Major Project

on

FITNESS TRACKER

2024 - 2025

Guided by,

Urooj Khan
Mentor – Data Science
Meta Scifor Technologies

Submitted by,

Mrudula A P
Data Science Intern
Meta Scifor Technologies

Abstract

In today's fast-paced society, maintaining a healthy lifestyle amidst busy schedules and sedentary habits is increasingly challenging. Fitness tracker applications, coupled with wearable devices like Fitbit, offer a promising solution by providing convenient monitoring of physical activity, sleep patterns, and overall health metrics. However, existing solutions often lack personalized insights and recommendations tailored to individual needs, limiting their effectiveness.

To address these challenges, this project proposes the development of a fitness tracker application that leverages machine learning techniques to provide personalized recommendations and predictions based on user data collected from Fitbit devices. The project aims to fill the existing gaps in the market by delivering a solution that offers actionable insights and guidance to improve fitness outcomes.

The project's objectives include collecting and preprocessing user data, developing machine learning models for prediction and recommendation, implementing the fitness tracker application with seamless integration with Fitbit devices, and evaluating its effectiveness through rigorous testing and analysis.

A structured methodology guides the project, encompassing data collection, preprocessing, model development, implementation, result analysis, and iterative refinement. By following this approach, the project aims to deliver a user-centric solution that promotes healthier lifestyles and improved well-being.

Through clear articulation of the problem statement, defined objectives, and systematic methodology, this project seeks to make a significant contribution to the field of fitness tracking applications, ultimately empowering users to achieve their fitness goals more effectively.

Introduction:

In today's society, where technology plays a significant role in our lives, health and fitness have become increasingly important. However, maintaining a healthy lifestyle amidst busy schedules and sedentary habits can be challenging. This is where fitness tracker applications come into play. These apps, coupled with wearable devices like Fitbit, offer a convenient way for individuals to monitor their physical activity, sleep patterns, and overall health metrics. However, traditional fitness tracking methods often lack personalized insights and recommendations, limiting their effectiveness.

By leveraging machine learning techniques, this application aims to provide personalized recommendations and predictions based on user data collected from Fitbit devices. The integration of machine learning not only enhances the accuracy of predictions but also enables the application to adapt and provide tailored guidance to each user's unique fitness journey.

Problem Statement:

Traditional methods often provide generic recommendations and lack personalized insights tailored to individual needs. Users may struggle to interpret the data collected by their Fitbit devices and translate it into actionable steps to improve their fitness levels. Moreover, there is a lack of comprehensive solutions that intelligently analyse user data to provide personalized guidance and support.

The project aims to address these challenges by developing a fitness tracker application that fills the existing gaps in the market. The application will utilize machine learning algorithms to analyse user data and provide personalized recommendations, thereby enhancing user engagement and improving fitness outcomes.

Objective:

The objective of this project is to develop a comprehensive fitness app that provides users with personalized solutions for achieving their health and wellness goals. This objective is divided into three main components:

- Firstly, we aim to develop a predictive model to estimate the number of steps taken by Fitbit users accurately. This model will leverage machine learning algorithms and incorporate demographic and lifestyle factors such as age, gender, weight, height, BMI, and sleep duration, along with historical activity data. By analysing these variables, the model will forecast the number of steps a user is likely to take within a given time frame, providing valuable insights into activity levels and enabling users to set realistic fitness goals.

- Secondly, we seek to explore the impact of demographic and lifestyle factors on physical activity levels among Fitbit users. Through detailed analysis and visualization techniques, we will uncover trends and correlations between various variables and users' activity patterns. Understanding how factors like age, gender, and sleep habits influence activity levels will enable us to provide personalized recommendations tailored to individual profiles and preferences, thereby promoting more effective engagement and behaviour change.

- Lastly, our objective is to provide personalized exercise recommendations aligned with users' fitness goals and preferences. Leveraging insights from the predictive model and demographic/lifestyle analysis, we will develop algorithms to generate customized workout plans, exercise routines, and activity goals for each user. These recommendations will consider factors such as muscle building, fat loss, or maintenance of tone, ensuring relevance and achievability for each individual. By offering tailored guidance and support, we aim to empower users to make informed decisions about their health and fitness journey and achieve better outcomes in terms of overall well-being and lifestyle satisfaction.

