COMP90089 - Project Proposal

Group Number: 14
Jiayuan Li (Student ID: 1404463)
Ron Chen (Student ID: 1181506)
Xinyi Wang (Student ID: 1360106)
Zixuan Guo (Student ID: 1124606)
Word Count: 816

Augest 2024

1 Research Objective

This project aims to leverage machine learning to predict the likelihood of mortality in stroke patients and analyse the impact of patient features on mortality risk. Stroke is a significant global health issue, responsible for 4.5 million deaths annually worldwide and causing substantial disability to survivors[11]. As stroke continues to be a leading cause of death and disability, better understanding contributing factors and improving prediction models for stroke outcomes can aid better healthcare management and personalised treatment strategies[5]. Deep learning has already been applied to many health applications[3, 9], demonstrating improvements to stroke mortality prediction performances[1]. We aim to apply and hypothesise that Transformers as the state-of-the-art deep classification model can further improve stroke mortality prediction performance[10].

2 Data Source & Phenotyping

For this research project, the MIMIC-IV dataset version 2.2 will be utilised as the primary data source. MIMIC-IV is a comprehensive, publicly available database. This dataset includes a wide range of clinical data (i.e. vital signs, laboratory test results, medication records, demographic information, and imaging data) making it suitable to apply data modelling to better understand illnesses and impact of treatments, including those related to stroke[7, 8].

To perform digital phenotyping, the following steps will be undertaken:

2.1 Variable Extraction

To perform phenotyping, this study will extract variables associated with stroke outcomes based on previous literature. Key patient characteristics including:

age, sex, and comorbidities such as hypertension, diabetes, and previous stroke were found to significantly impact stroke mortality in previous studies[6]. Additionally, the severity of stroke on admission will be assessed using the National Institutes of Health Stroke Scale (NIHSS), which is recorded in the MIMIC-IV dataset and serves as a critical predictor of stroke outcomes[5]. Lab test values such as blood glucose and renal function are critical factors influencing stroke prognosis, with elevated glucose levels associated with larger infarct volumes[4]. Other comorbidities, including heart disease and dyslipidemia had documented effects on post-stroke outcomes. In addition, advanced age and length of hospital stay will be captured, given their association with increased mortality risk within the first year post-stroke[2]. All aforementioned features are included for this study.

2.2 Stroke Definition and Classification

The definition of stroke used in this study will follow the World Health Organization's guidelines: 'clinical signs of focal or global disturbance of cerebral function lasting more than 24 hours or leading to death, excluding other causes such as infection or tumour'[11]. We will select ICU patients who match both this criteria and the ICD-10/ICD-9 codes for stroke to form our cohort.

3 Methodology

- **Exploratory Data Analysis (EDA) & Pre-processing:** Visualise data distributions, inter-correlations and associations to stroke outcomes, removing outliers and resolving missing value problems (the technique used will depend on the specific problem).

- **Data Splitting:** Divide dataset into training and testing sets (70%, 30%) and perform 10-fold Cross Validation (10CV) on the training data.

- **Addressing Class Imbalance:** Experiment with techniques such as Synthetic Minority Over-sampling Technique (SMOTE) on training data to improve model performance.

- **Feature Engineering:** Fit one hot encoding transformation and normalisation to discrete and continuous features respectively on each 10CV training set, and apply the fitted transformations to matching 10CV validation sets.

- **Feature Selection:** Apply correlation filter to retain only one feature in each highly correlated feature pair and apply F-test filter to remove irrelevant features. Features will be further selected as a hyperparameter during tuning.

- **Model Selection and Hyperparameter Tuning:** Experiment with various supervised machine learning algorithms (i.e. logistic regression, XGBoost, multilayer perceptron and transformers), tune to find optimal hyper-parameters based on 10CV's validation set balanced accuracy.

- **Model Evaluation:** Find best model based on 10CVs validation balanced accuracy. Then, fit best model using 70% training data and assess the best model's test dataset performance using metrics like balanced accuracy, precision, recall, F1-score, ROC-AUC, and calibration curves.

