NutriTrack - AI-Based Health Wellness System

ON

Submitted in partial fulfillment of the requirements of the degree of

Bachelor of Engineering

(Information Technology)

By

Brijesh Sharma (48)

Alok Yadav (59)

Sandesh Yadav (61)

Under the guidance of

Dr. Ravita Mishra

Department of Information Technology

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY,

Chembur, Mumbai 400074

(An Autonomous Institute, Affiliated to University of Mumbai) April 2024

AIDS Lab Exp 11
Aim: Mini Project – NutriTrack - AI-Based Health Wellness System

---

1.1 Introduction

In today's fast-paced lifestyle, individuals often struggle to maintain a healthy diet and fitness regime tailored to their specific body requirements and goals. Generic dietary recommendations often fall short in catering to personalized needs, which may vary based on age, gender, activity levels, and health conditions.

2.2 Design and Implementation

The design and implementation of the personalized diet and nutrition recommendation system are centered around creating an intuitive, responsive, and intelligent platform that seamlessly integrates a user-friendly frontend with a robust backend powered by machine learning models. The system is developed using React for the frontend and FastAPI for the backend to ensure efficient, scalable, and real-time communication between the user interface and the machine learning recommendation engine.

2.3 Feature Scaling

RFM values were standardized using StandardScaler to ensure equal contribution to clustering models.

2.4 Evaluation Analysis

KNN (with Cosine Similarity):

KNN recommends recipes by comparing the target nutritional vector (e.g., Calories, Protein) with each recipe's 9 scaled nutritional features using cosine similarity, after applying ingredient/allergy filters. Chosen for its simplicity, no training requirement, compatibility with filtering, and baseline performance (MSE: 51095.76), it's ideal for direct nutritional matching on small datasets.

Weighted Euclidean Distance:

A KNN variant using a weighted distance metric to emphasize key nutrients (e.g., Calories weight=2.0), improving control over deviations (Calories deviation reduced from 76.07% to 56.38%). Achieved better MSE (38406.71) than cosine, supports user-driven nutrient prioritization, and integrates with existing filters.

K-Means Clustering:

Clusters recipes into k groups (e.g., 50) based on nutritional features. For a target input, the closest cluster is selected, and top recipes are recommended. It improves diversity, efficiency, and handles cases with few exact matches, complementing KNN with pattern-based grouping.

Support Vector Regressor (SVR):

Trains 9 regressors (one per nutrient) using TF-IDF ingredients and time features to predict nutritional values. Recommendations are based on distance between predicted profiles and the target. Chosen for modeling ingredient-nutrient relationships, handling high-dimensional data, and reducing nutrient deviations with supervised learning.

Chapter 4: Results and Discussion

4.1 Dataset Description

The recipes dataset contains 522,517 recipes from 312 different categories. This dataset provides information about each recipe like cooking times, servings, ingredients, nutrition, instructions, and more.

RecipeId, Name, CookTime, PrepTime, TotalTime, RecipeIngredientParts, Calories, FatContent, SaturatedFatContent, CholesterolContent, SodiumContent, CarbohydrateContent, FiberContent, SugarContent, ProteinContent, RecipeInstructions

4.1 KNN:

KNN, a simple supervised algorithm, predicts based on the average of K similar data points. With MSE of 30517.99 and MAE of 101.99, it offers balanced nutrient deviations (e.g., Calories 24.01%), making it the most consistent and reliable model in this setup.

```
Evaluation Metrics for Breakfast:
MSE: 30517.99
MAE: 101.99
Deviation Percentages (%):
  Calories: 24.01%
  FatContent: 30.20%
  SaturatedFatContent: 30.50%
  CholesterolContent: 49.94%
  SodiumContent: 21.27%
  CarbohydrateContent: 66.75%
  FiberContent: 53.67%
  SugarContent: 46.50%
  ProteinContent: 54.74%
```

4.2 KMeans Clustering

K-Means clusters recipes by nutritional similarity and showed lower MSE (19047.16) and MAE (79.63). However, higher deviations (e.g., Protein 58.04%) and its unsupervised nature limit its use for precise predictions, making KNN a better choice overall.

```
Evaluation Metrics for Breakfast:
MSE: 19047.16
MAE: 79.63
Deviation Percentages (%):
  Calories: 31.24%
  FatContent: 19.13%
  SaturatedFatContent: 20.33%
  CholesterolContent: 45.38%
  SodiumContent: 8.42%
  CarbohydrateContent: 52.84%
  FiberContent: 24.80%
  SugarContent: 46.70%
  ProteinContent: 58.04%
```

4.3 KNN (Weighted Euclidean Distance)

This KNN variant prioritizes nutrients via weights but had higher MSE (38406.71), MAE (108.02), and poor calorie accuracy (76.07% deviation). Despite benefits like customizable importance, it lacked the precision for effective recommendations.

