Fitness Bootcamp Body Performance Data Model



At **Fitness Bootcamp**, we are a **healthcare**-focused gym committed to helping our members achieve their fitness goals. As part of our ongoing effort to enhance member outcomes, we are introducing a new personalized training program. Before selecting the appropriate number of personal training sessions, we utilize a sophisticated classification system to group members into four categories based on key physiological measurements: body fat percentage, weight and blood pressure.

**Our Machine Learning Classification Model:**We aim to develop a supervised machine learning model to predict member classifications into four distinct categories (A, B, C, D) based on body performance data. These categories are stratified, with Class A representing the highest fitness level and Class D representing the lowest.

The body performance data analyzes key measurements such as body fat percentage, diastolic blood pressure, and systolic blood pressure. Once a member is classified into a category, **Fitness Bootcamp** recommends a tailored number of personal training sessions to ensure that the program aligns with the member's specific health and fitness needs. The gym’s goal is to determine the optimal number of sessions that can be recommended for everyone, based on their classification, to maximize their fitness outcomes.

**Input Features:**

* **Age:** 20 to 64 years
* **Gender:** F (Female), M (Male)
* **Height:** Height in centimeters
* **Weight:** Weight in kilograms
* **Body Fat Percentage:** Percentage of body fat
* **Diastolic Blood Pressure:** Minimum diastolic blood pressure (in mmHg)
* **Systolic Blood Pressure:** Minimum systolic blood pressure (in mmHg)

**Output Classifications:**

* **Class:** A, B, C, D (where Class A represents the highest level of fitness and Class D represents the lowest).

Using this dataset, our goal is to train a model to predict the appropriate fitness classification (A, B, C, or D) for each member based on their input features.

**Source of Dataset:**

<https://www.kaggle.com/datasets/kukuroo3/body-performance-data?resource=download>

**Data Model Implementation and Optimization:**

Python Script model performance

Model utilizes data from Spark

Google colab implémentation

Google colab Import Dependencies: sklearn.model\_selection, train\_test\_split, sklearn.preprocessing, StandardScaler, sklearn.metrics, accuracy\_score, pandas as pd, tensorflow

Model utilizes data from Spark

**Tasks:**

* **Derilee Walters:** cleaning, machine learning model
* **Steven Quintero:** machine learning model
* **Kim Aissa:** cleaning, machine learning model
* **Liseth Rubio:** presentation, machine learning model