Methodology:

Data Collection: The dataset used in this project was obtained from a distributed survey conducted on Amazon Mechanical Turk. Thirty eligible Fitbit users consented to submit personal tracker data spanning from 03.12.2016 to 05.12.2016. The dataset comprises minute-level output for physical activity, heart rate, sleep monitoring, and other relevant metrics. Each user's data is organized into multiple files, such as dailyActivity, dailyCalories, hourlySteps, etc.

Data Preprocessing: Before conducting analysis, the Fitbit dataset undergoes preprocessing to ensure its quality and consistency. This involves several steps, including handling missing values, removing duplicates, and standardizing data formats. Additionally, categorical variables are encoded into numerical format using techniques such as one-hot encoding, and numerical features are scaled to ensure uniformity in their magnitudes. Preprocessing techniques are crucial for preparing the data for subsequent analysis tasks.

Exploratory Data Analysis (EDA): EDA is performed to gain insights into the dataset's structure, characteristics, and patterns. Descriptive statistics, data visualization techniques, and correlation analysis are employed to explore relationships between variables and identify any outliers or anomalies. EDA helps in understanding the distribution of data, detecting patterns, and informing subsequent analysis steps.

Feature Engineering: Feature engineering is carried out to improve the predictive model's performance by selecting, transforming, and creating new features from the existing dataset. Feature selection techniques are employed to identify the most relevant predictors for estimating step counts, while feature transformation and creation aim to enhance the model's predictive power. This iterative process involves experimentation and refinement based on the model's performance.

Model Development: Machine learning algorithms, particularly regression models, are utilized to develop the predictive model for estimating the number of steps taken by Fitbit users. The dataset is split into training and testing sets to evaluate the model's performance. Various regression techniques, including linear regression, decision trees, and ensemble methods like XGBoost, are explored and compared based on evaluation metrics such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE).

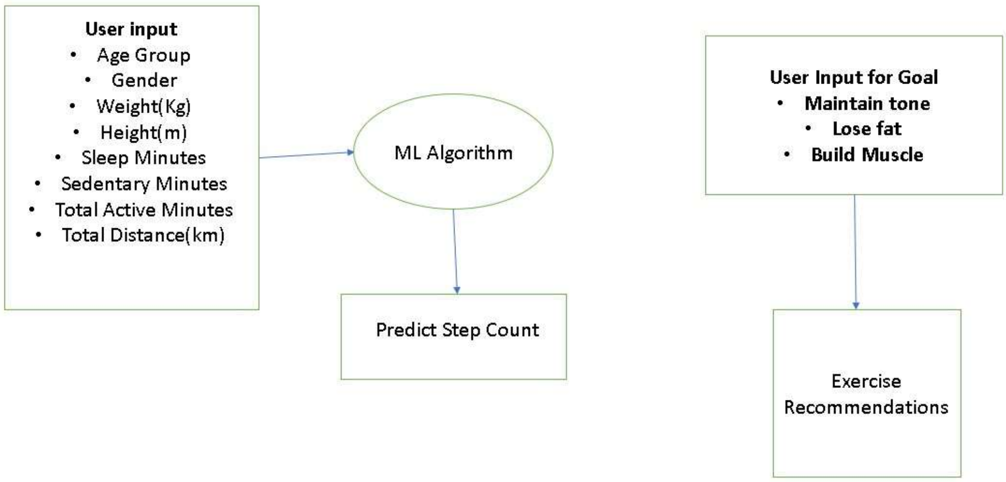
Model Evaluation and Validation: The developed model is evaluated using appropriate metrics to assess its accuracy and generalization performance. Cross-validation techniques are employed to ensure the model's robustness and reliability across different subsets of the data. Model validation is crucial to ensure that the model performs well on unseen data and can effectively generalize to new observations.

Personalized Exercise Recommendations: Based on the insights gained from the predictive model and demographic/lifestyle analysis, personalized exercise recommendations are generated for Fitbit users. Algorithms are developed to tailor workout plans, exercise routines, and activity goals to each user's fitness goals and preferences. Recommendations are customized based on factors such as muscle building, fat loss, or maintenance of tone, ensuring relevance and effectiveness for each individual's fitness journey.

