International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2s, 2025

Electric Vehicle Charge Classification Technique For Optimized Battery Charge Based On Ml

Mohammed Abdul Mannan[1], Ibrahim Siddiqui[2], Bilal Rasheed[3], Arjumand Jamal[4]

[1,2,3]UG Scholar, Dept of IT ISL Engineering College, Hyderabad, India.

[4]Assistant Professor, Department Of IT, ISL Engineering college (Autonomous), India.

mdabdulmannan5256@gmail.com, ahmedbilal5941@gmail.com, ibrahimsiddiqui5426@gmail.com

ABSTRACT

The severity of ischaemic stroke ,the world Intricate connection in the 5110 patient health profiles Leading cause of death and disability ,can be that made up the health care dataset. Three potent Considerably decreased by early detection of stoke . Classifiers light gradient boosting machine(LGBM), The purpose of this research is to use cutting edge extreme gradient boosting(XGBoost),random forest Machine learning techniques to create a predictive (RF)_were integrated to create a stacking ensemble Model for ischaemic strock. This method u.se an model that improved the prediction capacity. Artificial neutral network(ANN).

1. INTRODUCTION

One of the main causes of death and permanent impairment in the world is ischaemic stroke, which is defined by an obstruction in the blood arteries supplying the brain. In order to lessen its terrible effects, early warning signals and risk factors must be promptly identified. Because ischaemic stroke is a complicated and multifaceted disorder, reliable prediction of the condition remains a substantial issue despite breakthroughs in medical diagnostics. By facilitating data-driven insights and predictive analytics in the healthcare industry, the development of artificial intelligence (AI) and machine learning (ML) presents a revolutionary solution to this problem. Using the advantages of both feature extraction and ensemble learning, this work presents a unique framework for ischaemic stroke prediction. This study attempts to improve the precision and dependability of ischaemic stroke predictions by utilising an artificial neural network (ANN) as a feature extractor and integrating the predictive capabilities of sophisticated machine learning classifiers—Light Gradient Boosting Machine (LGBM), Extreme Gradient Boosting (XGBoost), and Random Forest (RF). Furthermore, robust learning from under-represented classes in imbalanced datasets is ensured by the application of the synthetic minority oversampling technique (SMOTE). The suggested approach facilitates its implementation in clinical settings by offering interpretability in addition to state-of-the-art performance.

2 OBJECTIVE

Create an Advanced Predictive Model: To attain better performance metrics, build a stacking ensemble model that combines ANN, LGBM, XGBoost, and RF.For class balancing, use SMOTE: To address the class imbalance in the healthcare dataset, use data augmentation approaches. Analyse Model Performance: Use important performance indicators including accuracy, precision, recall, F1-score, AUC, and confusion matrix to evaluate the predictive model. Encourage clinical usability by making sure that medical professionals can understand and act upon the model's results.

2.1 PROBLEM STATEMENT

Ischaemic stroke is a leading global cause of death and disability, with early detection being critical to reducing its severity and improving patient outcomes. However, identifying stroke warning signs early is challenging due to the complexity and variability of patient health profiles, compounded by class imbalances in medical datasets that can hinder predictive modelling. Existing approaches may lack the accuracy, interpretability, or robustness needed for effective clinical decision support in resource-constrained settings. Therefore, there is a need to develop a reliable, interpretable, and high-performing predictive model using advanced machine learning techniques to accurately identify ischaemic stroke risk from patient health data, enabling timely intervention and personalized decision-making.

2.2 EXISTING SYSTEM

Current approaches frequently have trouble with unbalanced datasets, which results in subpar minority case prediction. Numerous conventional models, such as CNN, LSTM, and logistic regression, are unable to handle complicated data patterns and do not attain high accuracy. Furthermore, deep learning algorithms may function as "black boxes," making it challenging for medical professionals to interpret the predictions. Another restriction that affects prediction accuracy is the inability to extract significant information and include patient history. Furthermore, a large number of current algorithms are unable to produce individualised forecasts based on unique patient profiles and generalise effectively across a variety of populations.

