Abhi International Journals of Applied Science (AIJAS)
ISSN(online): ……..
Vol. 2, Issue. 1, January 2025

Evaluating Machine Learning and Deep Learning Techniques in Stroke Risk Prediction

Kanchan Vishwakarma

NIET, NIMS University, Jaipur, India

ARTICLE INFO

Article History:
Received December 15, 2024
Revised December 30, 2024
Accepted January 12, 2025
Available online January 25, 2025

Keywords:
Stroke risk prediction,
Random Forest,
Model transparency,
External validation

Correspondence:
E-mail:
kanchanvishwakarma200416@gmail.com

ABSTRACT

The current study explores the application of Machine Learning (ML) and Deep Learning (DL) techniques to predict stroke risk, addressing major challenges such as model accuracy, adaptability, feature importance, transparency, and external validation. A quantitative approach is used to evaluate various ML algorithms, and in particular, Random Forest has been highlighted because of its superior predictive accuracy, while stressing the adaptability of DL models across different demographic contexts. The study further explores the role of feature significance in enhancing context-specific predictions, the challenges of model explainability in clinical adoption, and the critical importance of external validation in ensuring generalizability. The results underline the transformative potential of ML and DL in advancing personalized healthcare strategies for stroke prediction while identifying existing gaps in transparency and validation practices. This synthesis lays a foundation for future research to standardize external validation protocols and improve model transparency for wider clinical adoption.

Introduction

This chapter discusses the important issue of stroke, highlighting that it is leading up as the cause of disability and death. The unpredictability and severity in impact associated with the stroke make it crucial to develop advanced prediction methods. The core research question is one concerning the evaluation of ML and DL techniques for stroke risk prediction with five sub-research questions: effectiveness of different ML algorithms in stroke prediction, use of DL models in diverse contexts, role of feature importance in terms of giving recommendations in context, challenge of model explainability and transparency, and impact of external validation on the accuracy of models. This study employs a quantitative methodology, examining the relationship between independent variables (various ML and DL algorithms) and dependent variables (prediction accuracy and applicability across contexts). The paper follows a structured approach from literature review to methodology explanation, results presentation, and a conclusion discussing the theoretical and practical implications, highlighting the research's significance in advancing personalized healthcare strategies for stroke.

Literature Review

It looks at existing literature on ML and DL techniques used for the prediction of stroke risk, discussing the five topics described above: performance of ML algorithms, application of DL models, importance of feature significance, challenge of model transparency, and external validation. These questions give way to concrete observations: "Machine Learning Algorithm Effectiveness for Stroke Prediction," "Deep Learning Models in Varied Applications," "Feature Importance for Stroke Prediction," "Model Explainability and Interpretability Challenges," and "How External Validation Can Affect the Accuracy of the Model." With all these advancements, there are still gaps to fill, like that of model transparency and external validation. Conjectures are provided to each domain because these gaps can be filled.

This section engages in a detailed review of the existing work on ML and DL approaches for the prediction of stroke risk. The analysis is organized around five key themes: the effectiveness of ML algorithms, the use of DL models, the importance of feature importance, the challenges around model transparency, and the role of external validation. Each theme offers different insights: the "Effectiveness of Machine Learning Algorithms in Stroke Prediction" explores several algorithms and their predictive potential; the "Application of Deep Learning Models in Diverse Contexts" describes how DL techniques can be tailored to various medical settings; "Importance of Feature Significance in Stroke Prediction" highlights which variables have the greatest influence on predictions; "Challenges of Model Explainability and Transparency" deals with the opaque nature of these models; and "Impact of External Validation on Model Accuracy" highlights the importance of testing these models in real-world settings outside the development stage. Although the strides made in these areas are notable, there is much left to be filled, especially concerning the improvement of model transparency and the development of more robust external validation protocols. In this regard, hypotheses are formulated for each theme, which provide potential strategies and methodologies for improvement.

Effectiveness of Machine Learning Algorithms in Stroke Prediction

Early studies established the feasibility of ML algorithms in stroke prediction, with the main argument based on their capability to process large datasets. Yet, early studies often failed to compare the algorithms used thoroughly. Later studies did better by comparing the performance of various ML models and identified Random Forest as the best-performing model for accuracy in stroke prediction. Still, the challenge persists with model explainability. Hypothesis 1: Random Forest is more accurate than other ML algorithms for stroke risk prediction.

Application of Deep Learning Models in Various Contexts

Early studies on DL models have shown their ability to process large datasets with high accuracy. However, such studies usually did not focus on the applicability of DL models in different demographic contexts. Further research was expanded by applying DL models in various populations and emphasized their adaptability but also raised difficulties in model interpretation. Hypothesis 2: Deep learning models are highly adaptable to different demographic contexts in stroke risk prediction.

Significance of Feature in Stroke Prediction

Early researches on feature significance were aimed at determining significant risk factors for stroke, often only within a certain population. Further research introduced more holistic approaches that emphasized the requirement for context-specific recommendations since the risk factors for stroke vary geographically. Although there is a step forward, the integration of feature significance into predictive models is still not fully explored. Hypothesis 3: The inclusion of feature significance enhances the context-specific accuracy of stroke risk prediction models.

Challenges of Model Explainability and Transparency

Early research pointed out that ML and DL models are inherently black-boxed, which does not allow for their use in clinical settings. Later studies have focused on increasing model transparency by using methods such as feature importance analysis; however, full explainability is still not achieved. Hypothesis 4: Increasing model transparency increases the acceptance of ML and DL models in stroke prediction among clinicians.

Impact of External Validation on Model Accuracy

Initial studies often relied on internal validation, leading to overfitting and limited generalizability. Later research emphasized external validation, demonstrating its critical role in improving model robustness and reliability. Despite advancements, external validation practices are not yet standardized. Hypothesis 5: External validation significantly enhances the accuracy and generalizability of stroke risk prediction models.

Method

This chapter provides the methodology used in a quantitative research design that assesses the application of ML and DL in stroke risk prediction. This will include describing how data was collected, what variables were applied, and the type of statistical procedures applied to validate the accuracy of findings.

Data

The data were derived from a systematic review of 31 articles that had the inclusion criteria with respect to the application of ML and DL in stroke prediction. The PRISMA method guided the data collection process. It involved various datasets from countries such as China, India, and Bangladesh. Secondary datasets were chosen because they are comprehensive in nature and relevant to stroke risk factors. Sampling involved choosing studies that have a proper validation process so that the dataset selected is robust enough for analysis.

Variables

The independent variables include ML and DL algorithms such as Random Forest, neural networks, and decision trees. The dependent variables are based on prediction accuracy, model adaptability, and clinical applicability. The control variables include patient demographics, dataset size, and model complexity. Measurement of variable reliability is ensured with literature support in ML and DL methodologies to cover all aspects of the algorithms in predicting stroke.

Results

The results start with descriptive statistical analyses of the performance of ML and DL algorithms in stroke prediction, followed by regression analyses to validate the proposed hypotheses. The results clearly indicate that Random Forest is more accurate, DL models are adaptable, feature significance is important, achieving model transparency is challenging, and external validation is critical. These results point out that ML and DL methods can really improve the current personalized healthcare stroke strategies by highlighting the lack of transparency and proper validation in some models.

Random Forest and Prediction Accuracy

This result proved Hypothesis 1 since Random Forest produced better accuracy on stroke risk predictions compared to the other ML-based algorithms. Using data from different experiments, Random Forest is shown to have higher predictive accuracy because of its ability to process complex data as well as interacting features. The key variables include algorithm type and metrics of prediction accuracy, and the highest scores are achieved by Random Forest on all datasets.

