Human Activity Recognition Project Report

Introduction & Objective

This project focuses on Human Activity Recognition (HAR) using smartphone sensor data. The primary objective is to classify and predict user activities with high accuracy by leveraging both Machine Learning (ML) and Deep Learning (DL) models.

Dataset Overview

The dataset consists of smartphone accelerometer and gyroscope sensor readings collected from participants performing various activities such as walking, standing, sitting, walking upstairs, walking downstairs, and lying down.

Data Preprocessing & Feature Engineering

Data preprocessing involved handling missing values, normalization, and extraction of relevant features from raw sensor signals. Feature engineering techniques were applied to create meaningful variables that enhance predictive performance.

Exploratory Data Analysis (EDA)

EDA was conducted to explore label distributions, visualize activity frequencies, and detect potential inconsistencies. Insights included balanced distribution of most activities and slightly higher speed observed in walking downstairs.

Dimensionality Reduction (PCA & t-SNE)

Principal Component Analysis (PCA) was applied for dimensionality reduction, enabling visualization of feature separability across activities. t-SNE was used to capture non-linear structures, providing a clearer separation of activity clusters.

Model Implementation

Multiple ML models were implemented and compared: Logistic Regression, Support Vector Machines (SVM), Decision Trees, Gradient Boosting, and Linear SVC. Additionally, a Deep Learning model using LSTM was developed to capture temporal dependencies in sensor data.

Model Evaluation & Comparison

Models were evaluated using accuracy score. The LSTM model demonstrated superior performance due to its ability to handle sequential time-series data, while classical ML models provided competitive baseline results.

Key Insights & Conclusion

The project highlighted the effectiveness of combining ML and DL approaches for HAR. Dimensionality reduction improved interpretability, while LSTM enhanced accuracy. This approach can be extended to real-time applications such as fitness monitoring, elderly care, and activity-based health interventions.