DOI: 10.24996/ijs.2025.66.2.20

ISSN: 0067-2904

ADASYN Oversampling and SGB Ensemble Algorithm for Migraine Classification

Israa Mohammed Hassoon

Department of Mathematics, Collage of Science, University of Mustansiriyah (UOM), Baghdad-Iraq

Received: 5/9/2023 Accepted: 27/2/2024 Published: 28/2/2025

Abstract

A migraine is a neurological disorder that causes severe headaches. Although it is not life-threatening, it greatly affects people's lives. Migraine classification is considered a complicated research area. Early detection and accurate classification help in determining the appropriate care, speeding recovery from the disease, and avoiding its effects. So, developing a migraine classification system is very necessary. The aim of this work is to suggest a migraine classification model based on a stochastic gradient-boosting ensemble algorithm. First, the dataset is pre-processed. Adaptive Synthetic (ADASYN) is used to balance a migraine dataset that contains seven groups of migraine patients. Twenty-three migraine attributes were extracted from 400 patients. The ANOVA feature selection method is applied to select the most relevant features. Four experiments have been carried out based on apply/not apply ADASYN or ANOVA. The SGB model is trained using the number of hyperparameter tunings used in the four experiments in order to improve model performance. The SGB model is evaluated in terms of precision, recall, f1-score, and accuracy. The results showed that SGB achieved outstanding outcomes in the fourth experiment when applying ADASYN and ANOVA. The model achieved 97.01% accuracy, 0.970905 precision, 0.966250 recall, and a 0.967278 f1-score.

Keywords: Stochastic Gradient Boosting, Analysis of Variance, Features Selection, ADASYN, Migraine.

خوارزمية SGB وترميز ADASYN لتصنيف الصداع النصفي

اسراء محمد حسون

قسم الرياضيات، كلية العلوم، الجامعة المستنصرية، بغداد, العراق

الخلاصة

الصداع النصفي هو اضطراب عصبي يسبب صداعًا شديدًا، على الرغم من أنه لا يهدد الحياة، إلا أنه يؤثر بشكل كبير على حياة الناس. يعتبر تصنيف الصداع النصفي مجال بحث معقد. يساعد الاكتشاف المبكر والتصنيف الدقيق في: تحديد الرعاية المناسبة، وتسريع الشفاء من المرض وتجنب آثاره. وبالتالي، هناك حاجة إلى تطوير نظام لتصنيف الصداع النصفي. الهدف من هذا العمل هو اقتراح نموذج تصنيف للصداع النصفي يعتمد على خوارزمية تعزيز التدرج العشوائي. بدايةً تتم معالجة مجموعة البيانات. يستعمل Adaptive Synthetic (ADASYN)لموازنة مجموعة بيانات الصداع النصفي التي تحتوي على سبع مجموعات من مرضى الصداع النصفي. تم استخراج 23 سمة من سمات الصداع النصفي من 400 مريض.

Email: isrmo9@uomustasiriyah.edu.iq

يتم تطبيق طريقة اختيار ميزات ANOVA لتحديد الميزات الأكثر صلة. تم إجراء أربع تجارب بناءً على تطبيق / عدم تطبيق ADASYN / ANOVA .تم تدريب نموذج SGB باستعمال عدد من ضبط المعلمات الفائقة الذي تم استعماله في التجارب الأربع من أجل تحسين أداء النموذج. تم تقييم نموذج SGB من حيث: الدقة والحساسية ودرجة f1 والنوعية، وأظهرت النتائج أن SGB حقق نتائج تصنيف متميزة في التجربة الرابعة عند تطبيق ADASYN و ANOVA .حقق النموذج دقة 97.01٪، حساسية 0.970905، نوعية 0.966250 و0.967278 درجة f1.

1. Introduction

Migraine is a chronic neurological disorder characterized by recurring, moderate-to-severe headaches. Migraine pains usually affect the right or left half of the head. There are several types of migraine based on symptoms; the most prevalent is classical migraine. Other types include:

silent migraine, ocular migraine, migraine without aura, and migraine with aura. Despite the development in the medical field, the diagnosis of migraine still relies on symptoms, questions asked by the doctor, and medical examinations [1]. On the other hand, people do not always visit doctors in the early stages of disease symptoms, and when they do, doctors may make a misdiagnosis. Misdiagnosed or undiagnosed migraine may occur due to its overlapping symptoms with other neurological diseases such as tension-type headache and sinus-type headache or because of the absence of symptoms [2].

Early diagnosis of migraine and obtaining appropriate health care with early treatment help to significantly reduce the severity of the disease symptoms. It is paramount to design an automated migraine diagnosis system that can detect migraines at an early stage. Ensemble algorithms are machine learning algorithms that merge the predictions of multiple classifiers in order to improve model performance. Ensemble algorithms combine different ML algorithms to obtain more accurate results than a single algorithm. The purpose of using ensemble algorithms is to decrease generalization errors [3].