Disadvantage of Existing System

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP146-151

- Struggle to achieve high accuracy due to their inability to capture complex patterns in the data effectively
- Fail to handle imbalanced datasets, resulting in poor prediction performance for minority classes, such as high-risk stroke cases.

2.3 PROPOSED SYSTEM

To obtain high accuracy and dependable predictions, the suggested model for stroke prediction incorporates the advantages of several machine learning methods. Utilising its capacity to recognise intricate patterns in the input data, an Artificial Neural Network (ANN) is employed as a feature extractor. Three potent models—Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LGBM)—are then fed the features that the ANN has extracted. XGBoost improves speed by gradient-based optimisation and effectively manages errors, while LGBM effectively manages huge datasets by iteratively improving predictions with less processing. In contrast, Random Forest reduces over fitting by building numerous decision trees and mixing their results to ensure robustness. A stacking ensemble strategy is used to integrate these models, combining each model's specific predictions before sending them to a meta-classifier for a final determination.

Advantages of Proposed System

- Its unique strengths, reducing prediction errors.
- Feature extraction ensures that complex and non-linear patterns in the data are captured effectively.
- Feature extraction ensures that complex and non-linear patterns in the data are captured effectively.
- Model is capable of handling large datasets efficiently, owing to the computational optimizations

2. RELATED WORKS

This section examines recently released research on the prediction of ischemic stroke from 2017 onwards from a variety of online sources. In this investigation, the deployed dataset, data prior to A. A REVIEW OF EARLY ISCHEMIC STROKE PREDICTION METHODS To date, a large number of regression-based or other statistically based ischemic stroke prediction models have been developed. However, the practical utility of these models by [9] can occasionally be substantially limited due to their limited incorporation of minor components. According to a study by [10], multivariable logistic models had an area under the receiver operating characteristic (AUROC) curve of 0.71–0.74, indicating good performance. These models were created based on 332 patients' clinical and retinal characteristics (20 variables). Large volumes of real-world patient-level data from EHRs combined with ML and DL techniques can enhance the number of attributes recorded, enabling the development of prediction models that are more accurate [11]. B. CURRENTLY APPLIED ISCHEMIC STROKE PREDICTION METHODS The "Cardiovascular Health Study" (CHS) records were collected in 2017 [12]. There were 212 strokes and non-strokes in the three datasets that were produced. There are 212 stroke occurrences among the 1,824 entities and 357 attributes in the finished collection. The proposed approach uses principal component analysis (PCA) for dimension reduction and C4.5 decision tree approaches for feature selection. After the reduction, an Artificial Neural Network (ANN) was used to create a classification model, which produced a 94.7% accuracy classification model.

3. METHODOLOGY OF PROJECT MODULE DESCRIPTION:

1. Data Collection: This module involves gathering diverse patient health profiles, including demographic information (age, gender), medical history (hypertension, diabetes), lifestyle factors (smoking, physical activity), and clinical data (blood pressure, cholesterol levels). Data is collected from trusted sources like hospitals, research institutions, or publicly available healthcare datasets. Ensuring the accuracy and completeness of data is critical to building a reliable predictive model.

2. Data Analysis: Data analysis involves exploring the collected dataset to identify patterns, trends, and correlations between features. Statistical methods and visualization tools such as histograms, box plots, and scatter plots are employed to understand data distribution and highlight potential outliers or inconsistencies. Insights from this step guide further pre-processing and feature engineering tasks.

3. Data Preparation: Pre-processing the raw data to make it suitable for machine learning. This includes: Handling missing values through imputation techniques. Normalizing or scaling continuous variables to bring them to a similar range. Encoding categorical variables using techniques like one-hot encoding or label encoding. Addressing data imbalance using methods like SMOTE to augment underrepresented classes.

4. Feature Extraction: Feature extraction focuses on transforming raw data into meaningful representations that capture the essence of the dataset. Using an ANN as a feature extractor, the model identifies non-linear relationships and latent patterns that improve classification performance. This step ensures that critical predictive information is retained.

