Student Stress Analysis Project Report

Project Overview

This project involved analyzing two student stress datasets, focusing on psychological, academic, social, and environmental stress factors. The goal was to uncover insights through Exploratory Data Analysis (EDA), apply statistical methods, and build machine learning models for stress classification.

Key Analysis Performed

- Conducted comprehensive EDA on student stress datasets, examining distributions, correlations, and outliers.
- Performed feature importance analysis using Mutual Information and dimensionality reduction with PCA.
- Applied statistical tests (ANOVA, normality tests) to identify significant factors.
- Provided preprocessing recommendations such as scaling, handling multicollinearity, and class balance adjustments.

Machine Learning Modeling

- Built and evaluated multiple machine learning models including Logistic Regression, Random Forest, XGBoost, LightGBM, and SVM.
- Compared model performance using accuracy, confusion matrices, and classification reports.
- Achieved robust performance in multi-class stress classification tasks.

Resume Highlights

- Performed comprehensive EDA on student stress datasets to uncover key academic, psychological, and social drivers.
- Built and evaluated multiple machine learning models, achieving strong performance in stress classification.
- Applied feature importance, PCA, and statistical testing to enhance interpretability and guide data preprocessing.

Conclusion

The project successfully demonstrated how data-driven approaches can identify critical stress factors among students and build predictive models to classify stress levels. This analysis can inform academic institutions and mental health initiatives to better support student well-being.