2023 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON) | 979-8-3503-9654-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ECTIDAMTNCON57770.2023.10139707

Migraine Categorization based on the Integration of EMD and Naive Bayes Classification

Nyi Nyein Aung and Wanus Srimaharaj*
The International College, Payap University, Chiang Mai, Thailand 50000
6305040056@payap.ac.th (nyilampang@gmail.com) and wanus_s@payap.ac.th
*Corresponding Author

***Abstract*— Migraine is a disabling headache that affects over one billion people in the world and is singlehandedly responsible for 42 million years of life lost to disability in 2019. Medical experts need to analyze a patient's medical history and symptoms to determine the type of migraine. It is important to produce a conclusive diagnosis because many types of migraine also have overlapping symptoms with other medical diseases. Making an accurate diagnosis may take a long time as additional medical tests in the complexity of the case. Therefore, this study proposes a computational diagnostic method through a machine learning scheme using Naïve Bayes integrated with entropy minimization discretization (EMD). The proposed method can analyze a patient's symptoms and medical history accurately to reduce the doctor's workload and fasten the diagnosis process. The results show that the Naïve Bayes classifier with EMD achieved high accuracy results; therefore, the proposed method can be a useful method for clinical practice.**

Keywords—Migraine, Machine Learning, Classfication, Naïve Bayes, Entropy Minimization Discretization, EMD

I. INTRODUCTION

A migraine is a severe headache that appears from one side of the face, typically behind the eye, and spreads to the whole head. The onset of symptoms of migraine includes nausea, sensitivity to light and sound, and an intense pulsating sensation on one side of the face [1]. The severity of the pain caused by migraine varies from tolerable to severe. It can lead to a disabling condition as the frequency and severity of pain begin to interrupt a person's daily activities [2-3].

In clinical and translational neuroscience, it estimated that migraine has affected one billion people worldwide and globally represented 8% of all men and 20% of all women [4]. Migraine is ranked as the third highest cause of disability for males and females under the age of 50. Moreover, it caused 42.1 million years of healthy life lost due to disability in 2019 alone [6].

The two most common types of migraine are migraine without aura and migraine with aura [3]. A migraine without aura is expected to appear when a headache suddenly occurs without any prior warning. In contrast, a migraine with aura can be found when temporary sensory disturbances, such as flashing lights or blind spots. This symptom can appear before other migraine circumstances, like intense head pain and nausea. Migraines with aura tend to happen 30 minutes before the headache and last less than one hour [2]. A less common type of migraine is the hemiplegic migraine which mimics the symptoms of a stroke. It also causes weakness on one side of the body along with slurred speech.

Furthermore, the International Headache Society has established a criterion to differentiate migraine with aura and without aura and other types of migraines [5]. Doctors diagnose migraines by analyzing the patient's medical history and symptoms and a physical and neurological examination. Definitive diagnosis may require computerized tomography or magnetic resonance imaging. Typically, it is complicated by other medical conditions with the same symptoms, requiring expert knowledge to make a valid diagnosis.

Despite a growing global prevalence of migraines, there is a lack of headache specialists causing a rapid demand for digital diagnostic tools. It is estimated that 52% of the world suffers from active headache disorder every year [7]. There has been a 4.5-fold increase in the development and assessment of tools for headache diagnosis in the last decade [8]. Since the technology in artificial intelligence has been developed continuously, the machine learning method has been proposing to solve several kinds of medical symptoms including migraine. Supervised learning is a type of machine learning that can apply medical training data in which contain inputs and correct outputs for further clinical prediction. This method can learn the previous statistics of patients to produce the desired output over time. Data classification is very helpful for medical diagnosis because it lessens the doctors' workload and supports the clinical prediction accuracy [9]. This classification process can group the new input data into predefined categories. Classification algorithms have been used in many fields of medicine, such as the classification of medical images [10-11], and heart disease classification [12]. While the machine learning scheme can be used for the determination cof headaches [13], it should also be applicable for the identification of different types of migraines.

Although a computerized diagnostic method by medical professionals can determine the type of migraine, it requires further studies and alternative tools implementing in urgent cases. Therefore, this study proposes an adoptive classification method based on an analysis of the patient's medical history and symptoms using supervised learning techniques, Naïve Bayes classification with the entropy minimization discretization (EMD). The EMD was applied for the data preprocessing technique while the Naïve Bayes classification algorithm is mainly performed as the classification process and provides the expecting results in migraine type identification.

