND $Tinnitus$ > 0.5 AND $Vertigo$ <= 0.5 AND $Intensity$ > 0.5 AND $Vis...	2	2
Row16	$ï»¿Age$ <= 26.5 AND DPF > 0.5 AND $Tinnitus$ > 0.5 AND $Vertigo$ <= 0.5 AND $Int...	3	3
Row17	$ï»¿Age$ > 26.5 AND DPF > 0.5 AND $Tinnitus$ > 0.5 AND $Vertigo$ <= 0.5 AND $Inten...	2	1
Row18	$Phonophobia$ <= 0.5 AND $Vertigo$ > 0.5 AND $Intensity$ > 0.5 AND $Visual$ > 0.5 => ...	5	5
Row19	$Visual$ <= 1.5 AND $Defect$ <= 0.5 AND $ï»¿Age$ <= 39.0 AND $Phonophobia$ > 0.5 A...	4	4
Row20	$Sensory$ <= 0.5 AND $Tinnitus$ <= 0.5 AND $Visual$ > 1.5 AND $Defect$ <= 0.5 AND $ï...	2	2
Row21	$Sensory$ > 0.5 AND $Tinnitus$ <= 0.5 AND $Visual$ > 1.5 AND $Defect$ <= 0.5 AND $ï»...	2	1
Row22	$Tinnitus$ > 0.5 AND $Visual$ > 1.5 AND $Defect$ <= 0.5 AND $ï»¿Age$ <= 39.0 AND $Ph...	2	2
Row23	$Defect$ > 0.5 AND $ï»¿Age$ <= 39.0 AND $Phonophobia$ > 0.5 AND $Vertigo$ > 0.5 AN...	5	5
Row24	$ï»¿Age$ > 39.0 AND $Phonophobia$ > 0.5 AND $Vertigo$ > 0.5 AND $Intensity$ > 0.5 AN...	4	4

Tool Implementation

Constructing a graphical tool, with a user-friendly interface, for migraine diagnosis is highly beneficial, for both clinical and individual uses.

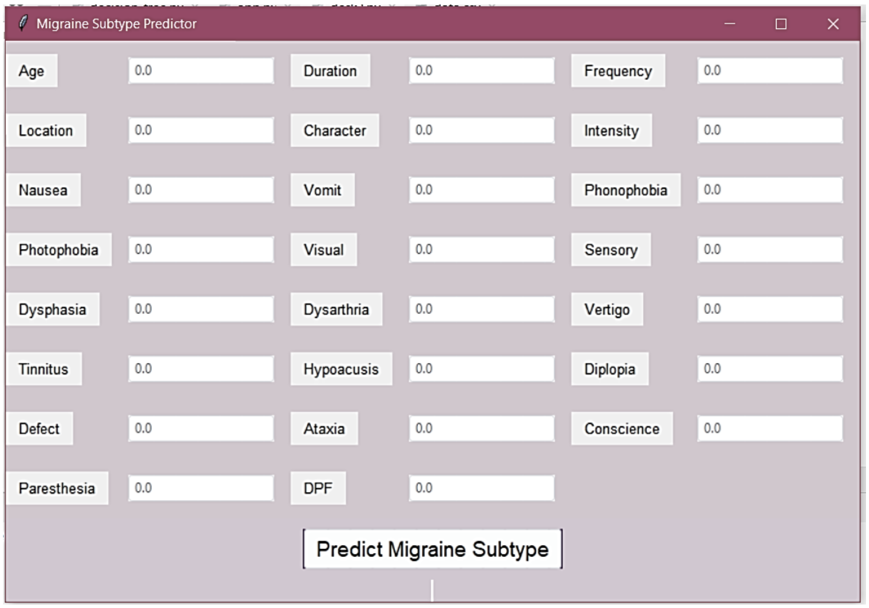
Traditional migraine diagnosis methods are often expensive and time-consuming.

The implementation of our tool, based on the machine learning models outlined in this paper, presents a cost-effective and efficient alternative.

Using the Python programming language, our tool offers a user-friendly solution to facilitate migraine diagnoses, as can be seen in Figure 6.

RESULTS AND MODEL EVALUATION

Confusion matrices have been implemented to evaluate the SVM and Decision Tree models, revealing the subtleties of their classification performance on the test set. The matrices show the correct and incorrect predictions for every subtype of migraine, providing information about both SVM and Decision Tree's possible advantages and disadvantages. Important performance indicators, including accuracy, precision, recall, and F1-score, have been extracted from the confusion matrices allowing for a more nuanced understanding of how successfully the models classified the various migraine subtypes. The SVM and Decision Tree confusion matrices are displayed in Table 5.

