Enhanced Classification of Migraine Headaches: Leveraging Ensemble Models and Boosting Techniques with Data Augmentation

Introduction :

Migraine disease is a neurological condition characterized by recurrent, pulsating headaches that can range from moderate to severe in intensity. These headaches are often accompanied by symptoms such as nausea, sensitivity to light and sound, and in some cases, visual disturbances known as aura. Migraines can last anywhere from a few hours to several days, significantly impacting daily activities and quality of life for millions worldwide.

The exact cause of migraines is multifactorial and not fully understood, but it involves a combination of genetic predisposition and environmental triggers. Common triggers include stress, hormonal fluctuations, certain foods (like aged cheese and processed meats), lack of sleep, dehydration, and environmental factors such as strong Odors or weather changes. A study [1] revealed that more than 90% of the population suffers from headaches.

There are several types of migraines, including migraine with aura (where sensory disturbances precede the headache), migraine without aura, hemiplegic migraine (accompanied by motor weakness), and chronic migraine (occurring on 15 or more days per month). Managing migraine involves identifying triggers through diary-keeping, lifestyle modifications (such as regular sleep patterns and stress management), and sometimes medication prescribed by healthcare providers.

The diverse clinical presentations of migraine include:

- **Migraine without Aura**: Commonly presents as a headache without preceding aura symptoms.
- **Migraine with Typical Aura**: Involves aura symptoms occurring before or during the headache, such as visual disturbances (flashes of light, blind spots), sensory changes (tingling, numbness), or motor impairments.
- **Typical Aura without Migraine**: Rarely, individuals may experience aura symptoms without subsequent headache.
- **Familial Hemiplegic Migraine**: Genetic condition characterized by aura symptoms and temporary weakness or paralysis on one side of the body.
- **Sporadic Hemiplegic Migraine**: Similar to familial hemiplegic migraine but occurs sporadically without a clear genetic link.

- **Basilar-type Migraine**: Involves aura symptoms originating from the brainstem, often leading to neurological symptoms such as dizziness, double vision, and difficulty speaking.

In this study, we employed advanced machine learning techniques to analyse data related to migraine episodes. Methods utilized include Support Vector Machines (SVM), Random Forest, k-Nearest Neighbors (KNN), Decision Trees (DT), Deep Neural Networks (DNN), and ensemble methods. Among these, the Deep Neural Network (DNN) demonstrated the highest accuracy in predicting migraine episodes. By leveraging data augmentation techniques, the DNN model effectively captured complex patterns within the migraine dataset, offering potential insights for improved treatment strategies and personalized management plans.

Literature Survey :

Research paper	Authors	Research gap	Model	Accuarcy
Migraine (headaches) disease data classification using data mining classifiers. [2]	Sah, R. D., Sheetlani	Lack of exploration into advanced machine learning techniques and the integration of diverse data types to improve the accuracy and robustness of migraine classification.	SVM, RF, KNN, Navie Bayers	90.5% - 95.3%
A decision support system for primary headache developed through machine learning. [3]	Liu, F., Bao	Limited application of real-world clinical data to validate the machine learning-based decision support system for primary headache diagnosis.	DST, RF, LR, Gradient Boosting	91% - 93%
Automatic diagnosis of primary headaches by machine learning methods. [4]	Krawczyk, B., Simić	insufficient investigation of ensemble and boosting techniques to enhance the accuracy and reliability of primary headache diagnosis using machine learning methods.	Ada boost, Ensemble Bagging	77%, 80%
Automatic migraine classification using artificial neural networks. [5]	Sanchez, Rúa Ascar	broader evaluation of artificial neural networks with larger, more diverse datasets to improve the generalizability and accuracy of automatic migraine classification.	Logistic Regression	87.5%, 92.15%
Classification of migraine disease using supervised machine learning. [6]	Gulati, S., Guleria	Limited exploration of hybrid and deep learning models to further enhance the classification accuracy of migraine disease using supervised machine learning.	Naive Bayes	94%