Model implementation:

To implement the step count prediction model using XGBoost, we'll follow these steps:

1. **Data Preparation:**
    - Gather data from Fitbit devices, including user inputs such as age group, gender, weight, height, sleep minutes, sedentary minutes, total active minutes, and total distance covered.
    - Cleanse and preprocess the data, handling missing values, outliers, and ensuring consistency.
2. **Feature Engineering:**
    - Extract relevant features from the collected data, such as BMI (Body Mass Index) from weight and height, and any other derived features that may enhance model performance.
3. **Model Training:**
    - Utilize the XGBoost algorithm to train a regression model on the prepared dataset.
    - Split the dataset into training and testing sets to evaluate model performance.

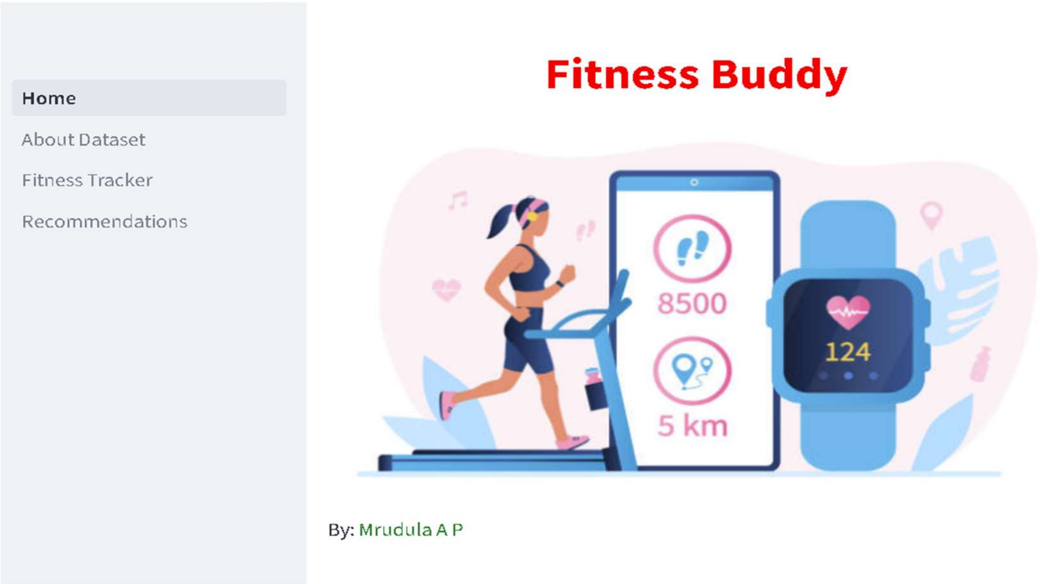


4. **Model Evaluation:**
    - Evaluate the trained XGBoost model using appropriate metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R2) score.
    - Validate the model's performance on the testing dataset to ensure its generalization ability.

	R2	MSE	MAE
XGBoost	0.99999	124.33883	1.96247
Random Forest	0.99999	165.64144	3.60635
Linear Regression	0.95500	880624.17200	617.27500

5. **Integration with Streamlit:**
    - Develop a user interface using Streamlit to collect user inputs such as age group, gender, weight, height, sleep minutes, sedentary minutes, total active minutes, and total distance covered.
    - Pass these inputs to the trained XGBoost model to predict the step count for the user.
6. **Frontend Design:**
    - Design an intuitive and user-friendly frontend interface using Streamlit, ensuring smooth navigation and a visually appealing layout.
    - Include features for input validation and error handling to enhance user experience.
7. **Deployment:**
    - Deploy the Streamlit application along with the trained XGBoost model on a server or cloud platform to make it accessible to users.
    - Ensure scalability and reliability of the deployed application to handle potential user traffic.

The user-interface of Fitness tracker application is given below.



8. **Testing and Iteration:**
   - Conduct thorough testing of the integrated system to identify any bugs or issues.
   - Gather feedback from users to understand their experience and make necessary iterations to improve the application's functionality and usability.

By following these steps, we could implement a robust system that leverages XGBoost for step count prediction based on various user inputs collected from Fitbit devices. The integration with Streamlit provides an interactive and user-friendly interface, enhancing the overall user experience.

Conclusion:

To conclude, the project aims to develop a fitness tracker application that leverages machine learning to provide personalized recommendations based on user data from Fitbit devices. By addressing the identified challenges and following a structured methodology, the project seeks to deliver a robust solution that empowers users to achieve their fitness goals effectively. Ultimately, the goal is to promote healthier lifestyles and improved well-being through technology-enabled fitness tracking.