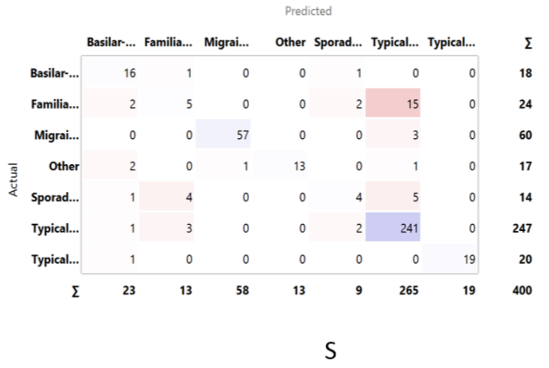


Age	0.0	Duration	0.0	Frequency	0.0
Location	0.0	Character	0.0	Intensity	0.0
Nausea	0.0	Vomit	0.0	Phonophobia	0.0
Photophobia	0.0	Visual	0.0	Sensory	0.0
Dysphasia	0.0	Dysarthria	0.0	Vertigo	0.0
Tinnitus	0.0	Hypoacusis	0.0	Diplopia	0.0
Defect	0.0	Ataxia	0.0	Conscience	0.0
Paresthesia	0.0	DPF	0.0		

Figure 6: Migraine Subtype Predictor Tool

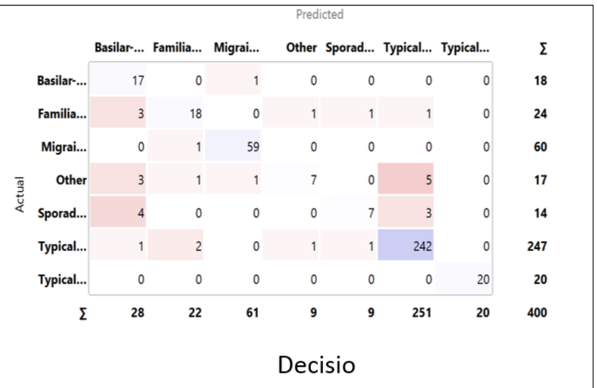
Table 6 displays the scores for each model, encompassing accuracy, Area under Curve, F1 Score, Precision, and Recall. The results are represented graphically in Figure 7.

Table 5: Confusion Matrices for SVM and Decision Tree



	Basilar-...	Familia...	Migrai...	Other	Sporad...	Typical...	Typical...	Σ
Basilar-...	16	1	0	0	1	0	0	18
Familia...	2	5	0	0	2	15	0	24
Migrai...	0	0	57	0	0	3	0	60
Other	2	0	1	13	0	1	0	17
Sporad...	1	4	0	0	4	5	0	14
Typical...	1	3	0	0	2	241	0	247
Typical...	1	0	0	0	0	0	19	20
Σ	23	13	58	13	9	265	19	400

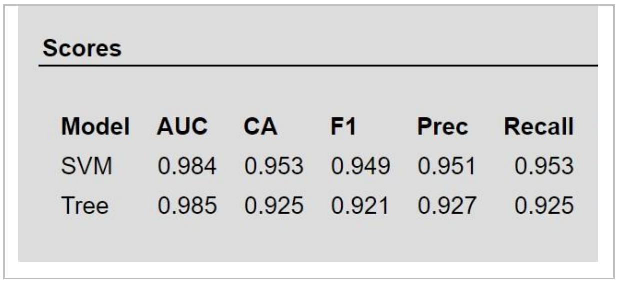
S



	Basilar-...	Familia...	Migrai...	Other	Sporad...	Typical...	Typical...	Σ
Basilar-...	17	0	1	0	0	0	0	18
Familia...	3	18	0	1	1	1	0	24
Migrai...	0	1	59	0	0	0	0	60
Other	3	1	1	7	0	5	0	17
Sporad...	4	0	0	0	7	3	0	14
Typical...	1	2	0	1	1	242	0	247
Typical...	0	0	0	0	0	0	20	20
Σ	28	22	61	9	9	251	20	400

Decisio

Table 6: SVM and Decision Tree Models Results



Scores

Model	AUC	CA	F1	Prec	Recall
SVM	0.984	0.953	0.949	0.951	0.953
Tree	0.985	0.925	0.921	0.927	0.925