Automatic migraine classification via feature selection committee and machine learning techniques over imaging and questionnaire data. [7]	Garcia-Chimeno, Y., Garcia-Zapirain	Inadequate assessment of the integration of feature selection methods with advanced machine learning algorithms to optimize migraine classification using multimodal data.	SVM, ada boost, navie Bayers	90% - 98%
Machine learning-based automated classification of headache disorders using patient-reported questionnaires. [8]	Kwon, J. et al	Further validation of machine learning models for headache classification across diverse populations and integration with clinical assessments is needed.	XGBoost	82%
Data augmentation and deep neural networks for the classification of Pakistani racial speakers recognition. [9]	Ammar Amjad, Lal Khan	Exploration of the generalizability of data augmentation techniques and deep neural networks for racial speaker recognition.	DNN(with smote)	95%
Classification of multi-channel EEG signals for migraine detection. [10]	Akben SB, Tuncel D	Generalizability of multi-channel EEG signal classification models for migraine detection across different patient demographics and conditions.	SVM, CNN	-
A clinical decision support system for the diagnosis of probable migraine and probable tension-type headache based on case based reasoning. [11]	Yin Z, Dong Z	There is a need for more effective and accurate CDSS tools specifically designed for diagnosing migraine and tension-type headaches. Existing systems may lack precision or comprehensiveness.	Case-Based Reasoning (CBR)	79.3%

Research Gaps :

The research on using machine learning for migraine classification reveals several critical gaps that must be addressed. The availability of publicly accessible datasets is notably limited, restricting the ability to develop and validate more precise models. There is a need for practical frameworks and tools to integrate machine learning models into healthcare settings, particularly in underdeveloped regions. Most studies to date are cross-sectional, highlighting the need for longitudinal research to evaluate model performance over time and their adaptability to evolving conditions in migraine patients. Additionally, the integration of structured and unstructured data sources has been insufficiently explored, calling for further research to develop methods that effectively combine diverse data types for better diagnostic accuracy.

Proposed Solutions:

Ensemble Models –

- Bagging method
- Stacking method
- Voting method

Boosting Techniques –

- Ada boosting
- Gradient boosting
- Gradient boosting Decision Tree

Identified Solution :

The use of ensemble models and boosting techniques significantly improved the accuracy of migraine classification in the study. By implementing models such as Ensemble Models, and Boosting Techniques, and leveraging data augmentation techniques, the performance of the classifiers was notably enhanced. These advanced methods outperformed traditional machine learning models, achieving higher accuracy and robustness in predicting migraine types. Specifically, the deep neural network (DNN) model, combined with data augmentation, reached an accuracy of 99.66%, demonstrating the effectiveness of using ensemble and boosting models for better diagnostic accuracy in migraine classification.

Materials and Methods:

Algorithms Proposed :

In the study of migraine classification, several advanced ensemble and boosting methods were employed to enhance predictive accuracy and model robustness. These techniques leverage the strengths of multiple models to improve the overall performance and reliability of the classification system.

Ensemble Models :

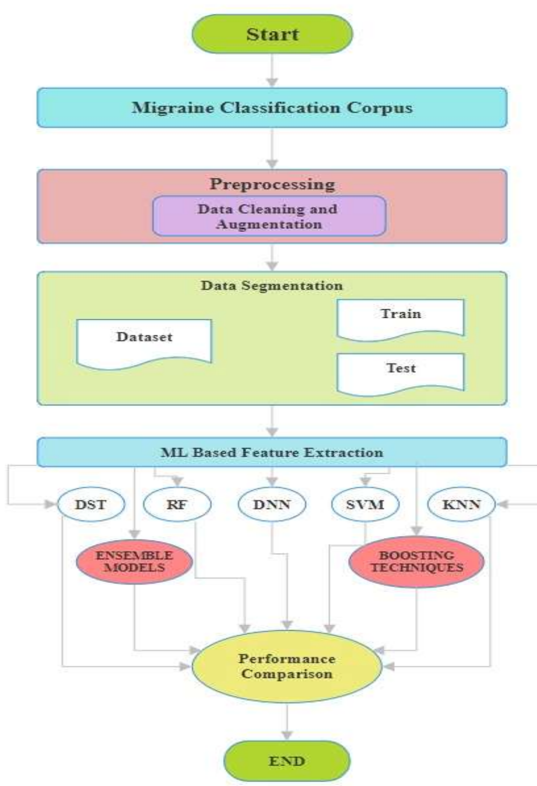
- **Bagging Method**: Bagging, or Bootstrap Aggregating, involves training multiple models on different subsets of the data and averaging their predictions. This method reduces variance and helps prevent overfitting, enhancing model stability and accuracy.