II. LITERATURE REVIEW

Computer software systems are very helpful for the accurate identification of headaches. Automatic diagnosis of primary headaches was proposed by utilization of machine learning algorithms [15]. It found that the diagnostic performance of machine learning technology of the proposed computer decision support system attained a greater diagnostic accuracy than the diagnostic performance of clinicians. Meanwhile, the clinical decision support system (CDSS) was proposed for the automatic diagnosis of probable migraine and probable tension-type headache on case-based reasoning [16]. It was constructed for a case library by analyzing 676 clinical cases in the Chinese PLA general hospital. This method offerred a better accuracy in the identification of primary headaches achieved by the CBR-based CDSS than the guideline-based CDSS. Previously, an early classification system for automatic migraine classification using an artificial neural network (ANN) was proposed [17]. The system used a multilayer perceptron (MLP) with a varying number of hidden layers to achieve 97% of accuracy and precision levels. In addtion, the aforementioned classification models with great accuracy and precision levels can be adopted in different situations. However, there are many barriers preventing the widespread use of these clinical decision support systems in clinical practices. A futher research needs to be done using different classification methods to develop a better understanding in clinical prediction. The medical professionals can practically use the CDSS in their diagnostic process.

Therefore, this study applies a machine learning method, Naïve Baye to produce a classification method for identifying the type of migraine automatically.

III. Research Methodology

The Naïve Bayes-based classification model proposed by this study is composed of 4 steps: data collection, data preprocessing, data classification, and data testing, shown in Figure 1.

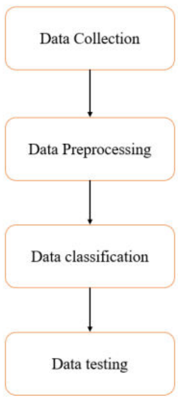

Figure 1. Data Collection

A. Data Collection

The study applied 400 medical records of patients diagnosed data with various symptoms related to migraines. It was collected by trained medical professionals at Hospital Materno Infantil de Soledad in the first quarter of 2013 [14]. Table 1 contains 23 variables of the dataset related to the patient's symptoms, and medical history: age, duration, frequency, location, character, intensity, nausea, vomit, phonophobia, photophobia, visual, sensory, dysarthria, vertigo, tinnitus, hypoacusis, diplopia, visual defect, ataxia, conscience, paresthesia, and family background (DPF).

TABLE 1 Descriptive Statistics of Features in the Dataset

Variable	Mean	Standard Deviation	Standard Error	Confidence Interval
Age	31.705	12.139	0.607	1.193
Duration	1.61	0.771	0.039	0.076
Frequency	2.365	1.676	0.083	0.165
…	…	…	…	…
Conscience	0.018	0.131	0.007	0.013
Paresthesia	0.008	0.086	0.004	0.008
DPF	0.41	0.492	0.025	0.048

All variables denote the continuous numeric data types excluding the type variables, which represent the diagnosis of migraine on the patient's medical records and symptoms made by a doctor. The seven types of migraine and the number of instances for each type are shown in Figure 2.

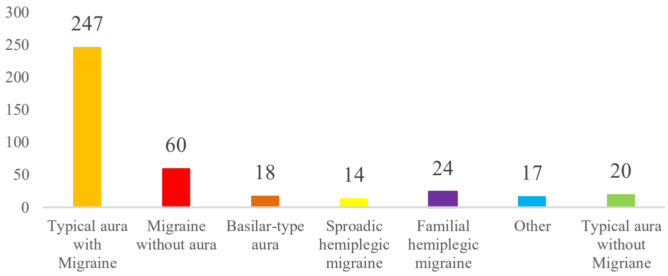

Figure 2. Classification of Migraine Types

B. Data Preprocessing

This study applied the entropy minimization discretization (EMD) to filtering input data [19]. The EMD is a normalization technique that transforms a numeric attribute into a categorial attribute and simplifies the data into more digestible information. EMD is a supervised top-down splitting discretization technique. According to the Equation (1), it selects a value from the whole attribute range, discretizes the data into two intervals, and calculates the entropy or information gained.

$$Entropy(D_1) = -\sum_{i=1}^{m} p_i \log_2 p_i \quad (1)$$

$m = number\ of\ classes\ in\ the\ dataset, i = an\ element\ of\ class$
$p_i = probability\ of\ randomly\ picking\ element\ i$

The minimum entropy is chosen as the split point. It will be recursively divided into more intervals until the minimal description length (MDL) stopping criterion is met [19].