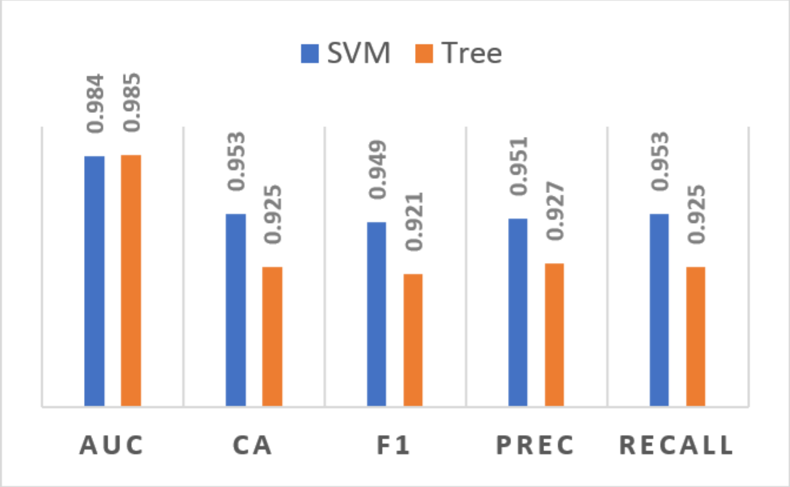


Figure 7: SVM and Decision Tree Models Results Represented Graphically

Our Support Vector Machine (SVM) model exhibited exceptional performance in classifying migraine subtypes, as evidenced by a comprehensive set of evaluation metrics. The scores of this model are:

Accuracy: A significant percentage of correctly identified occurrences across the migraine subtypes was shown by the SVM's excellent accuracy of 95.3%.

Area under Curve (AUC): With a 98.4% area under the ROC curve, the model demonstrated its strong ability to differentiate between various subtypes of migraines.

F1 Score: The model successfully balanced precise positive predictions with thorough retrieval of positive examples, as evidenced by the 94.9% F1 score, a harmonized measure of accuracy and recall.

Precision: The SVM showed a great capacity to prevent false positive predictions while maintaining a high degree of accuracy in positive classifications, with a precision score of 95.1%.

Recall (Sensitivity): Recall, or sensitivity, was 95.3%, indicating that the model was successful in identifying a large percentage of positive cases across the subtypes of migraines.

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830

The assessment of our Decision Tree model has proven to have positive outcomes, confirming its effectiveness in categorizing different forms of migraines. The scores of this model are:

Accuracy: With an accuracy of 92.5%, the Decision Tree was able to accurately classify a significant percentage of cases across the various migraine subtypes.

Area under Curve (AUC): With an incredible 98.5% Area under the Curve (AUC), the model's strong discriminating power in differentiating across migraine subtypes was highlighted.

F1 Score: The model was able to accurately anticipate positive outcomes and capture a significant proportion of positive occurrences, as demonstrated by its F1 score of 92.1%, which is a complete statistic that balances accuracy and recall.

Precision: The Decision Tree showed a remarkable capacity to reduce false positive predictions and retain a high degree of accuracy in positive classifications, earning a precision score of 92.7%.

Recall (Sensitivity): The sensitivity, or recall, was 92.5%, which demonstrated the model's ability to correctly detect a large percentage of positive cases across the various migraine subtypes.

DISCUSSION

An in-depth analysis of the Decision Tree and Support Vector Machine (SVM) models for migraine subtype classification reveals complex and persuasive performance characteristics.

In particular, the SVM outperformed the Decision Tree in terms of overall classification performance, with a somewhat better accuracy of 95.3% compared to 92.5% for the Decision Tree.

This tiny difference suggests that the SVM's inherent capacity to construct optimum decision limits contributes to its somewhat better performance on the carefully examined dataset.

The effectiveness of the SVM in capturing intricate correlations within the dataset is revealed when one delves deeper into the performance details.

The model can identify subtle patterns that might be difficult for a Decision Tree to pick up on, thanks to its support vector-defined, nuanced decision bounds.

In the case of migraine subtypes, this characteristic becomes more important since the way symptoms appear and interact may show complex interactions.

Ultimately, the little difference in accuracy between the Support Vector Machine (SVM) and the Decision Tree emphasizes the SVM's ability to both define appropriate decision boundaries and navigate the complex landscape of migraine subtype classification.

The accuracy of both models for each subtype of migraine is displayed in Figure 8 and Figure 9.