- **Stacking Method**: Stacking is an ensemble technique that combines multiple models by training a meta-learner to combine the base models' predictions. It leverages the strengths of various models to improve predictive performance.
- **Voting Method**: The Voting method combines the predictions of several base models. The final prediction is determined by the majority vote (for classification) or average (for regression). It helps improve robustness and generalization.

Boosting Models :

- **AdaBoost**: AdaBoost, or Adaptive Boosting, combines multiple weak classifiers to create a strong classifier. It works by sequentially training classifiers, where each subsequent classifier focuses on the errors made by the previous ones. This method increases model accuracy by reducing bias and variance.
- **Gradient Boosting**: Gradient Boosting builds models sequentially, with each new model correcting the errors made by the previous ones. It uses gradient descent to minimize the loss function, enhancing the model's predictive performance.
- **Gradient Boosting Decision Tree (GBDT)**: GBDT is a specific implementation of gradient boosting that uses decision trees as the base learners. It combines the strengths of decision trees and gradient boosting to produce a powerful predictive model with high accuracy and robustness.

Flow Chart :



Results and Discussion :

Description of the Proposed Datasets :

Name of the Dataset	Dataset for Migraine/Headache Classifications
Number of Samples Recorded	400
Source	Hospital Materno Infantil de Soledad, Columbia
Number of Features	24
Number of Class Labels	7
Number of Sample after Data Augmentation (On this paper)	1449

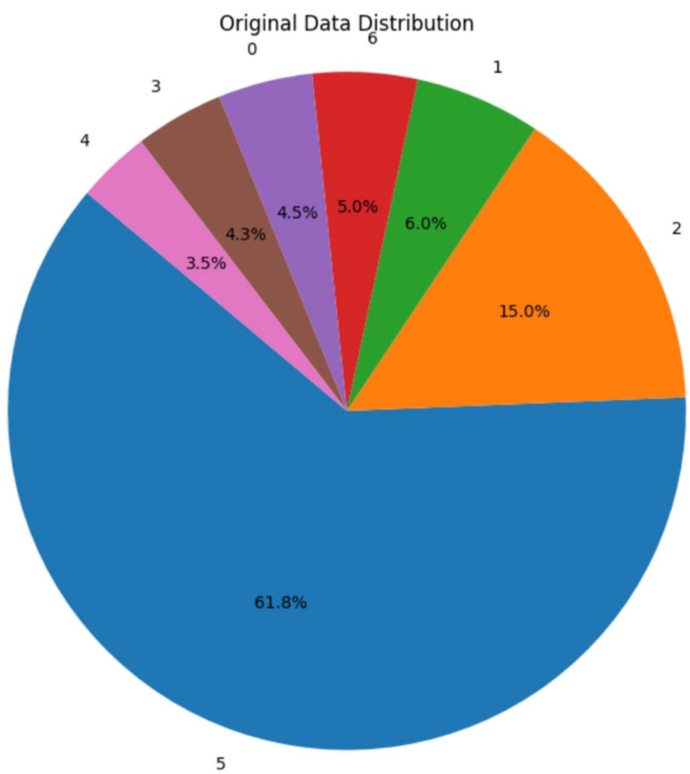


Figure 1 : Dataset Class Imbalance before Augmentation

In this study, we enhanced the dataset by increasing the number of instances from 400 to 1449. This augmentation ensured a balanced distribution with 207 instances for each class label. The data augmentation Technique is SMOTE [13].