5. Feature Selection: Selecting the most relevant features to improve model efficiency and a curacy.

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP146-151

ISSN 2347–3657

Volume 13, Issue 2s, 2025

Techniques such as correlation analysis, mutual information, or feature importance rankings from algorithms like Random Forest are used. This reduces computational overhead and minimizes the risk of over fitting.

6. Data Dividing: Splitting the dataset into training, validation, and testing subsets. Typically, 70-80% of the data is allocated for training, 10-15% for validation, and the remaining for testing. This division ensures the model is trained on one set, validated for hyper parameter tuning, and tested for performance evaluation on unseen data.

7. Applying Algorithm: This module involves implementing the proposed stacking ensemble model. The ANN is used for feature extraction, and the outputs are fed into three classifiers—LGBM, XGBoost, and RF. These classifiers are combined using a meta-classifier in the stacking ensemble to make final predictions, leveraging their individual strengths.

8. Model Evaluation: The model is evaluated using performance metrics like: Accuracy: Overall correctness of the model. Precision: Proportion of true positive predictions among all positive predictions. Recall: Ability of the model to identify all relevant instances.F1-Score: Harmonic mean of precision and recall. AUC: Measures the ability of the model to distinguish between classes. Confusion Matrix: Provides a detailed breakdown of predictions into true positives, true negatives, false positives, and false negatives.

9. Model Deployment: Deploying the trained model into a production environment where it can be accessed by healthcare professionals. This involves creating APIs or integrating the model into existing healthcare systems, ensuring security and scalability**.**

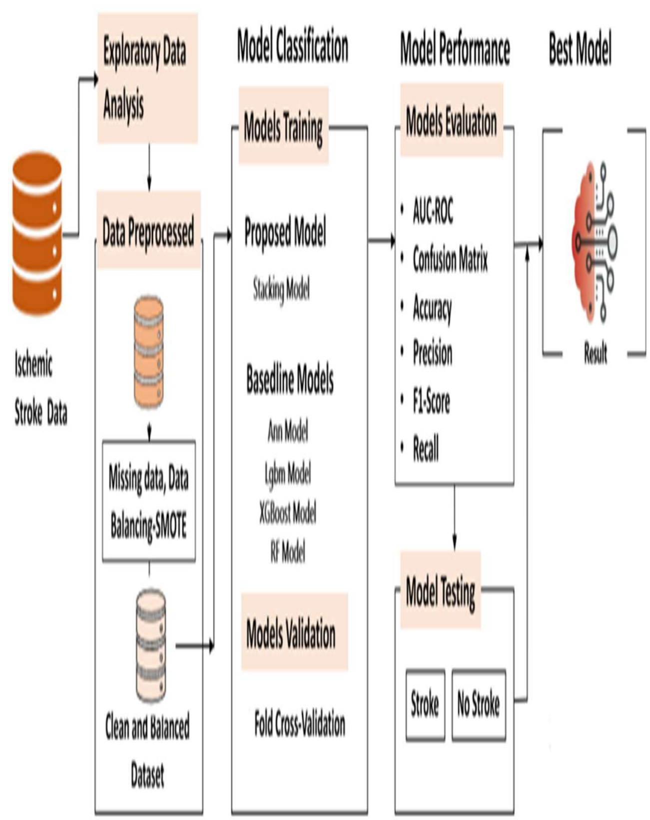
10. User Interface: Developing an intuitive and user-friendly interface that allows users to input patient data, view predictions, and access interpretability features. The interface may include visualizations like risk scores, feature importance, and recommendations for further action.

11. Prediction: Generating predictions using the deployed model. The output indicates the probability or risk of ischemic stroke for a given patient. This information supports early intervention and aids clinicians in making data-driven decisions, ultimately improving patient care.