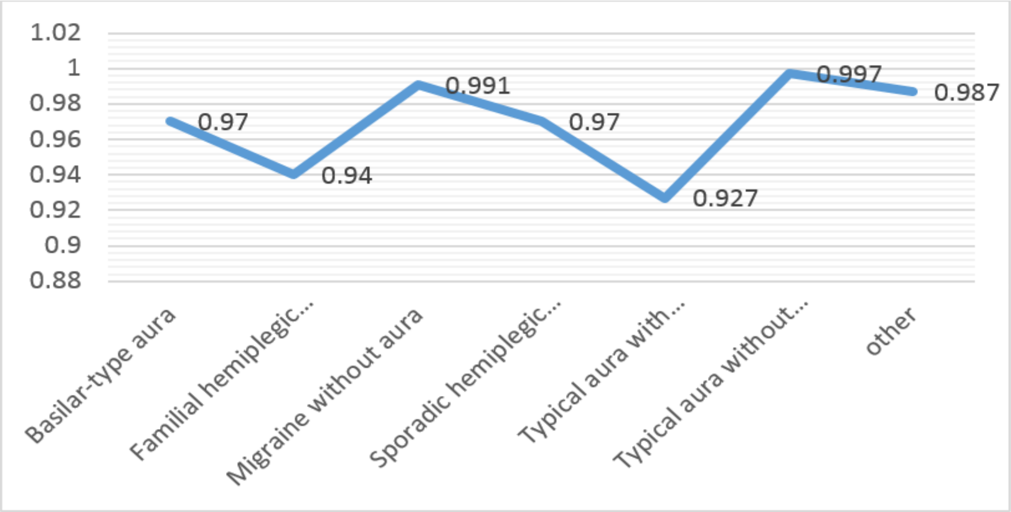


Figure 8: Accuracy of SVM for Every Subtype of Migraine

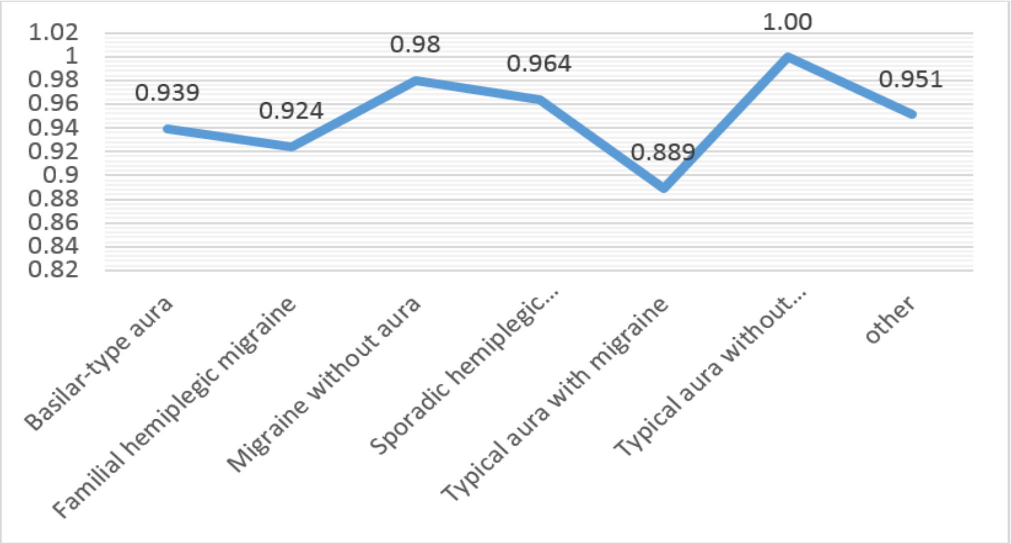


Figure 9: Accuracy of Decision Tree for Every Subtype of Migraine

CONCLUSION

Migraine is a neurological disorder, impacting 1 in 7 individuals globally. The accurate diagnosis and classification of migraine subtypes is essential for developing effective medication plans, that could significantly reduce the severity of this disorder and manage its symptoms. In this research, we employ two machine learning techniques, Support Vector Machine (SVM) and decision trees, to predict migraines and classify them into different subtypes based on patients' reported symptoms. This approach not only enhances treatment efficiency but also contributes to minimizing the possibility of medication overuse (MO) and

DOI: 10.5281/zenodo.10730830

hence, reduces rebound headaches. The results have shown that SVM overcomes decision trees in diagnosing and classifying migraine, with *95.3% accuracy. We implemented a simple tool that can be used in the clinical diagnosis of migraine.* A future work is to refine and improve the implemented migraine diagnostic tool, by integrating more advanced features within it and incorporating other machine learning algorithms.