The purpose of this work is to design a migraine classification system based on an ensemble learning algorithm. Determining the type of migraine helps determine the severity of the disease and then determines the type of care and treatment required. In addition to the migraine classification itself, the feature selection combines with an ensemble algorithm to reduce the number of features and improve system performance. Features selection techniques play an influential role in data pre-processing. It is used to increase accuracy by choosing the most fundamental features and removing the unimportant ones. A small number of features is extremely desirable because it decreases model complexity [4].

Another factor affecting model accuracy is database balancing. Balancing data is a dataset where each category has the same number of inputs. Balancing a dataset makes the training process simpler because it prevents the system from being biased toward one category. Balancing data can be performed by utilizing over- and under-sampling techniques [5]. Oversampling techniques add new samples to the minority group. It utilized it considerably more than under-sampling techniques. Under sampling techniques, data is eliminated from the majority group, causing the loss of significant information. Oversampling techniques can yield better outcomes than undersampling techniques [6].

The main contributions of the presented work are:

1-Developed an ensemble learning model to accurately classify seven different types of migraine.

2-Demonstrated the impact of ADASYN/ANOVA on model performance, with ADASYN for balancing the dataset and ANOVA for feature selection.
3-Conducted investigations into the effectiveness of combining the SGB algorithm with ADASYN oversampling and ANOVA feature selection for classifying a wide variety of migraine types.
The proposed work is organized as follows: In Section 2, the related works are revised. In Section 3, various methods and materials are explained. In Section 4, the experimental results of the migraine classification model are provided. Section 5 concludes this work.

2. Related Works

The author has presented a brief review of related studies, indicating a literature survey was conducted.

In [2], a method for detecting migraines based on three distinct classifiers—RF, CART, and C4.5—is shown. Variations in window lengths were applied. They used EEG signals obtained from the Department of Kahramanmaras Sutcu Imam University for their research. The 60 individuals in the dataset were divided into two classes: nine females and twenty males with migraines, and eleven males and nineteen females in the second class who were in good health. The data was recorded using an eighteenth-channel Nichollet One system. The first involved using DWT to break down EEG data; the second involved extracting statistical features; and the third involved determining the relationship between accuracy and flash stimulation. They experimented with various lengths of windows; the window with a length of 768 samples produced the best results, with a high accuracy of 85.18% reached using random forest. Their work showed that patients with migraines have been found to be more sensitive to triggers like flash stimulation. The results of their investigation have implications for the brain's physiological reaction.

In [7], a system is presented to classify migraines based on three algorithms: naive Bayes boosting and SVM. Questionnaires and DTI images were used; the dataset contained 52 instances that were divided into four groups. 41 attributes were extracted, and the number of attributes was reduced by using four feature selection methods: gradient tree boosting, univariate, random forest, and L1-based. Each of these methods selected the relevant features, which were used by one of the three classifiers. They were able to attain a higher percentage in their work because features from migraine-specific questionnaires were included in the classification that was generated using the features derived from DTI images. Their work was evaluated in two terms: accuracy and f1-score. The results showed that SVM achieved the best accuracy of 90 to 95%. Even with the best outcomes, their work had some limitations, such as: 1) the whole dataset of their work contains fewer than 20 people per class, indicating a limited sample size. Consequently, a smaller number of training samples may contribute to the overfitting issue, and the classification method may be modified to account for extremely particular aspects of the training set. and 2) the inability of model performance metrics to account for unobserved data.

[8] presented a new SSEP approach to classifying migraines based on machine learning methods. The dataset contained 42 patients: 29 interictal and 13 ictal, which were compared with 15 healthy. The SSEP signals were recorded based on healthy subjects and migraineurs in two steps: ictal and interictal. The signals were plotted and construed to find the DCA, and then FT was used to extract the biomarkers. Handcrafted and one-way analysis of variance methods were used to select relevant features, and then deep CNN and hyperparameter tuning were performed to obtain the highest accuracy. Their method's effectiveness was assessed by the application of 10-fold cross-validation. The f1-score, specificity, sensitivity, and accuracy of the model were assessed. Their method produced an ictal accuracy of 89.7% for healthy

individuals. Their work had some limitations, such as: 1) the absence of prospective data about the lateralization of MWA assaults, which may have an impact on neuroimaging investigations. 2) Further migraine patients should be examined in order to validate the validity of the approach and the results, as there was insufficient evidence to support their technique.

[9] proposed a system for initial headache diagnosis based on hybrid machine learning. Data for 614 real patients was collected at different hospitals in Jordan. 26 ML classifiers were used with the default number of sections (10); each classifier uses nine sections for training data and one section for testing data. This process was repeated in each classifier stage. Nine feature selection methods were used to reduce the number of features. 19 features were selected. At last, the system could classify three types of headaches: migraine, tension headache, and cluster headache. The highest accuracy was obtained by K-means and random forest, with migraine accuracy of 99.1% and 92.7%, respectively, and an overall accuracy of 93%. Their work showed that SMOTE data augmentation helped to reduce dataset bias and increased the accuracy of the classifiers on an unbalanced dataset.

