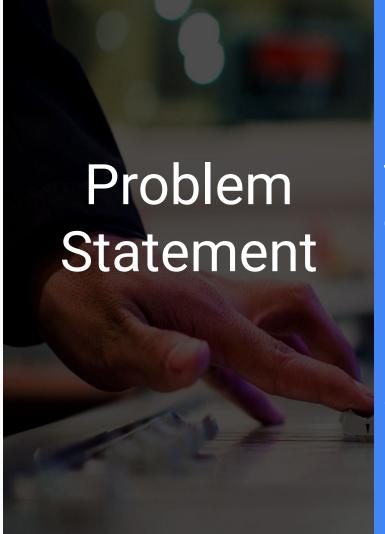# CAPSTONE PROJECT

Predicting Weight Loss

Ratan Yalamanchili
Sep 2025

Problem Statement

To build ML models to predict the calories of an individual, whether a person would lose weight or not if we use features like exercise types, food intake, water consumption, supplement , dieting etc

Applications

Linear Regression (LR)

LR is a way to estimate the relationship between two variables. This can make predictions about future values of the variables (i.e., Calories loss) The application uses variables like Total Distance and No of Steps taken

Algorithm: GradientBoostingRegressor

Training Score	85%
Test Score	61%

This algorithm can predict calories loss with an accuracy of over 60%+

Decision Tree Classifier (DTC)

DTC is a ML model that can be used to classify data. It does this by building tree of decisions based on the value of a variable (I.e., no of steps etc.).

The tree is then used to classify new data like weight loss flag

Algorithm: Decision Tree Classifier

Training Score	62%
Test Score	63%

This algorithm can predict calories loss with an accuracy of over 60%+

ML Classifications

ML Classifications uses a process of dividing data into groups based on shared characteristics.

The goal of classification is to identify patterns (i.e., no of steps taken etc.) in the data and use those patterns to predict the class of new data points (i.e., weight loss)

Algorithm: K-Nearest Neighbor (KNN)

Training Score	75%
Test Score	59%

This algorithm can predict calories loss with an accuracy of over 60%+

Outcome

Overview:

- The project includes three notebooks demonstrating how various Machine Learning (ML) models can predict calories burned and weight loss potential.

- Data was collected from 33 Fitbit users over one month.

- Model training accuracy ranged from 65% to 100%, depending on the algorithm used.

Key Findings:

- ML Application 1:

  - Tested multiple models using all variables and key inputs (Total Steps and Total Distance).

  - Training accuracy: 62%–100% | Test accuracy: 37%–67%

  - Most influential variables: Total Steps and Total Distance

- ML Application 2:

  - Models achieved an average accuracy around 60%.

- ML Application 3:

  - K-Nearest Neighbor (KNN) performed best: Training accuracy: 75% | Test accuracy: 58%

- Each notebook includes a Findings section detailing model performance.

Recommendation

Recommended Model: GradientBoostingRegressor

- **Training accuracy**: 86% | **Testing accuracy:** 61%

- Strong candidate for predicting calories burned and weight loss potential

- Performance can be further improved with more diverse and high-activity data (e.g., users with >30k steps/day and >5 miles distance)

Data Collection for Model Enhancement:

- **Lifestyle Metrics**: Daily water intake, sleep patterns, and calorie intake (Protein, Carbs, Fats)

- **User Demographics**: Age, Gender, Weight, Height, BMI

- **Activity Metrics**: Heart rate, active minutes, and step intensity

- **Including these features will help the model capture personalized patterns, improving prediction accuracy**

This ML model would be deployed as an AI application and Weight loss application can use the model for users to determine what they need to do to lose weight or estimate the amount of Calories loss based on those features