Machine Learning Integrated Software for Prediction, Diagnosis, and Prognosis of Amyotrophic Lateral Sclerosis

Hoang, Viet Ha (School: Tran Phu Gifted High School)
Bui, Phuong Tue (School: Tran Phu Gifted High School)

Our software system is designed to assist medical personnel in early diagnosing and prognosticating of Amyotrophic Lateral Sclerosis (ALS), a neurodegenerative disease that is often misdiagnosed or overlooked. The software features five tabs, each with distinct functions: three for users and two for administrators. Tab 3 predicts the likelihood of ALS using clinical data including electromyography (EMG), blood, and urine tests, analyzed through Neural Networks and Logistic Regression. Tab 4 uses Decision Trees and K-Nearest Neighbors algorithms, leveraging the ALS Functional Rating Scale-Revised to determine patient stages. Tab 5 employs Linear Regression to estimate patient remaining time, utilizing specially refined indices. The input datasets contain basic, easily collectible, and cost-effective data, avoiding complex tests like MRI and spinal tap, which makes it particularly suitable for detecting early-stage symptoms such as muscle twitches. The data collection process involved compiling demographic features and clinical tests from a control group of 300 individuals without neurological disorders and an experimental group of over 11,700 ALS patients sourced from the PRO-ACT global database and Vietnam hospitals. Data were prepared and standardized for model input, with noise filtering, model training, and validation taking place in tab 2 to ensure optimal index compatibility and maximum accuracy. After training, tab 1 is used to identify, implement, and customize the optimal hyperparameters for each model. Through flexible adjustments in data selection, model settings, and fine-tuning, an advanced tool has been developed achieving accuracy rates of up to 82%, 91%, and 78% for ALS prediction, diagnosis, and prognosis, respectively.