4. ALGORITHM USED IN PROJECT

The proposed stacking ensemble model for ischaemic stroke prediction integrates an Artificial Neural Network (ANN) as a feature extractor with three powerful classifiers: Light Gradient Boosting Machine (LGBM), Extreme Gradient Boosting (XGBoost), and Random Forest (RF). The ANN processes input data from 5110 patient health profiles through multiple layers of neurons, capturing complex patterns and relationships to generate high-quality features that enhance the predictive power of subsequent models. These features are then fed into the ensemble of classifiers. LGBM, a decision-tree-based boosting algorithm, efficiently splits data into subsets and iteratively refines predictions by minimizing errors, leveraging its lightweight design for handling large datasets. XGBoost, another boosting algorithm, constructs multiple decision trees and optimizes predictions using gradient descent, ensuring high speed and accuracy. Random Forest complements the ensemble by building multiple decision trees and aggregating their outputs, reducing over fitting and providing robust predictions for classification tasks. By combining these models, the stacking ensemble achieves superior performance, with a 95.9% accuracy rate, offering an interpretable and reliable framework for early stroke detection in clinical decision support systems.

5. SYSTEM ARCHITECTURE



DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP146-151

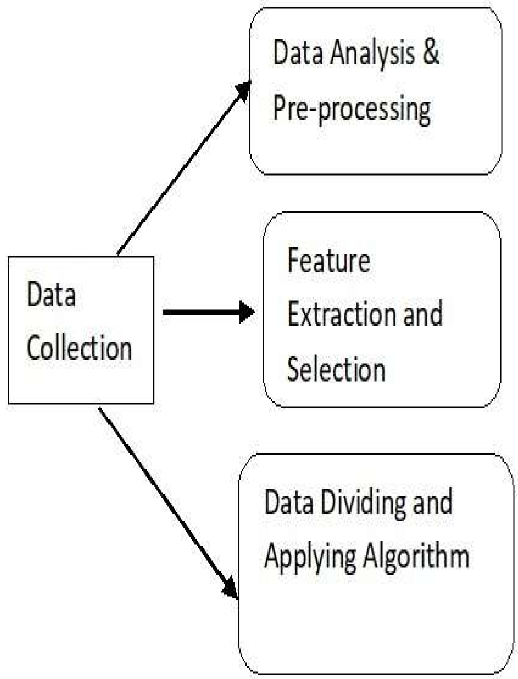
International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

