**PROJECT CODE: 2024-P06**

**Title: Innovations in Stroke Identification: A Machine Learning-Based Diagnostic Model Using Neuroimages**

**Abstract**

Cerebrovascular diseases such as stroke are among the most common causes of death and disability worldwide and are preventable and treatable. Early detection of strokes and their rapid intervention play an important role in reducing the burden of disease and improving clinical outcomes. In recent years, machine learning methods have attracted a lot of attention as they can be used to detect strokes. The aim of this study is to identify reliable methods, algorithms, and features that help medical professionals make informed decisions about stroke treatment and prevention. To achieve this goal, we have developed an early stroke detection system based on CT images of the brain coupled with a genetic algorithm and a bidirectional long short-term Memory (BiLSTM) to detect strokes at a very early stage. For image classification, a genetic approach based on neural networks is used to select the most relevant features for classification. The BiLSTM model is then fed with these features. Cross-validation was used to evaluate the accuracy of the diagnostic system, precision, recall, F1 score, ROC (Receiver Operating Characteristic Curve), and AUC (Area Under The Curve). All of these metrics were used to determine the system’s overall effectiveness.

In addition to the core methodology, our system incorporates data augmentation techniques to enhance the robustness of the model and mitigate overfitting. Advanced preprocessing steps, including noise reduction and normalization, are applied to the CT images to ensure high-quality input data. The integration of a genetic algorithm not only optimizes feature selection but also accelerates the training process. Our approach leverages parallel computing for efficient model training and validation. Extensive experimentation with various BiLSTM architectures was conducted to identify the most effective configuration. The system’s adaptability to different datasets highlights its potential for widespread clinical application.