Empirical significance implies that the ensemble approach of Random Forest effectively captures the nuances of stroke risk factors and thus aligns with theories about robust model performance. This finding emphasizes the need for appropriate algorithm selection for accurate stroke prediction.

Adaptability of Deep Learning Models

This finding supports Hypothesis 2, indicating that deep learning models are highly adaptable to different demographic contexts in stroke risk prediction. Analysis of data from diverse populations shows that DL models maintain high prediction accuracy across various demographic groups. Key variables include model type and adaptability metrics, with DL models demonstrating consistent performance in different contexts. The empirical importance reiterates the flexibility of DL models, implying that their complicated architecture can absorb varied risk factors. This indicates that DL models have a bright future in personalized healthcare for stroke.

Inclusion of Feature Significance

This conclusion verifies Hypothesis 3, and feature significance is added to enhance the context-specific accuracy of the stroke risk prediction models. From the analysis of the papers that include feature significance, improved prediction accuracy with relevance to particular populations can be seen. Key variables are feature importance metrics and prediction accuracy, with models that include feature significance outperforming those that do not. The empirical significance of the results indicates that knowing and including key risk factors improves model precision and applicability. This result highlights the requirement for context-specific approaches in stroke prediction, as suggested by theories on personalized healthcare.

Improving Model Transparency

This result supports Hypothesis 4. The increase in model transparency leads to the improved clinical acceptance of ML and DL models in stroke prediction. Studies focused on transparency techniques show models with improved explainability and clinical acceptance. Key variables involve transparency techniques, including acceptance metrics, where transparent models have greater integration with the clinician. The empirical significance suggests that improving model explainability addresses concerns about the black-box nature of ML and DL models, facilitating their clinical adoption. This finding underscores the importance of transparency in advancing the practical application of stroke prediction models.

Role of External Validation

This finding confirms Hypothesis 5, emphasizing that external validation significantly enhances the accuracy and generalizability of stroke risk prediction models. Studies using external validation show improved predictive accuracy and robustness. Validations include both types: that of accuracy and metrics. More importantly, external validation shows the superiority of the model over an internally validated one. Empirical significance gives the rationale of why it needs to be externally validated before the application for reliability and universality across diverse contexts. Hence, rigorous practices in validation of models would go a long way in achieving more effective stroke predictive models.

Conclusion

This synthesis discusses the application of ML and DL techniques in stroke risk prediction, pointing to their roles in improving the accuracy of predictions, adaptability, and clinical acceptability. In this regard, the research suggests that Random Forest is the dominating ML algorithm. It further puts emphasis on the adaptability of DL models while underlining the significance of feature significance and model transparency. Despite these improvements, there remain challenges in terms of model explainability and validation practices. Future research should focus on standardizing external validation and exploring new transparency techniques to enhance model integration into clinical practice. By addressing these areas, future studies can provide deeper insights into the potential of ML and DL techniques in transforming stroke risk prediction and advancing personalized healthcare strategies.

References

[1] Deo, R. C. (2015). Machine learning in medicine. *Circulation*, 132(20), 1920-1930.

[2] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (pp. 4765-4774).

[3] Kam, H. J., & Kim, H. Y. (2017). Learning representations for the early detection of stroke using deep learning. *BMC Medical Informatics and Decision Making*, 17(1), 121.

[4] Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347-1358.

[5] van der Velden, A., Jaspers, M. W. M., & Elbers, P. W. G. (2021). Machine learning algorithms in clinical settings: A review of their application and impact. *Journal of Biomedical Informatics*, 118, 103790.

[6] Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics*, 22(5), 1589-1604.

[7] Lipton, Z. C. (2016). The mythos of model interpretability. *arXiv preprint arXiv:1606.03490*.

[8] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135-1144).

[9] Zhang, Z., & Zou, Q. (2019). Advances and challenges of deep learning applications in biomedical informatics. *Briefings in Bioinformatics*, 20(2), 626-641.

[10] Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causality and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(4), e1312.