In [10], a hybrid intelligent approach for primary headache diagnosis is proposed. A number of techniques and methods were used to construct their hybrid approach: the weighted FCM algorithm, the Calinski-Harabasz index, and the analytical hierarchy process. The approach classified primary headaches into three classes: 1) tension headaches; 2) migraines; and 3) other primary headaches. The dataset included 479 patients; 224 of them had tension headaches, 169 of them had migraines, and 186 had another type of primary headache. Twenty attributes were extracted from the dataset; nine of them were selected at first, and then the five most relevant attributes were selected. Their approach was evaluated in terms of f1-score, accuracy, and precision. The accuracy of their approach was 75%. Their work revealed the relationship between migraine and TTH in two newly created cluster groups that were created through the application of cluster analysis. Their work has an advantage in that it uses a hybrid intelligent system for primary headache diagnosis and calculates computational complexity. It also evaluates the application of the hybrid intelligent system for primary headache diagnosis experimentally and compares it with a few chosen state-of-the-art methods

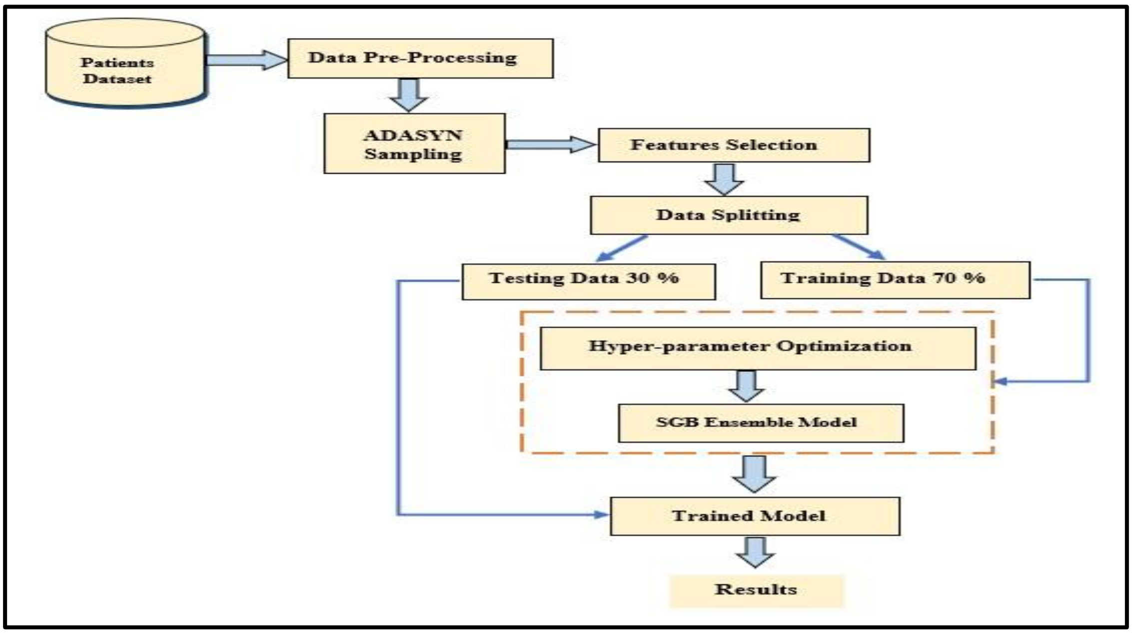


Figure 1: The Presented Work Architecture

Discussion of the Presented Work Architecture

The architecture of the proposed work is illustrated in a block diagram (Fig. 1). The figure outlines the workflow of the migraine classification system, including data preprocessing, feature selection, model training, and testing steps.

Figure 1 displays the phases of the migraine classification system; initially, data preprocessing and visualization were carried out on the migraine dataset. After that, feature selection is performed. The dataset is split into 70% for training and 30% for testing. For classification, five hyper-parameters are used with the SGB model. In this work, four experiments are performed: In the first experiment, migraine classification is implemented by using the initial unbalanced dataset with all the dataset attributes. A hyperparameter tuning of model parameters is utilized to tweak model performance for high accuracy. The SGB model used the following hyperparameters for the four experiments: n_estimators=100, random_state=42, learning_rate=0.5, max_depth=5, tol=0.0002. The results of the first experiment were: 86.87% accuracy, 0.913842 precision, 0.718785 recall, and 0.785565 f1-score. In the second experiment, the ADASYN oversampling method is applied to the dataset, and a new balanced migraine dataset is used with full dataset attributes. The results of the second experiment were improved to: 94.82% accuracy, 0.946789 precision, 0.944935 recall, and 0.945333 f1-score. In the third experiment, an initial unbalanced dataset and new selected features are utilized, and the new selected features are obtained by applying the ANOVA feature selection method. The results have decreased to: 83.84% accuracy, 0.760652 precision, 0.652735 recall, and 0.680159 f1-score. In the fourth experiment, the ADASYN and ANOVA methods are applied. With 97.01% accuracy, 0.970905 precision, 0.966250 recall, and 0.967278 f1-score, the SGB model did a great job when used with a balanced dataset that included the newly chosen features.
The detailed discussion of the presented work phases is given in the following sections.

3.1 Data Set and Pre-processing

Pre-processing is a standard step in data analysis to ensure the quality and consistency of the data before applying machine learning algorithms. Even if a dataset appears clean, pre-processing may include operations such as encoding categorical variables, normalizing or scaling features, and removing duplicates to prepare the data for modeling. Pre-processing can also involve more nuanced steps, like feature selection, to improve model performance by removing irrelevant or redundant features. The mention of pre-processing does not necessarily imply that the data was dirty or unclean; it is part of the workflow to optimize the data for the specific requirements of the chosen algorithms.

In the presented study, a significant dataset is imported from [11]. It contains 400 instances, which are divided into seven classes of migraine. The dataset contains 24 columns, 23 features, and 1 result. The result is a category label used to classify migraines into seven classes. The contents of the result column should be converted to numerical values. The seven classes of migraine have values ranging from 0 to 6, respectively. For more details on the migraine dataset, see Table 1 below.

Table 1: Migraine dataset details

Dataset	No. of patients	Attributes	Classes
Migraine Dataset [11]	400	23 attributes: 1-Age 2-Duration 3-Frequency 4-Location 5-Character 6-Intensity 7-Nausea 8-Vomit 9-Phonophobia 10-Photophobia 11-Visual 12-Sensory 13-Dysphasia 14-Dysarthria 15-Vertigo 16-Tinnitus 17-Hypoacusis 18-Diplopia 19-Defect 20-Ataxia 21-Conscience 22-Paresthesia 23-DPF	7 classes: 1-Typical aura with migraine 2-Typical aura without migraine 3-Migraine without aura 4-Basilar-type aura 5-Sporadic hemiplegic migraine 6-Familial hemiplegic migraine 7-Other

Pre-processing operations include removing/filling missing data, encoding data, and removing duplicate data. Missing values are checked; in this dataset, there are no missing values. Duplicate data is removed. The migraine dataset is unbalanced; it contains seven classes with unequal numbers of patients. Balancing data is needed, and the ADASYN oversampling method is utilized [13].

3.2 Data Visualization

Data visualization and data correlation analysis are used to improve the migraine dataset [12]. Data visualization is the process of translating system information into a visual form to make it easier to understand. Data visualization includes different techniques that help in understanding and analyzing complex and big data, such as charts, histograms, scatter plots, boxplots, and maps. Data correlation analysis is utilized to illustrate the strength of association between variables; it can manifest the direction of a two-variable relationship by using quantitative methods [8]. Data correlation analysis, which is employed to measure the relationship between variables, is illustrated in Figure 2.

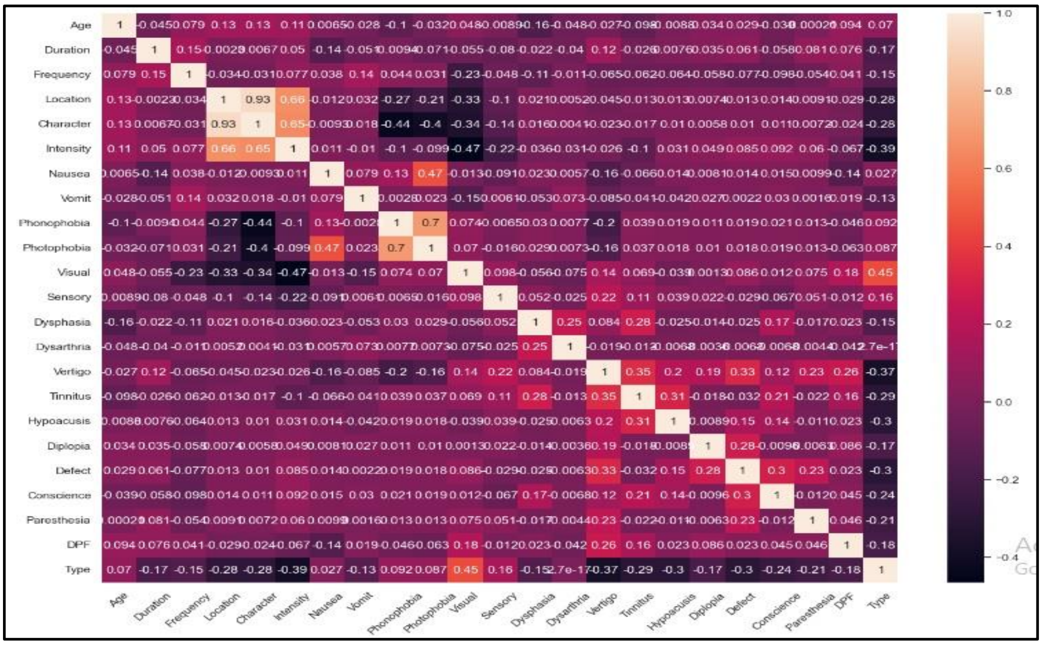


Figure 2: Correlation heatmap revealing the degree of relationship between variables

3.3 ADASYN Oversampling

It is a method used to solve the problem of an unbalanced database based on a local distribution estimation of the category to be oversampled. The main idea of the ADASYN method is to use a weighted distribution for minority categories depending on the grade of difficulty of learning [4]. ADASYN utilized the nearest neighbor technique to inset more imitation samples for the minority categories exhibited to balance the dataset. There are two advantages to using ADASYN: 1) Minimizing bias caused by unbalanced categories 2) changing the prediction decision boundary toward difficult samples [14].
The algorithm 1 below explains the steps of ADASYN [15]:

Algorithm 1: Steps of ADASYN method
Input: Unbalanced dataset
Output: Balanced dataset

- **Step1:** Find the ratio of minority samples to majority samples (R)

$$R = \frac{\text{mins}}{\text{majs}}$$

Where:
mins: minority samples.
majs: majority samples.