Since the discretization is optimal to a dataset with large amounts of numeric attributes, presenting in this study, it reduces noise by grouping all the data into groups of categories. This technique reduces the skewness of the dataset and improves the dataset's internal consistency. The Naive Bayes classifiers integrated with the EMD are promising to achieve lower classification errors than those using common probability density assumptions [20].

C. Data Classification

This study performed a classification process using the Naïve Bayes. The classifier is called 'Naïve' because it assumes all attributes to be independent given the class [18-21]. Naïve Bayes has been selected as the main method of this study as it is based on probability, assuming to strong independence between features. Naïve Bayes is a simple classifier based on Bayes' theorem that calculates the probability of an attribute being assigned to a class, shown in Equation (2).

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (2)$$

$A = piror\ probility, B\ = marginal\ probability$
$P(A|B) = posterier\ probability$
$P(B|A) = likelihood\ probability$

In general, the unrealistic assumption of independence of is notably effective in practice since the classification decision turns out to be correct when the probability estimates are inaccurate. For nominal attributes, the Naïve Bayes uses the multinomial to simply counts the number of times an event occurs for each attribute and normalizes the counts into consequent probability [18]. This approach takes a transpired event and predicts the probability of the various known causes. Also, the assumption of independence between features allows Naïve Bayes to be highly scalable with the number of predictors. The data points improves the practicality and the ease of computational implementation for medical purposes.

For the performance and reliability comparison, the K-Nearest Neighbor, Random Forest, MLP Neural Network, j48 Decision Tree, and Naïve Bayes without EMD algorithms were also applied to the data without any preprocessing method.

D. *Data Testing*

The data is tested using the holdout method which splits the dataset into two parts: a training set and a testing set. Out of the 400 instances in the dataset, 80% (320) correspond to the training set and 20% (80) correspond to the testing set. The data is split into 80/20 based on the Pareto Principle, which states that 80% of all outcomes come from 20% of the causes. It has also been proven that the 80-20 split offers the best result [20].

IV. Experimental Results

Table 2 shows the result of the experiment. The rows stand for different classifier algorithms accuracy. The output value of Naïve Bayes with EMD represents the best outcome in the respective column.

TABLE 2 Performance Metrics of the Classification Models based on Accuracy

Classifier Algorithm	Accuracy
Naïve Bayes with EMD	**97.5%**
Naïve Bayes without EMD	96.25%
KNN	85%
J48 decision tree	87.5%
Random Forest	90%
MLP	92.5%

The results from the classification models are verified by 10-fold cross validation, comparing the results with the diagnosis made by the treating physician.

As a result, in Table 3, it shows that every classification model achieved accuracy levels of 85% or higher which means that theoretical machine learning methods can achieve good quality of classification of migraines.

The proposed method of this study, Naïve Bayes integrated with EMD achieved the maximum accuracy in the experiment of 97.5%. This means 78 of the 80 test cases match with classification of migraine made by proposed technique. The highest recall value of 0.98 was also obtained by the proposed method. Nevertheless, the Naïve Bayes integrated with EMD attained a better accuracy than Naïve Bayes without any feature scaling (97.5 % versus 96.25%). It shows that data preprocessing plays a role in improving the accuracy of the Naïve Bayes classifier.

Further analysis of the experiment results shows that the proposed method produced better accuracy levels than other machine learning algorithms with a noticeable gap. The proposed method successfully produced an accurate classification model despite the small dataset size. The proposed method is flexible enough to implement in different conditions because it requires minimal data pre-processing and does not need complicated features to produce accurate results. Due to the simplicity of the proposed method, there is no overfitting observed during the experiment, shown in Table 3.

TABLE 3 The Classification Report for Naïve Bayes Integrated with EMD

Class	Precision	Recall	f1-score	MCC
Typical Aura with Migraine	0.96	1.00	0.98	0.95
Migraine without Aura	1.00	1.00	1.00	1.00
Typical Migraine without Aura	1.00	1.00	1.00	1.00
Familial Hemiplegic Migraine	1.00	0.6	0.75	0.76
Sporadic Hemiplegic Migraine	1.00	1.00	1.00	1.00
Basilar-Type Aura	1.00	1.00	1.00	1.00
Others	1.00	1.00	1.00	1.00
Average	0.976	0.975	0.972	0.954

Predicted Type \ Actual Type	Typical aura with migraine	Migraine without aura	Basilar-type aura	Sporadic hemiplegic migraine	Familial hemiplegic migraine	Other	Typical aura without migraine
Typical aura with migraine	48	0	0	0	0	0	0
Migraine without aura	0	17	0	0	0	0	0
Basilar-type aura	0	0	2	0	0	0	0
Sporadic hemiplegic migraine	2	0	0	3	0	0	0
Familial hemiplegic migraine	0	0	0	0	3	0	0
Other	0	0	0	0	0	1	0
Typical aura without migraine	0	0	0	0	0	0	4

Figure3. Confusion Matrix of Naïve Bayes Integrated with EMD

Figure 3 shows the confusion matrix of Naïve Bayes Integrated with EMD. It shows that the proposed method predicted two instances as sporadic hemiplegic migraine instead of typical aura with migraine.

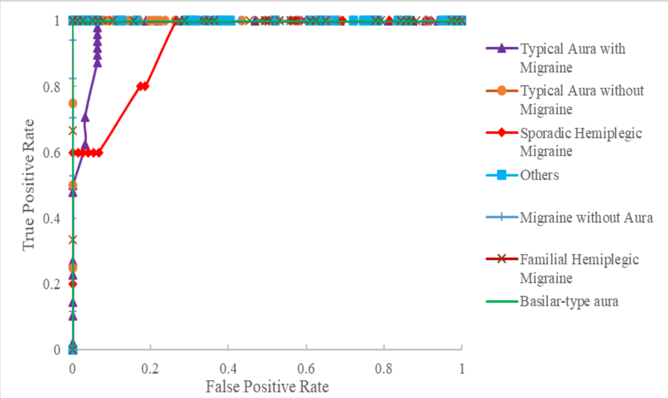

Figure 4. ROC Curve for Naïve Bayes Integrated with EMD

The classification performance of the best model by the proposed method was compared to prior works on the classification of migraine using accuracy, shown in Table 4.

TABLE 4 COMPARISON OF ACCURACY FOR MIGRAINE CLASSIFICATION OBTAINED FROM PREVIOUS STUDIES

Algorithms	Classification Methods	Accuracy
This study	Naïve Bayes	97.5%
Classification of multi-channel EEG signals for migraine detection [23]	SVM	88.4%
Automatic migraine classification using artificial neural networks [17]	ANN	97.5%
Automatic diagnosis of primary headaches by machine learning methods [15]	Random Forest	81%
Effect of Flash Stimulation for Migraine Detection Using Decision Tree Classifiers [24]	Random Forest	85.18%

As shown in Table 4, the classification of migraines using machine learning methods has been explored previously. The proposed Naïve Bayes method (97.5 accuracy level) achieved better performance compared to the methods in [15,23-24] 88.4%, 81%, and 85.18% respectively. The multilayer perceptron type artificial neural network method in the existing work [17]. achieved the same accuracy of 97.5% as the proposed method of this study which shows that the proposed method has the classification capabilities of a deep learning classification algorithm for the classification of different types of migraine.

V. CONCLUSION

Migraine is a type of primary headache that causes intense head pain and nausea. Migraine affects over 1 billion people each year [4] and is solely responsible for 42.1 million years of life lost due to disability in 2019 [6]. This study proposed the development of a computerized diagnostic method for migraine classification using a supervised machine learning algorithm, Naïve Bayes, with entropy minimization discretization (EMD).

The results show Naïve Bayes integrated with EMD provided an excellent performance result. It achieved higher accuracy and precision levels than other machine learning classification methods. The proposed method also attained the same accuracy of 97.5% as the artificial neural network method proposed in the existing work [17]. This demonstrates the usefulness of the proposed method for the classification of different types of migraine.

In the future work, additional studies with different datasets are needed to further optimize the classification performance and determine the validity of the proposed method in real-life situations.

ACKNOWLEDGMENT

This research work is supported by the International College, Payap University.

REFERENCES

[1] National Health Service, “Migraine,” NHS choices, 10-May-2019. [Online]. Available: https://www.nhs.uk/conditions/migraine/. [Accessed: 03-Feb-2023].

[2] Mayo Clinic Staff, “Migraine,” Mayo Clinic, 10-May-2019. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/migraine-headache/symptoms-causes/syc-20360201. [Accessed: 20-Feb-2023].

[3] WebMD Editorial Contributors, “Migraine: Symptoms, causes, diagnosis, treatment, and prevention,” WebMD, 13-Dec-2022. [Online]. Available: https://www.webmd.com/migraines-headaches/migraines-headaches-migraines. [Accessed: 20-Feb-2023].

