Analysis and classification of body performance data

Abstract

The aim of this project was to explore the characteristics affecting the classification of body performance, and to apply several models to the data to help the community in the future to classify their physical performance and reduce obesity diseases, based on the characteristics below: Gender, Height_cm , Weight_kg , Body_fat% , Diastolic , Systolic , GripForce , Sit and bend forward_cm , Sit-up counts, Broad jump_cm.

Design

This project originates for virtual assistant this is data that confirmed the grade of performance with age and some exercise performance data.

Data

The dataset is a collection of performance records collected from Korea Sports Promotion Foundation with 12 attributes and Rows there are 13393 instances each represents a person's information.

Algorithms

- Feature engineering.

- Split datafarme by class and gender.

- Removing outliers.

- Calculate weights of each class.

Tools

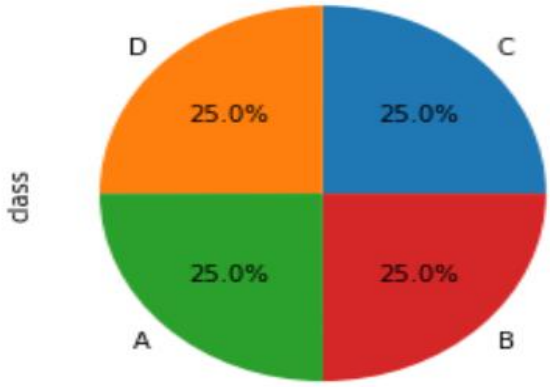
For Exploratory data analysis and visualisation we used:

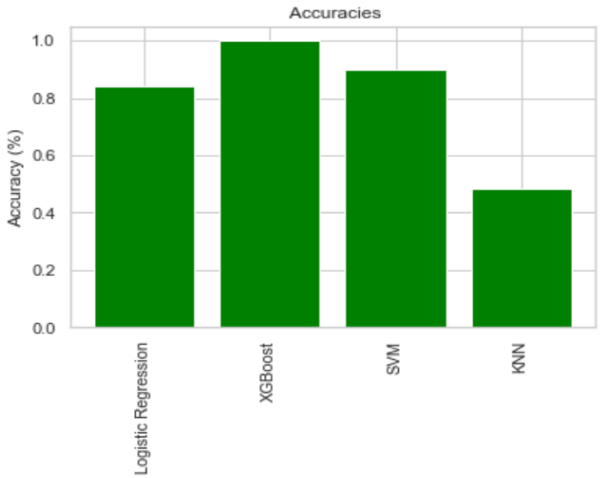
Numpay– Pandas – Seaborn – Matplotlib.

For Modeling and validation/testing we used:

– Sklearn.

Communication


dass
D
C
25.0%
25.0%
25.0%
25.0%
A
B


Accuracies
Accuracy (%)
1.0
0.8
0.6
0.4
0.2
0.0
Logistic Regression
XGBoost
SVM
KNN