- **Step2:** Find number of synthetic data (C)

C= (majs - mins) β
Where:
β: desired R value

- **Step3:** Find number of neighbors that hail from the majority category (Ni)

$$\text{Ni} = \frac{\text{majs}}{k}$$

Where:
k: number of desired KNN

- **Step4:** Normalize Ni

$$NR = \frac{Ni}{\sum Ni}$$

- **Step5:** Find the amount of synthetic data to generate for each neighborhood (Ci)

Ci= C* NR

- **Step6:** Create Ci data for each neighborhood, then create the new synthetic data (Gi)

$Gi=s_i+(s_{zi}-s_i)\,\lambda$

Where:

s_i, s_{zi}: two minority instances within a same neighborhood

λ: Random integer number between 0 and 1.

3.4 Features Selection

The feature selection technique is the process of decreasing the dimension of input variables when designing a model. Reducing the size of the dataset can be performed by removing irrelevant and duplicated features. It plays a substantial role in improving model efficiency, reducing model building costs, reducing overfitting, and speeding up the model learning phase [16]. The correlation between features can be categorized as strong or weak. Features with weak correlation are considered unimportant to the model learning process. Therefore, eliminating these irrelevant features does not reduce the performance of the model. However, features with strong correlation are considered essential to the model learning process and irreplaceable [17]. In this work, feature selection is performed using the analysis of variance (ANOVA), which is a well-known analysis technique utilized in statistics that compares the means of two independent categories. The ANOVA tool calculates the F-ratio test between and within categories [18].

3.5. Stochastic Gradient Boosting

SGB is one of the ensemble algorithms used in classification and regression problems. It is a hybrid of bagging and boosting algorithms. It can handle quantitative and qualitative values; it is robust to outliers and null values. SGB is an improved version of the original gradient boosting algorithm; it combines the advantages of boosting and regression trees [19]. At each new boosting iteration, a random sub-sample is elected from the training data to fit a tree; furthermore, SGB is focused on misclassified training data that are close to their correct classification; finally, at each new iteration, small trees are developed. The advantages of using SGB are: it is easy and convenient to construct; it requires fewer parameters; it is the best algorithm to handle large datasets; and it is faster and more accurate [20, 21].

3.6 Performance Evaluation

Evaluation of performance is a measurement process for how a trained system performs on a particular task. Classification metrics can be used to describe how often the trained model predictions were incorrect or correct [22]. It illustrates the number of correct predictions and incorrect predictions in each class. In this work, a prediction matrix is used that is based on four variables: TP, TN, FP, and FN. This matrix sums up the performance of a model in terms of precision, recall, f1-score, and accuracy. These terms are explained in equations (1-4) [23].

$$PRE = \frac{TP}{TP+FP} \tag{1}$$

$$RC = \frac{TP}{TP+FN} \tag{2}$$

$$FSC = 2 * \frac{RC*PRE}{RC+PRE} \tag{3}$$

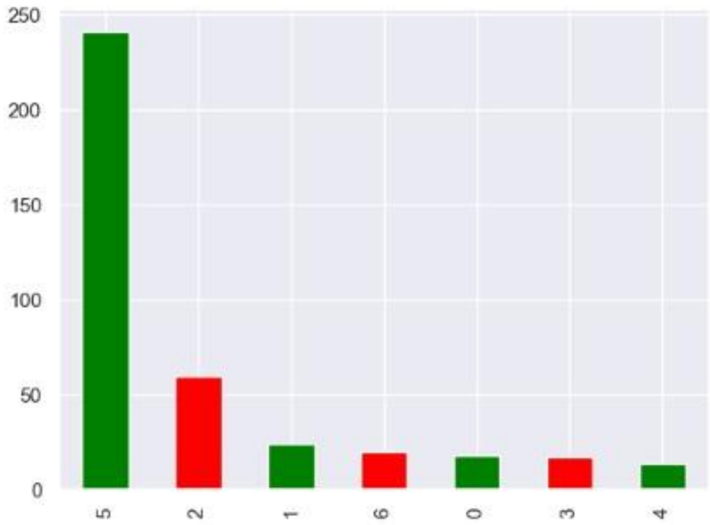
The accuracy of the model can be calculated using the following equation:

$$\text{ACC} = \frac{TP+TN}{TP+TN+FP+FN} * 100 \tag{4}$$

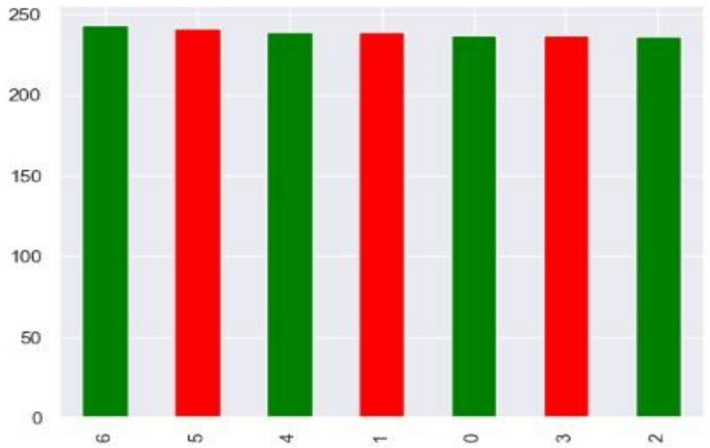
4. Dataset Split and Experimentation

The dataset is split into 70% for training and 30% for testing after pre-processing. Four experiments are conducted, with the first using the initial unbalanced dataset. The discussion about the dataset split and the experiments is related to Figure 1.

In this study, stochastic gradient boosting has been used for migraine classification. Processor Intel (R) Core (TM) i5-2430M CPU @ 2.40GHz, RAM 4GB, x64-based processor, Windows 10 and ANACONDA 3.5.2.0 64-bit are used to implement this work. A significant migraine dataset is employed; it contains 400 records with seven classes of migraine. Following the diagram for the SGB migraine classification model in Fig. (2), the first phase is pre-processing. During the data pre-processing phase, six duplicate records are removed from the migraine dataset. Uniqueness attributes in the dataset that are irrelevant are checked and removed; one unique attribute is removed from the migraine dataset. The dataset classes are unbalanced: 241 patients have class 5 migraine, 60 patients have class 2 migraine, 24 patients have class 1 migraine, 20 patients have class 6 migraine, 18 patients have class 0 migraine, 17 patients have class 3 migraine, and 14 patients have class 4 migraine. In order to balance migraine dataset classes, the adaptive synthetic oversampling method is used, as shown in Fig. 3.



(a) Dataset classes before data balancing



(b) Dataset classes after balancing

Figure 3: Dataset before and after using the ADASYN method to balance data

The next phase is feature selection; the ANOVA method is used to select important features. The last phase is classification, which includes two steps: training and testing. Four experiments have been implemented. The details for the four experiments are illustrated in Table 2 below:

Table 2: The proposed model evaluation in terms of precision, recall, f1-score, and accuracy with and without ANOVA, ADASYN

	ADASYN	ANOVA	Precision	Recall	F1-Score	Accuracy (%)
SGB Model	No	No	0.913842	0.718785	0.785565	86.87%
	Yes	No	0.946789	0.944935	0.945333	94.82%
	No	Yes	0.760652	0.652735	0.680159	83.84%
	yes	yes	**0.970905**	**0.966250**	**0.967278**	**97.01%**

As shown in Table 2, the SGB model achieved outstanding outcomes with accuracy (97.01%) when using the ADASYN oversampling and ANOVA feature selection methods, whereas the results decreased to 83.84% accuracy when using the unbalancing dataset with the feature selection method. Figure 4 below illustrates the results when using or not using the ADASYN oversampling method and the ANOVA feature selection method.

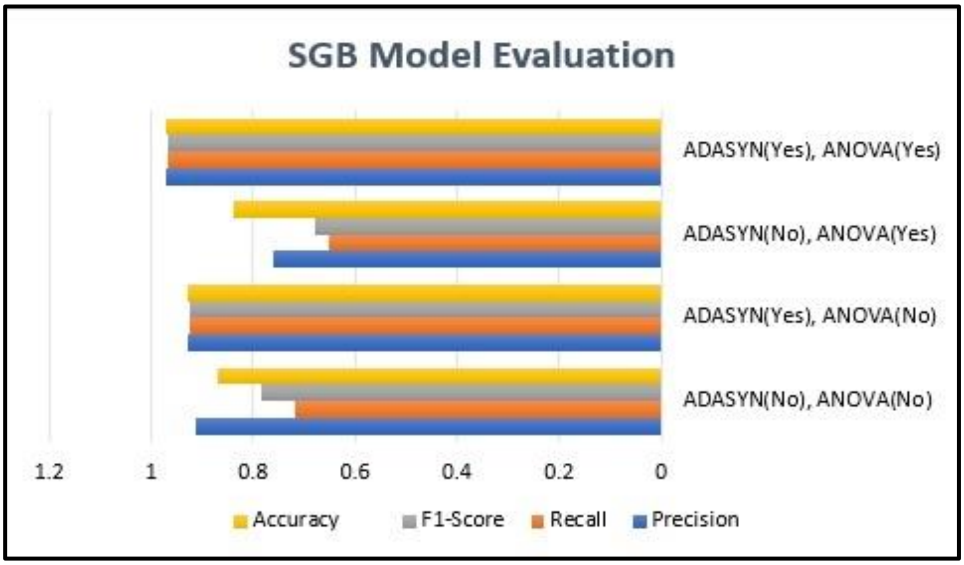


Figure 4: Results when used or not: ADASYN and ANOVA

Figure 5 below illustrates the accuracy of the four experiments carried out on the presented migraine classification model.