References

1) https://www.mayoclinic.org/diseases-conditions/migraine-headache/symptoms-causes/syc-20360201
2) Patrizia Ferroni, Fabio M. Zanzotto, Noemi Scarpato, et. al. Machine learning approach to predict medication overuse in migraine patients. Computational and Structural Biotechnology journal, 2020. https://doi.org/10.1016/j.csbj.2020.06.006
3) Garcia-Chimeno, Y., Garcia-Zapirain, B., Gomez-Beldarrain, M. et al. Automatic migraine classification via feature selection committee and machine learning techniques over imaging and questionnaire data. BMC Med Inform Decis Mak 17, 38 (2017). https://doi.org/10.1186/s12911-017-0434-4.
4) Stubberud A, Ingvaldsen SH, Brenner E, et al. Forecasting migraine with machine learning based on mobile phone diary and wearable data. Cephalalgia. 2023;43(5). doi:10.1177/03331024231169244
5) Hsiao, FJ., Chen, WT., Pan, LL.H., et al. Resting-state magnetoencephalographic oscillatory connectivity to identify patients with chronic migraine using machine learning. J Headache Pain 23, 130 (2022). https://doi.org/10.1186/s10194-022-01500-1
6) Pérez-Benito FJ, Conejero JA, Sáez C, García-Gómez JM, Navarro-Pardo E, Florencio LL, Fernández-de-Las-Peñas C. Subgrouping Factors Influencing Migraine Intensity in Women: A Semi-automatic Methodology Based on Machine Learning and Information Geometry. Pain Pract. 2020 Mar;20(3):297-309. doi: 10.1111/papr.12854. Epub 2019 Dec 2. PMID: 31677218.
7) Song YY, Lu Y. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry. 2015 Apr 25;27(2):130-5. doi: 10.11919/j.issn.1002-0829.215044. PMID: 26120265; PMCID: PMC4466856.
8) (2022).https://doi.org/10.1101/2022.06.15.496326

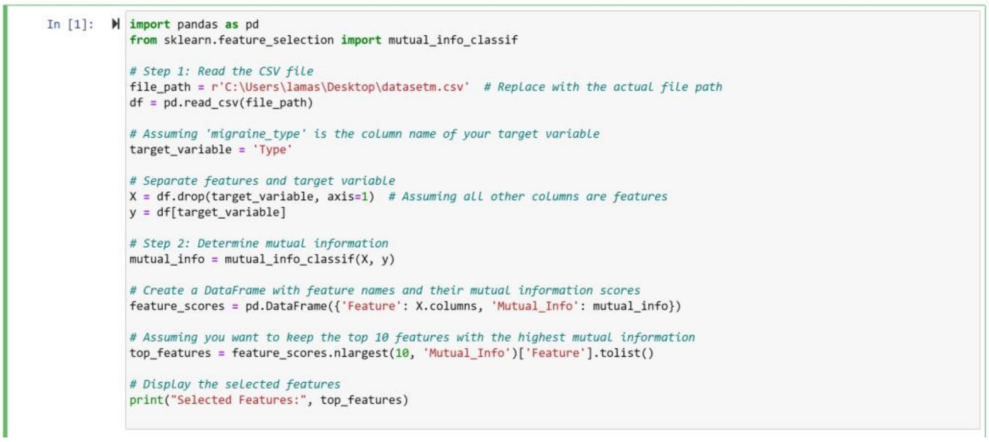
Appendix

1- Migraine dataset : https://www.kaggle.com/datasets/weinoose/migraine-classification/discussion

2- Feature selection code:

ISSN 1533-9211

DOI: 10.5281/zenodo.10730830



```
In [1]:  import pandas as pd
         from sklearn.feature_selection import mutual_info_classif

         # Step 1: Read the CSV file
         file_path = r'C:\Users\lamas\Desktop\datasetm.csv'  # Replace with the actual file path
         df = pd.read_csv(file_path)

         # Assuming 'migraine_type' is the column name of your target variable
         target_variable = 'Type'

         # Separate features and target variable
         X = df.drop(target_variable, axis=1)  # Assuming all other columns are features
         y = df[target_variable]

         # Step 2: Determine mutual information
         mutual_info = mutual_info_classif(X, y)

         # Create a DataFrame with feature names and their mutual information scores
         feature_scores = pd.DataFrame({'Feature': X.columns, 'Mutual_Info': mutual_info})

         # Assuming you want to keep the top 10 features with the highest mutual information
         top_features = feature_scores.nlargest(10, 'Mutual_Info')['Feature'].tolist()

         # Display the selected features
         print("Selected Features:", top_features)
```

3- Decision tree for migraine diagnosis: