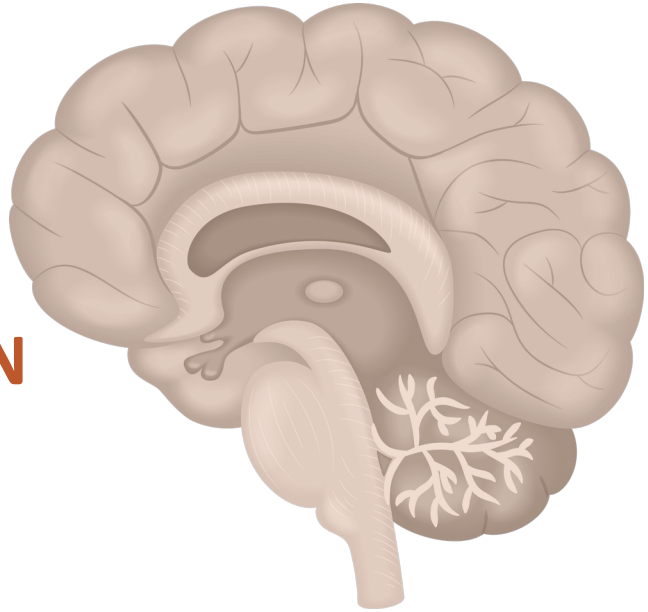

BRAIN STROKE RISK PREDICTION

STRATEGIES FOR IMPROVING EARLY BRAIN STROKE PREDICTION AND DIAGNOSIS USING MACHINE LEARNING

Presented By: Athira Reghunath (W9561613)

Supervisor: Glen Hopkinson

Second Reader: Lewis Golightly

Outline

1. Introduction
2. Existing Issues
3. Solutions
4. Evaluation
5. Results
6. Conclusion

Introduction

Stroke ranks as the second most common cause of mortality and a significant contributor to global disability. The objective of this research is to revolutionize the field of stroke disease prediction through the application of advanced machine-learning techniques. With a primary focus on early detection, the study incorporates a diverse range of features, such as gender, hypertension, body mass index, heart disease, average glucose level, smoking status, marital status, employment type, and residential area. By combining machine learning with a comprehensive set of patient attributes, this research strives to contribute to more accurate and timely stroke predictions, ultimately enhancing healthcare outcomes.

Existing Issues

1. **Limited Predictive Accuracy:**
   - Current predictive models may struggle to achieve high accuracy due to the complexity of stroke risk factors and interactions.
2. **Data Imbalance:**
   - Imbalances in datasets, where the number of stroke cases is significantly lower than non-stroke cases, can lead to biased models and affect their ability to generalize.
3. **Incomplete Risk Factor Coverage:**
   - Some existing models may not incorporate a comprehensive set of risk factors, potentially overlooking crucial elements that contribute to stroke risk.
4. **Lack of User-friendliness:**
   - The current interfaces of existing Brain Stroke Risk Prediction web applications lack user-friendliness.

Solutions

- Develop a robust machine learning model for stroke risk prediction.
- Explore strategies to enhance the accuracy of predictive models for early intervention.
- Develop a machine learning model with higher accuracy than the existing Brain Stroke Prediction models.
- Develop a web application with a user-friendly interface for Brain Stroke Risk Prediction.

Evaluation

- Data Collection
- Data Preprocessing
- Exploratory Data Analysis
- Predictive Machine Learning Models
  - Logistic Regression
  - Decision Tree
  - Random Forest
  - K-Nearest Neighbors(KNN)
  - Support Vector Machine(SVM)
- Implementation of Web Application

Results

- Logistic Regression - 93.83%
- Decision Tree - 89.53%
- Random Forest - 93.73%
- K-Nearest Neighbors(KNN) - 93.44%
- Support Vector Machine(SVM) - 93.93%

After cross validating 5 machine learning methodologies, Support Vector Machine is the best.

Conclusion

The major contributions of this study are;

- This study attains an accuracy of 93.93%, surpassing the previous results achieved by other researchers in this specific area.
- In this research, five classifiers and various machine learning techniques, such as label encoding, outlier removal, and cross-validation, are used to achieve the optimal outcome.
- A web application is created based on this research, capable of accurately calculating results using real-time inputs and a better user-friendly interface.
- Among the five classifiers utilized, Support Vector Machine and Logistic Regression exhibit the highest accuracies, reaching 93.93% and 93.83%, respectively.

THANK YOU