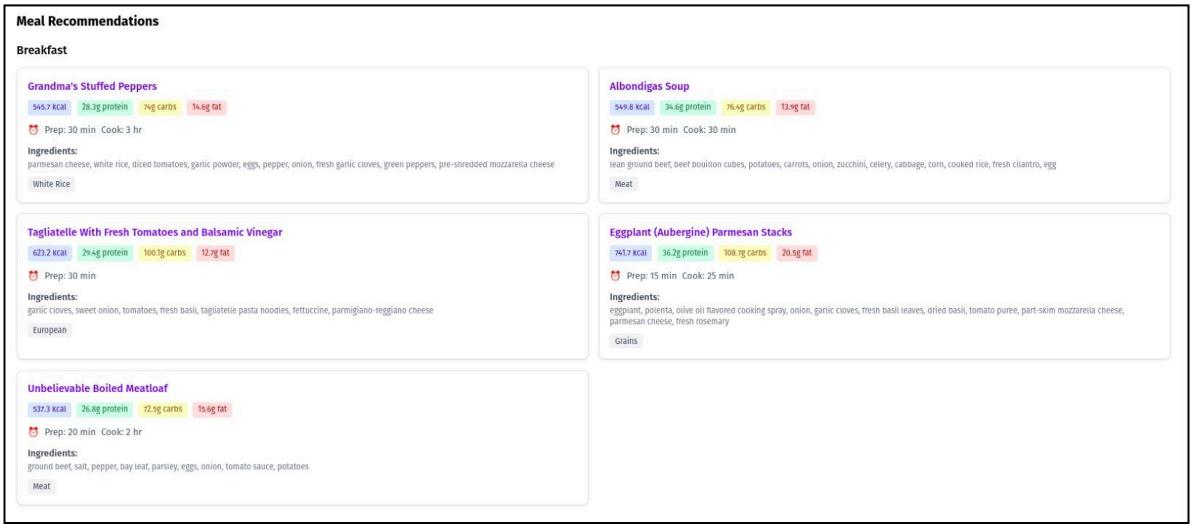
```
Evaluation Metrics for Breakfast:
MSE: 22125.59
MAE: 85.94
Deviation Percentages (%):
  Calories: 38.06%
  FatContent: 16.47%
  SaturatedFatContent: 22.83%
  CholesterolContent: 41.70%
  SodiumContent: 10.53%
  CarbohydrateContent: 47.70%
  FiberContent: 34.13%
  SugarContent: 41.90%
  ProteinContent: 55.12%
```

4.4 Support Vector Regression

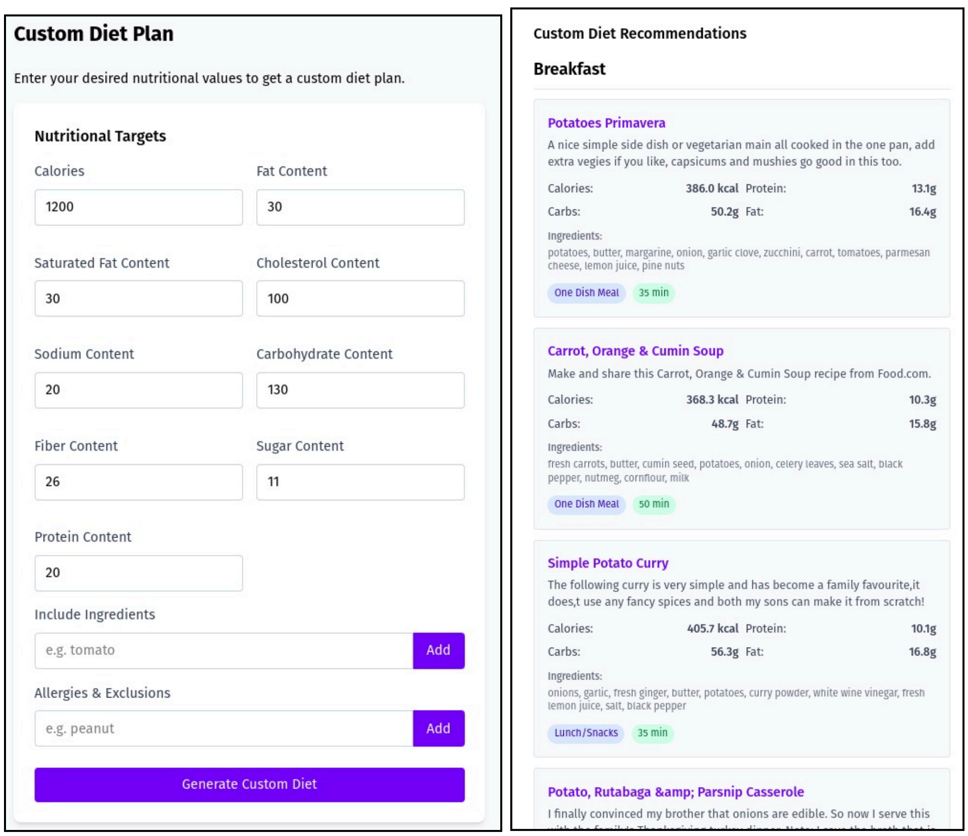
SVR, though powerful, had the worst performance (MSE: 174111.09, MAE: 218.83) with very high nutrient deviations (e.g., Protein 88.88%). Its sensitivity to noisy data led to poor generalization, making it unsuitable for this task.

```
Evaluation Metrics for Breakfast:
MSE: 174111.09
MAE: 218.83
Deviation Percentages (%):
  Calories: 49.27%
  FatContent: 34.47%
  SaturatedFatContent: 16.17%
  CholesterolContent: 43.15%
  SodiumContent: 73.07%
  CarbohydrateContent: 85.81%
  FiberContent: 92.27%
  SugarContent: 40.20%
  ProteinContent: 88.88%
```

4.6 Streamlit App Output



Meal recommendations based on user data



Custom diet recommendations based on user requirements

Custom Diet goals based on user intake

Conclusion

This project explored multiple models for recipe recommendation based on nutritional goals, comparing both supervised and unsupervised approaches. Among all, KNN emerged as the most reliable due to its consistent accuracy, balanced deviations, and simplicity. While K-Means showed promising MAE, its lack of predictive depth limits its utility. Overall, KNN offers the best balance of precision, interpretability, and adaptability for personalized nutritional recommendations.