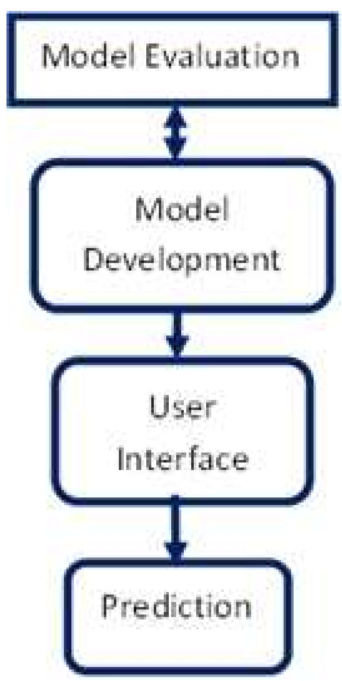
Volume 13, Issue 2s, 2025

6. DATA FLOW DIAGRAM:

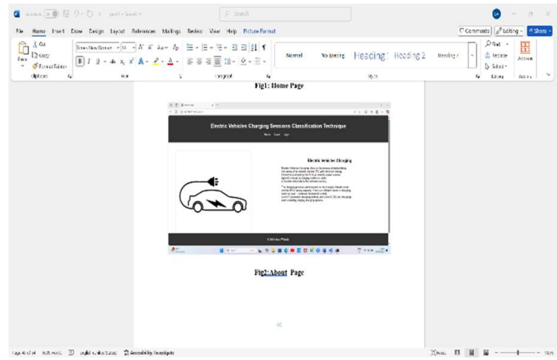
Level 0

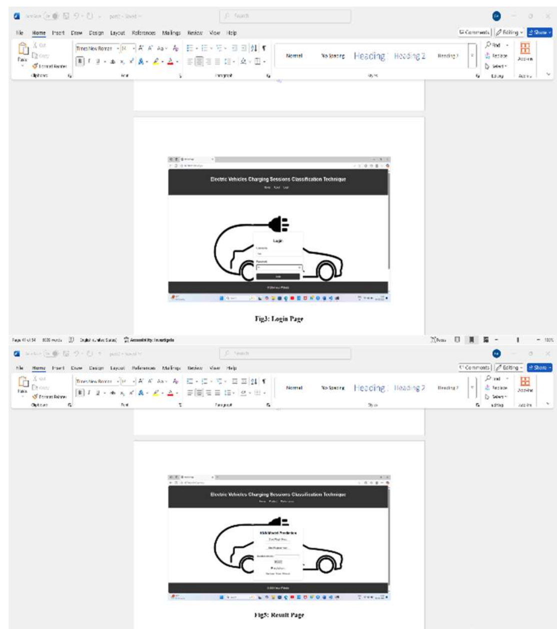


Level 1



7. RESULTS:





8. FUTURE ENHANCEMENT:

Future studies should concentrate on adding more and more varied datasets to improve the model's flexibility across various populations. For additional performance gains, efforts should also focus on investigating cutting-edge deep learning architectures and machine learning techniques. Furthermore, to confirm the model's effectiveness and usability in real-world healthcare contexts, real-world testing and incorporation into clinical processes would be crucial. Prediction models will be improved by cooperative research and multidisciplinary innovations, opening the door to more individualise and successful stroke prevention plans.

9. CONCLUSION:

This study concludes by highlighting how machine learning has the potential to revolutionise the prediction of ischaemic stroke. The suggested approach demonstrates its efficacy in diagnosing stroke risk by integrating ANN-based feature extraction with a strong stacking ensemble model, resulting in

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP146-151

impressive accuracy and AUCROC scores. Although there are still issues with data quality and interpretability, this study shows that using AI to detect strokes early is feasible. The application of this model in actual healthcare settings has enormous potential to enhance patient outcomes and maximise the use of available resources. To further improve and broaden the model's applicability and open the door to more individualised and effective stroke preventive techniques, more developments and cooperative efforts will be essential.

REFERENCES:

[1] L. Hwangbo, Y. J. Kang, H. Kwon, J. I. Lee, H.-J. Cho, J.-K. Ko, S. M. Sung, and T. H. Lee, ''Stacking ensemble learning model to predict 6-month mortality in ischemic stroke patients,'' Sci. Rep., vol. 12, no. 1, p. 17389, Oct. 2022.

[2] M. Mahmud, M. S. Kaiser, M. M. Rahman, M. A. Rahman, A. Shabut, S. Al-Mamun, and A. Hussain, ''A brain-inspired trust management model to assure security in a cloud based IoT framework for neuroscience applications,'' Cognit. Comput., vol. 10, no. 5, pp. 864–873, Apr. 2018.

[3] M. B. T. Noor, N. Z. Zenia, M. S. Kaiser, S. A. Mamun, and M. Mahmud, ''Application of deep learning in detecting neurological disorders from magnetic resonance images: A survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia,'' Brain Informat., vol. 7, no. 1, pp. 1–21, Oct. 2020.

[4] M. Mahmud, M. S. Kaiser, T. M. McGinnity, and A. Hussain, ''Deep learning in mining biological data,'' Cognit. Comput., vol. 13, no. 1, pp. 1–33, Jan. 2021.

[5] G. Fang, Z. Huang, and Z. Wang, ''Predicting ischemic stroke outcome using deep learning approaches,'' Frontiers Genet., vol. 12, Jan. 2022, Art. no. 827522.

[6] M. Chun, R. Clarke, B. J. Cairns, D. Clifton, D. Bennett, Y. Chen, Y. Guo, P. Pei, J. Lv, C. Yu, L. Yang, L. Li, Z. Chen, and T. Zhu, ''Strokerisk prediction using machine learning: A prospective cohort study of 0.5 million Chinese adults,'' J. Amer. Med. Inform. Assoc., vol. 28, no. 8, pp. 1719–1727, May 2021.