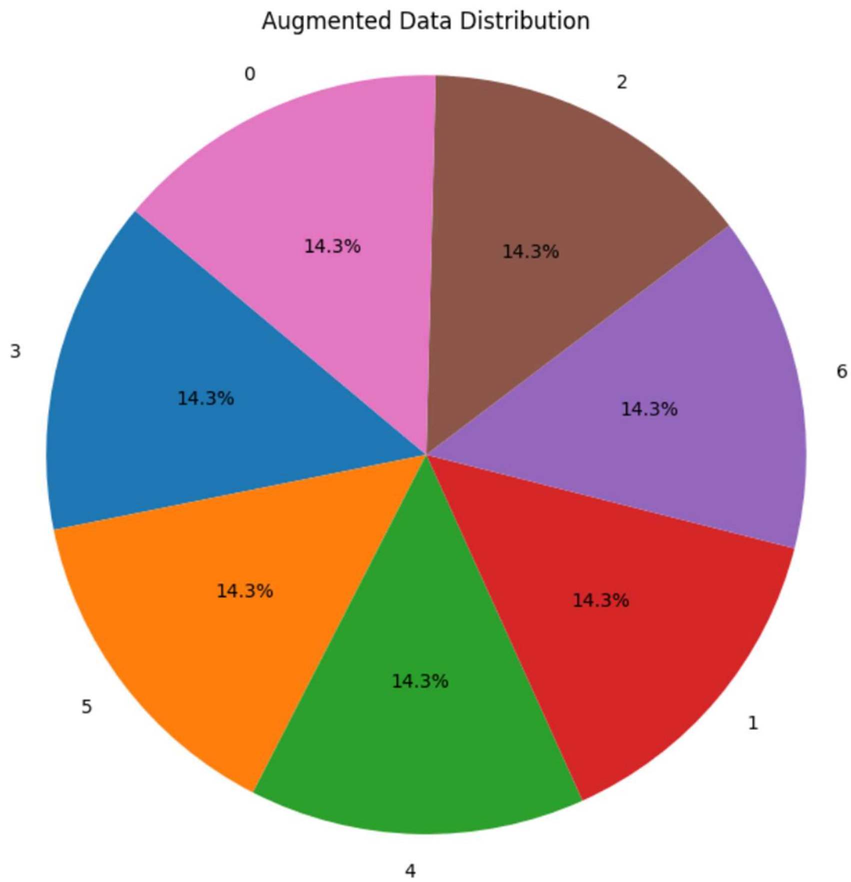


Figure 2 : Dataset Class Balance after Augmentation

Details of the Hardware and Softwares used :

Device Processor : 11th Gen Intel(R) Core(TM) i5-11320H @ 3.20GHz 3.19 GHz

Device Ram : 16.0 GB for RAM

Device System Type : 64-bit operating system, x64-based processor

Software used : Google Colab

Software Processor : Intel Xeon CPU with 2 vCPUs (virtual CPUs)

Software Ram : 13 GB of RAM

Literature results interms of the evaluation metrics :

Table 1 : Classification Report of all used algorithms with and without Augmentation :

Model	Precision	Recall	F1score	Accuracy(%)
SVM	99.38	98.75	98.75	98.75
SVM + Data Augmentation	99.33	99.31	99.31	99.31
KNN	96.22	95	94.77	95
KNN + Data Augmentation	99.66	99.65	99.65	99.65
DST	96.5	96.25	95.98	96.25
DST + Data Augmentation	99.67	99.66	99.66	99.66
RF	96.33	96.25	96.18	96.25
RF + Data Augmentation	99.01	98.97	98.96	98.97
DNN	92.5	92.5	92.4	92.5
DNN + Data Augmentation	99.65	99.66	99.65	99.66

Table 2 : Classification Report of Ensemble Models with and without Augmentation :