[4] Diener et al, “Treatment of migraine attacks and prevention of Migraine: Guidelines by the German migraine and headache society and the German Society of Neurology,” Clinical and Translational Neuroscience, vol. 3, no. 1, p. 2, 2019.

[5] Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. Cephalalgia. 2018;38(1):1-211. doi:10.1177/0333102417738202

[6] “Migraine - level 4 cause,” Institute for Health Metrics and Evaluation, 15-Oct-2020. [Online]. Available: https://www.healthdata.org/results/gbd_summaries/2019/migraine-level-4-cause. [Accessed: 20-Feb-2023].

[7] L. J. Stovner, K. Hagen, M. Linde, and T. J. Steiner, “The global prevalence of headache: An update, with analysis of the influences of methodological factors on prevalence estimates,” The Journal of Headache and Pain, vol. 23, no. 39, Apr. 2022.

[8] Y. W. Woldeamanuel and R. P. Cowan, “Computerized migraine diagnostic tools: A systematic review,” Therapeutic Advances in Chronic Disease, vol. 13, p. 20406223211065235, Jan. 2022.

[9] L. Wang and K. Zuo, “Medical Data Classification assisted by Machine Learning Strategy,” Computational and Mathematical Methods in Medicine, 10-Sep-2022. [Online]. Available: https://www.hindawi.com/journals/cmmm/2022/9699612/. [Accessed: 19-Feb-2023].

[10] S. Ismael, S. Kareem, and F. Almukhtar, “Medical Image Classification using different machine learning algorithms,” AL-Rafidain Journal of Computer Sciences and Mathematics, vol. 14, no. 1, pp. 133–145, 2020.

[11] H. Tang and Z. Hu, “Research on medical image classification based on machine learning,” IEEE Access, vol. 8, pp. 93145–93154, 2020.

[12] [12] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, “Heart disease identification method using machine learning classification in E-Healthcare,” IEEE Access, vol. 8, pp. 107562–107582, 2020.

[13] J. Kwon, H. Lee, S. Cho, C.-S. Chung, M. J. Lee, and H. Park, “Machine learning-based automated classification of headache disorders using patient-reported questionnaires,” Scientific Reports, vol. 10, no. 1, 2020.

[14] Paola A. Sánchez-Sánchez, José Rafael García-González, Juan Manuel Rua Ascar (2020) Migraine Classification Model [Source Code]. https://doi.org/10.24433/CO.2826453.v1

[15] B. Krawczyk, D. Simić, S. Simić, and M. Woźniak, “Automatic diagnosis of primary headaches by machine learning methods,” Open Medicine, vol. 8, no. 2, pp. 157–165, Jan. 2013.

[16] Z. Yin, Z. Dong, X. Lu, S. Yu, X. Chen, and H. Duan, “A clinical decision support system for the diagnosis of probable migraine and probable tension-type headache based on case-based reasoning,” The Journal of Headache and Pain, vol. 16, no. 29, Apr. 2015.

[17] P. A. Sanchez-Sanchez, J. R. García-González, and J. M. Rúa Ascar, “Automatic migraine classification using Artificial Neural Networks [version 2; peer review: 1 approved, 2 approved with reservations],” F1000Research, vol. 9, p. 618, 2020.

[18] M. A. Hall and I. H. Witten, Data Mining: Practical Machine Learning Tools and techniques 3rd edition ed. 3. Burlington, Vermont: Morgan Kaufmann Publishers, 2011.

[19] U. Fayyad and K. Irani, “Multi-interval discretization of continuous-valued attributes for classification learning,” 1995.

[20] Y. Yang and G. I. Webb, “Discretization for Naive-Bayes Learning: Managing discretization bias and variance,” Machine Learning, vol. 74, no. 1, pp. 39–74, 2008.

[21] R. Irina, “An empirical study of the naive Bayes classifier,” IJCAI 2001 workshop on empirical methods in artificial intelligence, vol. 3, no. 22, pp. 41–46, 2021.

[22] Gholamy, Afshin; Kreinovich, Vladik; and Kosheleva, Olga, "Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation" (2018). Departmental Technical Reports (CS). 1209. https://scholarworks.utep.edu/cs_techrep/1209

[23] B. S. Akben, D. Tuncel, and A. Alkan, “Classification of multi-channel EEG signals for migraine detection ,” Biomed Res, vol. 27, no. 3, pp. 743–748, 2016.

[24] A. Subasi, A. Ahmed, and E. Alickovic, “Effect of flash stimulation for migraine detection using decision tree classifiers,” Procedia Computer Science, vol. 140, pp. 223–229, 2018.