[7] M. S. Jabal, O. Joly, D. Kallmes, G. Harston, A. Rabinstein, T. Huynh, and W. Brinjikji, ''Interpretable machine learning modeling for ischemic stroke outcome prediction,'' Frontiers Neurol., vol. 13, May 2022, Art. no. 884693.

[8] A. Adiguzel, E. M. Arsava, and M. A. Topcuoglu, ''Temporal course ofperipheral inflammation markers and indexes following acute ischemic stroke: Prediction of mortality, functional outcome, and stroke-associated pneumonia,'' Neurological Res., vol. 44, no. 3, pp. 224–231, Sep. 2021.

[9] D. Chaudhary, V. Abedi, J. Li, C. M. Schirmer, C. J. Griessenauer, andR. Zand, ''Clinical risk score for predicting recurrence following a cerebral ischemic event,'' Frontiers Neurol., vol. 10, p. 1106, Nov. 2019.

[10] Z. Yuanyuan, W. Jiaman, Q. Yimin, Y. Haibo, Y. Weiqu, and Y. Zhuoxin, ''Comparison of prediction models based on risk factors and retinal characteristics associated with recurrence one year after ischemic stroke,'' J. Stroke Cerebrovascular Diseases, vol. 29, no. 4, Apr. 2020,
Art. no. 104581.

[11] N. Noorbakhsh-Sabet, R. Zand, Y. Zhang, and V. Abedi, ''Artificial intelligence transforms the future of health care,'' Amer. J. Med., vol. 132, no. 7, pp. 795–801, Jul. 2019.

[12] M. S. Singh and P. Choudhary, ''Stroke prediction using artificial intelligence,'' in Proc. 8th Annu. Ind. Autom. Electromechanical Eng. Conf. (IEMECON), Aug. 2017, pp. 158–161.

[13] Y. Zhang, W. Song, S. Li, L. Fu, and S. Li, ''Risk detection of strokeusing a feature selection and classification method,'' IEEE Access, vol. 6, pp. 31899–31907, 2018.

[14] L. García-Temza, J. L. Risco-Martín, J. L. Ayala, G. R. Roselló, and J. M. Camarasaltas, ''Comparison of different machine learning approaches to model stroke subtype classification and risk prediction,'' in Proc. Spring Simul. Conf. (SpringSim), Apr. 2019, pp. 1–10.

[15] M. S. Singh, P. Choudhary, and K. Thongam, ''A comparative analysis for various stroke prediction techniques,'' in Proc. Int. Conf. Comput. Vis.Image Process., Jaipur, India, Sep. 2019, pp. 27–29.

[16] M.A.Bari & Shahanawaj Ahamad," Process of Reverse Engineering of Enterprise InformationSystem Architecture" in International Journal of Computer Science Issues (IJCSI), Vol 8, Issue 5, ISSN: 1694-0814, pp:359-365,Mahebourg ,Republic of Mauritius , September 2011

[17] Dr.Abdul Bari ,Dr. Imtiyaz khan , Dr. Rafath Samrin , Dr. Akhil Khare , " VPC & Public Cloud Optimal Perfomance in Cloud Environment ",Educational Administration: Theory and

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP146-151

ISSN 2347–3657

Volume 13, Issue 2s, 2025

Practic,ISSN No : 2148-2403 Vol 30- Issue -6 June 2024

[18] Dr. Mohammed Abdul Bari,Arul Raj Natraj Rajgopal, Dr.P. Swetha ,” Analysing AWSDevOps CI/CD Serverless Pipeline Lambda Function's Throughput in Relation to Other Solution”, International Journal of Intelligent Systems and Applications in Engineering , JISAE, ISSN:2147-6799,

DOI: https://doi.org/10.62647/IJITCE2025V13I2sPP146-151