CHAPTER V
CONCLUSION

5.1 Conclusion

Based on the results obtained from evaluating the stacking model for each base-learner model (Logistic Regression, Random Forest, and Extreme Gradient Boosting) and the meta-learner (Artificial Neural Network), along with the final prediction results and the Explainable Artificial Intelligence (XAI) interpretations, several conclusions can be drawn by the author:

1. Based on the series of processes that have been conducted, the stacking model in this study combines the prediction results from three base-learners (initial layer/level 0) Logistic Regression, Random Forest, and Extreme Gradient Boosting and utilizes them as inputs (meta-features) for the meta-learner (second layer/level 1), which also involves a fold cross-validation process to prevent overfitting caused by data leakage.
2. Each individual base-learner and meta-learner model in the study was evaluated across all of its stacking combinations, encompassing accuracy, precision, recall, F1-score, and ROC-Score. It was observed that there were increases as well as decreases in each metric tested across the various combinations. The accuracy achieved by each model and combination did not reach the highest possible figures (~60% - 75%). However, the models possess a high sensitivity or Recall, enabling them to reduce the risk of missing stroke patients (false negatives) as much as possible, although this comes at the cost of a high number of false positive warnings. The ROC-score also yielded high values, proving the model can effectively distinguish between the two target classes.
3. XAI interpretations were conducted globally using SHAP for each base-learner and every stacking combination on the ANN meta-learner. It was found that the age feature is the most effective prediction determinant for each base-learner model, whereas at the meta-learner level, the Logistic Regression model's prediction result was assessed as the most dominant primary determining factor in driving the ANN's decisions. XAI interpretations were also conducted locally using LIME for each base-learner and every stacking combination on the ANN meta-learner using two patient case studies. This revealed that the age, bmi, and

hypertension features are the most influential features in determining whether or not a patient will have a stroke across all base-learners. For the meta-learner level, the model exerting the highest influence on the predictions remained the same: Logistic Regression.

4. With the creation of the Graphical User Interface (GUI), the developed models can be integrated into this interface and utilized locally by external users. Elements such as the feature input forms, the prediction results area, and both XAI interpretation areas are functioning perfectly with the assistance of Streamlit, thereby completing this research.

5.2 Suggestion

Based on the entirety of this research, there are several recommendations for improvement in future studies:

1. The tuning process during the model training phase (base-learner and meta-learner) conducted in this study was only executed a limited number of times, prioritizing effectiveness and factoring in the author's computational limitations. Further research is necessary to discover more effective tuning results, as there are numerous tuning combinations that can be explored for each model.
2. The accuracy results for each model in this study can still be improved in future research. Subsequent studies are recommended to acquire more diverse datasets, conduct superior hyperparameter tuning, and explore alternative model-building techniques so that the final prediction accuracy can be enhanced.
3. The XAI interpretation techniques from this study are expected to serve as a foundation for interpreting similar models in future research. Future studies are advised to investigate whether there are other XAI methodologies that can be utilized for model interpretation beyond merely analyzing features and model predictions.
4. The GUI developed in this study is not yet perfect, as it is utilized solely as a visualization tool for this research and operates locally using the simple Streamlit framework. Future research is recommended to develop a dynamic, interactive website that can be accessed by the wider public and genuinely deployed for practical use within the medical field.