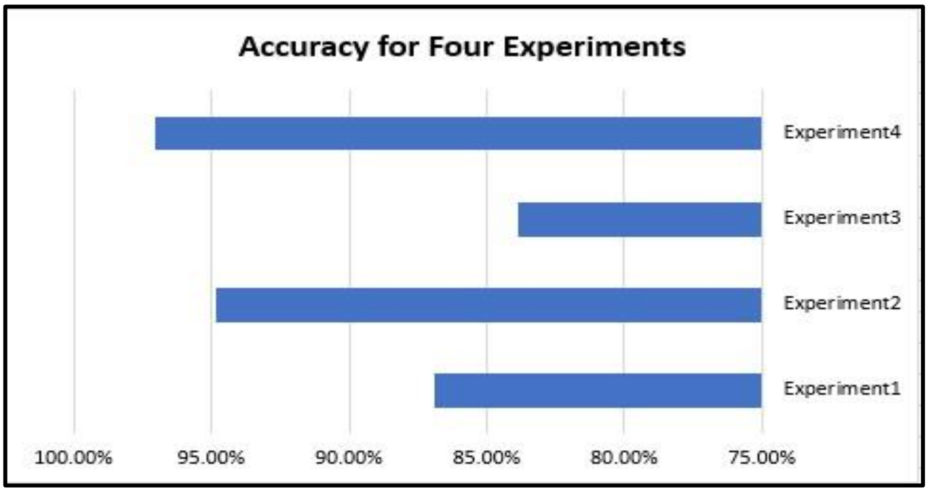


Figure 5: The accuracy of the four experiments carried out on the presented migraine classification model

5. Conclusion

Migraine is one of the most annoying neurological diseases. The aim of this study is to design a migraine classification model capable of accurately classifying seven types of migraine. A stochastic gradient-boosting ensemble machine learning algorithm is used. SGB is combined with ANOVA, which is an influential feature selection method, to provide a methodical solution for relevant feature assessment. The presented migraine classification system is tested on a significant dataset. The migraine dataset is balanced using the ADASYN method. Four experiments are carried out based on the used and not-used ANOVA and ADASYN methods. The proposed work obtained prominent outcomes when evaluated on precision, recall, f1-score, and accuracy. This work achieved high performance compared to prior studies; it achieved 0.970905 precision, 0.966250 recall, 0.967278f1-score, and 97.01% accuracy. These outcomes manifest that SGB has the required potential to be exploited in disease multi-classification. For future works, more samples, a complex dataset, and different feature selection techniques can be used to improve this work and classify more headache types.

Acknowledgments

The author would like to thank Mustansiriyah University (www.uomustansiriyah.edu.iq) in Baghdad, Iraq, for its support in the present work.

References

[1] O. E. Santangelo, V. Gentile, S. Pizzo, D. Giordano, and F. Cedrone, "Machine learning and prediction of infectious diseases: a systematic review," *Mach. Learn. Knowl. Extr.*, vol. 5, no. 1, pp. 175–198, Feb. 2023, Doi: 10.3390/make5010013. [Online]. Available: https://doi.org/10.3390/make5010013.

[2] A. Subasi, A. Ahmed, and E. Luckovich, "Effect of flash stimulation for migraine detection using decision tree classifiers," *Procedia Comput. Sci.*, vol. 140, pp. 223–229, 2018, Available: https://doi.org/10.1016/j.procs.2018.10.332.

[3] A. Mohammed and R. Kora, "A Comprehensive review on ensemble deep learning: opportunities and challenges," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 2, pp.757–774, Feb. 2023, Available: https://doi.org/10.1016/j.jksuci.2023.01.014.

[4] L. Göcs and Z. C. Johanyák, "Feature selection with weighted ensemble ranking for improved classification performance on the CSE-CIC-IDS2018 dataset." *Computers*, vol. 12, no. 8, p. 147, Jul. 2023, Doi: 10.3390/computers12080147. [Online]. Available: https://doi.org/10.3390/computers12080147.

[5] M. T. Vo, T. Nguyen, H. A. Vo, and T. Le, "Noise-adaptive synthetic oversampling technique," *Appl. Intell.*, vol. 51, no. 11, pp. 7827–7836, Mar. 2021, Available: https://doi.org/10.1007/s10489-021-02341-2.

[6] N. Mduma, "Data balancing techniques for predicting student dropout using machine learning," *Data*, vol. 8, no. 3, p. 49, Feb. 2023, Doi: 10.3390/data8030049. [Online]. Available: http://dx.doi.org/10.3390/data8030049.