Model	Precision	Recall	F1score	Accuracy(%)
Stacking	96.4	96.25	95.8	96.25
Stacking + Data Augmentation	98.32	98.28	98.28	98.28
Bagging	98.77	98.75	98.7	98.75
Bagging + Data Augmentation	98.69	98.62	98.6	98.62
Voting	96.92	96.25	96.16	96.25
Voting + Data Augmentation	99.66	99.66	99.65	99.66

Table 3 : Classification Report of Boosting Techniques with and without Augmentation :

Model	Precision	Recall	F1score	Accuracy (%)
Ada boost	93.5	95	93.86	95
Ada boost + Data Augmentation	96.89	96.89	96.85	96.89
Gradient Boosting Classifier	97.94	97.5	97.42	97.5
Gradient Boosting Classifier + DA	99.67	99.66	99.66	99.66
Gradient Boosting Decision Tree	97.94	97.5	97.42	97.5
Gradient Boosting Decision Tree + DA	99.67	99.66	99.66	99.66

Dataset Split up :

We have split the data into 80-20 ratio. 80% is used for training and remaining 20 % is used for testing.

Comparision of the proposed models performance with the state of the art results interms of tables and figures :

Study	**Models used**	**Accuracy(%)**
[2]	SVM	95.3%
[3]	Gradient Boosting	93%
[4]	Ensemble Bagging	80%
[5]	Logistic Regression	92.15%
[6]	Navie Bayers	94%
[7]	Ada Boost	94%
[8]	XG Boost	84%
[11]	Case-Based Reasoning(CBS)	79.3%
[12]	DNN(With SMOTE)	99.66%
Proposed Study	Stacking + Data Augmentation	98.28%
Proposed Study	Bagging + Data Augmentation	98.62%
Proposed Study	Voting + Data Augmentation	99.66%
Proposed Study	Ada boost + Data Augmentation	96.89%
Proposed Study	Gradient Boosting Classifier + DA	99.66%
Proposed Study	Gradient Boosting Decision Tree + DA	99.66%

Observations :

The data provides intriguing insights into the impact of data augmentation on various classifiers. Initially, the Support Vector Machine (SVM) demonstrates robust performance, maintaining high accuracy even with data augmentation, suggesting its strong generalization capabilities. Conversely, the K-Nearest Neighbors (KNN) and Decision Tree (DST) classifiers, which initially exhibited moderate accuracies, experienced significant improvements with augmented data, indicating their sensitivity to data variability and enhanced generalization. Random Forest (RF), while consistent initially, also benefited notably from data augmentation, showcasing its adaptability to augmented datasets for improved classification. These observations collectively highlight the transformative potential of data augmentation, particularly for models like KNN, DST, and RF, in enhancing classification accuracy and robustness by leveraging augmented data for training.

DNN achieved notably high accuracy, it suggests that they may have outperformed other classifiers in terms of predictive power and generalization. However, it's important to note that the effectiveness of DNN models can depend on various factors such as network architecture, data quality, and tuning parameters.

The data showcases the performance of ensemble learning techniques, including Stacking, Bagging, and Voting, with and without data augmentation. Stacking initially demonstrates respectable accuracy levels, with a notable improvement when augmented data is employed. Bagging exhibits consistently strong performance without augmentation, showing minimal impact from augmentation. Similarly, Voting initially shows decent accuracy, with a significant boost when augmented data is utilized. These observations underscore the varying impacts of data augmentation across ensemble learning techniques, with some experiencing substantial improvement, while others maintain strong performance even without augmentation. The highest achieved accuracies are approximately 98.32% for Stacking with data augmentation and around 99.65% to **99.66% for Voting with data augmentation**.