- **Model Interpretation:** Use explainable-AI tool SHAP to visualise and analyse the effect of each feature on the best model's predicted mortality likelihood in stroke patients, both on individual samples and entire testing set.

3.1 Performance Metrics

- **Balanced Accuracy:** Weighted accuracy of each class providing fair model performance evaluation on imbalanced evaluation datasets.

- **Precision:** Measures what proportion of predicted moralities were actual mortalities.

- **Recall:** Measures what proportion of actual mortalities were predicted as mortalities

- **F1-score:** Balances precision and recall and provides a fair model performance assessment for imbalanced evaluation datasets.

- **ROC-AUC:** Measures the ability of the model to discriminate between the two outcomes. It aggregates model performance across different thresholds, with a value of 1 indicating perfect classification and 0.5 representing random guessing.

- **Calibration Curves:** Assesses whether predicted probabilities are close to actual outcomes, thus identifying if model is overconfident or underconfident in its predictions.

3.2 Expected Outcomes

- **Improved Prediction Model:** A model that enhances the prediction accuracy of stroke outcomes, aiding healthcare professionals in identifying patients with high mortality risk.

- **Optimised Patient Management:** Better management of stroke patients through better understanding how factors impact stroke mortality risk, leading to reduced mortality.

- **Deep Learning Applications in Health:** Explore whether more complex deep learning models enhance predictions, hence improving healthcare decision-making despite the interpretability tradeoff.

References

[1] Armin Abdollahi, Xinghong Ma, Jiahao Zhang, Daijia Wu, Tongshou Wu, Zizheng Ye, and Maryam Pishgar. Enhanced mortality prediction in icu stroke patients via deep learning, 2024.

[2] Barbara E Bates, Dawei Xie, Pui L Kwong, Jibby E Kurichi, Diane Cowper Ripley, and Margaret G Stineman. One-year all-cause mortality after stroke: a prediction model. *PM&R*, 6(6):473–483, 2014.

[3] L Butler, F Gunturkun, L Chinthala, I Karabayir, MS Tootooni, B Bakir-Batu, T Celik, O Akbilgic, and RL Davis. Ai-based preeclampsia detection and prediction with electrocardiogram data. *Frontiers in Cardiovascular Medicine*, 11:1360238, March 4 2024.

[4] Wei-Min Ho, Jr-Rung Lin, Hui-Hsuan Wang, Chia-Wei Liou, Ku-Chou Chang, Jiann-Der Lee, Tsung-Yi Peng, Jen-Tsung Yang, Yeu-Jhy Chang, Chien-Hung Chang, et al. Prediction of in-hospital stroke mortality in critical care unit. *Springerplus*, 5:1–9, 2016.

[5] Timothy Ingall. Stroke-incidence, mortality, morbidity and risk. *JOURNAL OF INSURANCE MEDICINE-NEW YORK THEN DENVER–.*, 36:143–152, 2004.

[6] Annette Ingeman, Lars Pedersen, Heidi H Hundborg, Palle Petersen, Susanne Zielke, Jan Mainz, Paul Bartels, and Søren P Johnsen. Quality of care and mortality among patients with stroke: a nationwide follow-up study. *Medical care*, 46(1):63–69, 2008.

[7] Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. Mimic-iv. *PhysioNet. Available online at: https://physionet. org/content/mimiciv/1.0/(accessed August 23, 2021)*, pages 49–55, 2020.

[8] Alistair EW Johnson, Lucas Bulgarelli, Lu Shen, Alvin Gayles, Ayad Shammout, Steven Horng, Tom J Pollard, Sicheng Hao, Benjamin Moody, Brian Gow, et al. Mimic-iv, a freely accessible electronic health record dataset. *Scientific data*, 10(1):1, 2023.

[9] R Tourani, DH Murphree, A Sheka, GB Melton, DJ Kor, and GJ Simon. Consensus modeling: Safer transfer learning for small health systems. *Artificial Intelligence in Medicine*, 154:102899, August 2024. Epub 2024 May 24.

[10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, pages 5998–6008. Curran Associates, Inc., 2017.

[11] Charles D A Wolfe. The impact of stroke. *British Medical Bulletin*, 56(2):275–286, 01 2000.