[7] Y. Garcia-Chimeno, B. Garcia-Zapirain, M. Gomez-Beldarrain, B. Fernandez-Ruanova, and J. C. Garcia-Monco, "Automatic migraine classification via feature selection committee and machine learning techniques over imaging and questionnaire data," *BMC Med. Inform. Decis. Mak.*, vol. 17, no. 1, p.38, Apr. 2017, Available: Doi: 10.1186/s12911-017-0434-4.

[8] B. Zhu, G. Coppola, and M. Shoaran, "Migraine classification using somatosensory evoked potentials," *Cephalalgia*, vol. 39, no. 9, pp. 1143–1155, Mar. 2019, Available: Doi: 10.1177/0333102419839975.

[9] A. Qawasmeh, N. Alhusan, F. Hanandeh, and M. Al-Atiyat, "A High-performance system for the diagnosis of headache via hybrid machine learning model," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 5, pp. 655-663, Jan. 2020, doi:10.14569/IJACSA.2020.0110580.

[10] S. Simić, J. R. Villar, J. L. Calvo-Rolle, S. R. Sekulić, S. D. Simić, and D. Simić, "An application of a hybrid intelligent system for diagnosing primary headaches," *Int. J. Environ. Res. Public Health*, vol. 18, no. 4, p.1890, Feb. 2021, Available: Doi: 10.3390/ijerph18041890.

[11] A.U.R. Butt, "Data for Migraine Classification," Kaggle, 2022, Available: https://www.kaggle.com/code/azazurrehmanbutt/migraine-prediction-dnns-99-acc/input.

[12] E. Cho, T.-W. Chang, and G. Hwang, "Data preprocessing combination to improve the performance of quality classification in the manufacturing process," *Electronics*, vol. 11, no. 3, p. 477, Feb. 2022, Available: Doi: 10.3390/electronics11030477.

[13] T. S. Amelia, M. N. S. Hasibuan, and R. Pane, "Comparative analysis of resampling techniques on machine learning algorithm," *Sinkron*, vol. 7, no. 2, pp. 628–634, Apr. 2022, Available: Doi: 10.33395/sinkron.v7i2.11427.

[14] L. M. Shoohi, J. H. Saud, "Adaptation proposed methods for handling imbalanced datasets based on over-sampling technique," *MJS*, vol. 31, no. 2, p. 25, Apr. 2020, Available: Doi: https://doi.org /10.23851/mjs.v31i2.740.

[15] W. Falah and I. J. Mohammed, "Hybrid CNN-SMOTE-BGMM Deep Learning Framework for Network Intrusion Detection using Unbalanced Dataset," *Iraqi J. Sci.*, vol. 64, no. 9, pp. 4846–4864, Sep. 2023, Doi: 10.24996/ijs.2023.64.9.43.

[16] N. Ansari, "A Survey on Feature Selection Techniques using Evolutionary Algorithms," *Iraqi J. Sci.*, vol. 62, no. 8, pp. 2796–2812, Aug. 2021, Doi: 10.24996/ijs.2021.62.8.32.

[17] P. Usha and M. P. Anuradha, "Feature selection techniques in learning algorithms to predict truthful data," *Indian J. Sci. Technol.*, vol. 16, no. 10, pp. 744–755, Mar. 2023, Available: Doi: 10.17485/IJST/v16i10.2102.

[18] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Electr. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014, Available: https://doi.org/10.1016/j.compeleceng.2013.11.024.

[19] S. Panicker, S. Selot, and M. Sharma, "Improving Accuracy in Human Age Classification Using Ensemble Learning Techniques," *Iraqi J. Sci.*, vol. 60, no. 8, pp. 1830–1836, Aug. 2019, doi:10.24996/ijs.2019.60.8.20.

[20] S. Cerón-Figueroa, C. López-Martín, and C. Yáñez-Márquez, "Stochastic gradient boosting for predicting the maintenance effort of software-intensive systems," *IET Softw.*, vol. 14, no. 2, pp. 82–87, Apr. 2020, Available: https://doi.org/10.1049/iet-sen.2018.5332.

[21] M. R. Baker, A.H. Alamoodi, O.S. Albahri, A.S. Albahri, S. Garfan, A. Alamleh, M. L. Shuwandy, and I. Alshakhatreh, "Comparison of Machine Learning Approaches for Detecting COVID-19-Lockdown-Related Discussions During Recovery and Lockdown Periods", *J. Oper. Intell*., vol.1, no.1, pp.11-29, Oct. 2023, doi.org/10.31181/jopi1120233.

[22] I.M. Hassoon and S.A. Hantoosh, "CFNN for Identifying Poisonous Plants," *Baghdad Sci. J.*, vol. 20, no. 3(Suppl.), pp. 1122–1130, Jun. 2023, https://doi.org/10.21123/bsj.2023.7874.

[23] D. Božić, B. Runje, D. Lisjak, and D. Kolar, "Metrics related to confusion matrix as tools for conformity assessment decisions," *Appl. Sci.*, vol. 13, no. 14, p. 8187, Jul. 2023, Doi: 10.3390/app13148187. [Online]. Available: https://doi.org/10.3390/app13148187.