Project Proposal (Body Performance Data)
S15055

Predict Body Performance class, based on physical characteristics & performance on physical tests.

Introduction/background of the study.

Physical performance is influenced by various factors, including an individual's physical characteristics and their performance on specific physical tests. Predicting an individual's performance class, based on these factors can provide valuable insights into their overall fitness level and help guide personalized training programs. This project aims to develop a predictive model that accurately determines an individual's performance class based on their physical characteristics and performance on physical tests.

Objectives of the study.

a) To develop a predictive model that accurately determines an individual's performance class based on their physical characteristics and performance on physical tests.
b) To identify factors that affect the body performance of a person & identify the most influential factors on determining the performance data.

About Data .

The Data set is including 13393 observations with 12 variables which include 2 Qualitative variables and 10 Quantitative variables as below,

Variable	Description	Type
age	20 ~ 64 (year)	Quantitative
gender	F (Female) & M (Male)	Qualitative
body fat_%	Body fat in percentage	Quantitative
height_cm	Height in centimeters	Quantitative
weight_kg	Weight in kilograms	Quantitative
diastolic	Blood pressure, the bottom number, measures the force your heart exerts on the walls of your arteries in between beats	Quantitative
systolic	Blood pressure, the top number, measures the force your heart exerts on the walls of your arteries each time it beats	Quantitative
gripForce	Person grip strength in kilogram	Quantitative
sit and bend forward_cm	Forward bend measured in centimeter(flexibility)	Quantitative
sit-ups counts	Sit up in 1 repetition	Quantitative
Broad jump_cm	High jump measured in centimeter	Quantitative
class	Performance score A,B,C,D (A best) / stratified(Target Variable)	Qualitative

Significance of the study.

This study holds significant importance in various domains as below,

a) The predictive model can be applied in real-world settings for personalized training program recommendations, health assessments, and performance monitoring, aiding individuals in achieving their fitness goals.

b) The study's findings can contribute to the broader research field by expanding our understanding of the interactions between physical characteristics, performance on physical tests, and performance class. It can inspire further investigations and studies in related areas.

c) The comprehensive body performance dataset created through this study can serve as a valuable resource for researchers, healthcare professionals, and fitness enthusiasts. It can facilitate collaborations, knowledge sharing, and further studies on performance prediction and related fields.

Suggest methodology.

Methodology for studying steps can suggest as below,

1. Data Preprocessing part,
    - Clean the dataset by addressing missing values, outliers, and inconsistencies.
    - Implement appropriate techniques to handle missing data and outliers, ensuring data quality.
2. Feature Selection part,
    - Select the most relevant features from the dataset that are likely to influence an individual's performance class by using feature selection methods and consider factors such as statistical significance, correlation with performance class, and expert knowledge.
3. Model Development part,
    - Develop a predictive model using machine learning and logistic regression algorithms suitable for the task and Train the model using the selected features and the available data.

References.

Body Performance Data (Data Set)
https://www.kaggle.com/datasets/kukuroo3/body-performance-data