HUMAN STRESS LEVEL DETECTION : MACHINE LEARNING CLASSIFICATION FOR DETECTING HUMAN STRESS LEVELS

ABSTRACT

The Human Stress Level Detection project is a machine learning classification initiative designed to predict stress levels based on various physiological indicators. The dataset includes features such as physiological metrics related to sleep quality, respiration, body temperature, limb movement, blood oxygen, eye movement, and heart rate. Participants will utilize classification algorithms to analyze these physiological parameters and develop a model capable of detecting and categorizing stress levels. This project is of practical significance for applications in healthcare, well-being, and stress management, providing a valuable tool for early stress detection

DATA PREPROCESSING

The dataset used in this study is the sa yopillow dataset. It consists of 631 instances and includes eight medical predictor variables (attributes) and one target variable (outcome). The attributes are

- **sr** - Sleep Quality
- **rr** - Respiration Rate
- **t** - Body Temperature
- **lm** - Limb Movement
- **bo** - Blood Oxygen Levels
- **rem** - Eye Movement (REM Sleep)
- **sr** - Heart Rate
- **hr** - Heart Rate Variability
- **sl** - Stress Level

Missing values were handled by imputing with the mean or dropping columns with a high proportion of missing data. Categorical variables were encoded into numerical values to make the data suitable for machine learning models.

EXPLORATORY DATA ANALYSIS:

EDA was performed to understand the distribution of human stress occurrences across different features. Visualizations showed the distribution and relationships between attributes and the outcome variable, providing insights into the patterns within the dataset.

MODEL TRAINING AND EVALUATION:

Various machine learning algorithms were used, including Logistic Regression, Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) with Linear and Polynomial kernels, Random Forest Classifier, and Voting Classifier. The dataset was split into training and testing sets, and models were trained and evaluated on the test set using accuracy, confusion matrix, and classification report metrics.

* Logistic Regression: 80%

* LDA: 79%

* Linear SVC: 79%

* Polynomial kernel SVC: 79%

* Random Forest Classifier: 82%

* Voting Classifier: 80%

MODEL DEPLOYMENT:

The Random Forest model, which had the highest accuracy, was serialized using the ‘pickle’ library. This allowed the model to be saved and reloaded for future predictions without the need for retraining, facilitating deployment in a production environment for real-time predictions.

CONCLUSION:

The "Human Stress Level Detection" project successfully demonstrated the application of machine learning techniques to predict and classify stress levels based on physiological indicators. By leveraging a diverse set of features such as sleep quality, respiration, body temperature, limb movement, blood oxygen levels, eye movement, and heart rate, the project addressed a significant need in healthcare and well-being for early stress detection and management.