In the realm of classifier performance, AdaBoost initially demonstrates accuracy levels ranging from 93.5% to 95%, which significantly improves to approximately 96.89% with the integration of augmented data. However, the Gradient Boosting Classifier and Gradient Boosting Decision Tree, although performing strongly without data augmentation, achieve even higher accuracies of approximately 99.66% to 99.67% when augmented data is utilized. This remarkable performance enhancement positions the Gradient Boosting Classifier and **Gradient Boosting Decision Tree as the highest-achieving classifiers in the dataset with 99.66%**

Conclusion :

The study delved into advanced machine learning techniques, specifically ensemble models and boosting methods, to classify migraine headaches, addressing gaps in existing literature such as limited integration of diverse data types and the need for robust classification models. Ensemble methods like Bagging, Stacking, and Voting, alongside Boosting techniques such as AdaBoost and Gradient Boosting, were applied to a dataset enhanced using SMOTE, resulting in significantly improved classification accuracy. Notably, Voting and Gradient Boosting Classifier achieved the highest accuracy at 99.66% when utilizing augmented data.

The research emphasized the pivotal role of data augmentation in enhancing the generalizability and robustness of machine learning models, particularly benefiting classifiers sensitive to variations in data distribution. By balancing class labels and optimizing model performance across various evaluation metrics, this approach demonstrated substantial improvements over traditional methods.

The findings underscored the potential of ensemble and boosting techniques in advancing migraine classification accuracy, promising more effective diagnostic tools and personalized treatment strategies. Future avenues could explore longitudinal studies and the integration of structured and unstructured data sources to further bolster model reliability in clinical settings.

Overall, this study showcases how leveraging ensemble and boosting methods alongside data augmentation can not only refine migraine classification but also contribute to broader applications in healthcare, where precise and reliable predictive models are crucial for improving patient outcomes and treatment efficacy.

References :

1. Hagen, K. et al. The epidemiology of headache disorders: A face-to-face interview of participants in hunt4. J. Headache Pain 19, 1–6 (2018).
2. Sah, R. D., Sheetlani, J., Kumar, D. R. & Sahu, I. N. Migraine (headaches) disease data classification using data mining classifiers. J. Res. Env. Earth Sci. 3, 10–16 (2017).
3. Liu, F., Bao, G., Yan, M. & Lin, G. A decision support system for primary headache developed through machine learning. PeerJ 10, e12743 (2022).
4. Krawczyk, B., Simić, D., Simić, S. & Woźniak, M. Automatic diagnosis of primary headaches by machine learning methods. Open Med. 8, 157–165 (2013).
5. Sanchez-Sanchez, P. A., García-González, J. R. & Rúa Ascar, J. M. Automatic migraine classification using artificial neural networks. F1000Research 9, 618 (2020).
6. Gulati, S., Guleria, K. & Goyal, N. Classification of migraine disease using supervised machine learning. In '2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO), 1–7 (organizationIEEE, 2022).
7. Garcia-Chimeno, Y., Garcia-Zapirain, B., Gomez-Beldarrain, M., Fernandez-Ruanova, B. & Garcia-Monco, J. C. Automatic migraine classification via feature selection committee and machine learning techniques over imaging and questionnaire data. BMC Med. Inform. Decis. Mak. 17, 1–10 (2017).
8. Kwon, J. et al. Machine learning-based automated classification of headache disorders using patient-reported questionnaires. Sci. Rep. 10, 14062 (2020).
9. Ammar Amjad, Lal Khan and Hsien-Tsung Chang : Data augmentation and deep neural networks for the classification of Pakistani racial speakers recognition, PeerJ Comput. Sci., DOI 10.7717 (2022).
10. Akben SB, Tuncel D, Alkan A: Classification of multi-channel EEG signals for migraine detection. Biomed Res. 2016; 27(3): 743–748.
11. Yin Z, Dong Z, Lu X, et al.: A clinical decision support system for the diagnosis of probable migraine and probable tension-type headache based on case based reasoning. J Headache Pain. 2015; 16(29): 1–9
12. Khan, Shahreen, Qazi, Shah, Hussain & Chang : Migraine headache (MH) classification using machine learning methods with data augmentation, Scientific report (2024).
13. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. Smote: Synthetic minority over-sampling technique. J. Art. Intell. Res. 16, 321–357 (2002).