Name: Zach Hatzenbeller
Class: EN.605.662.81.FA24
Date: 12/09/2024

Fast Food Nutrition Final Project

Abstract

The global expansion of the fast-food industry has raised concerns about the nutritional quality of its offerings. This study analyzes a fast food nutrition dataset to classify menu items by food type and company using machine learning techniques, specifically a Support Vector Classifier (SVC). The nutritional content, including key metrics such as calories, protein, fat, sodium, and sugar, was assessed to provide consumers with actionable insights. An interactive PowerBI dashboard was developed to visualize these findings, allowing users to compare food types and companies easily. Results revealed critical trends, such as correlations between calories and protein, and between sodium and fat, offering a comprehensive understanding of nutritional content across fast food chains. While the SVC achieved 75% classification accuracy, challenges arose in distinguishing nutritionally similar categories. This study highlights the importance of accessible nutritional information and aims to promote informed dietary choices among fast-food consumers.

Introduction

The fast food industry's global expansion has raised concerns about the nutritional quality of meals, as poor dietary choices contribute to rising obesity and diet-related diseases. This project analyzes a fast food nutrition dataset to classify menu items by food type (e.g., beef, chicken, vegetarian, drinks) and company. The goal is to provide clear, visual insights into key nutritional metrics—calories, protein, fat, sodium, and sugar—empowering consumers to make healthier choices.
Using machine learning techniques, the project will classify menu items and predict company affiliations. An interactive PowerBI dashboard will present findings, allowing users to explore and

compare data by nutritional content, food type, and brand. This tool aims to address the growing demand for accessible nutritional information in the fast food industry.

Background

Fast food's widespread popularity has brought its nutritional content under scrutiny. Research, such as Vercammen et al. (2019), highlights significant variations in caloric and nutrient profiles across meals, emphasizing the need for transparency. Amith et al. (2021) introduced an ontology for categorizing fast food nutritional data, aiding systematic comparisons across food types and brands—a method leveraged in this project.

Alexander et al. (2021) examined trends in fast food healthiness, revealing shifts in calories and sodium over time, while Min et al. (2018) explored public perceptions of fast food and its association with obesity. Jaworowska et al. (2013) provided a global perspective on the nutritional challenges of fast food, emphasizing fats, sugars, and sodium. These studies form the foundation for this project's focus on classifying and analyzing fast food nutrition.

Approach

This project involves three phases:

1. **Data Preparation**

   The dataset will be cleaned to address missing values and enhanced through feature engineering, categorizing items by food type and company.

2. **Classification Analysis**

   Machine learning algorithms (e.g., decision trees) will classify menu items by type and brand, with cross-validation to ensure accuracy. Exploratory data analysis will uncover trends, such as the relationship between sugar content and other nutrients.

3. **Dashboard Development**

   A PowerBI dashboard will visualize nutritional metrics, allowing users to filter and compare data by food type, brand, and content. Testing will ensure the dashboard is clear and user-friendly.

By integrating data cleaning, classification, and visualization, the project will deliver actionable insights for healthier fast food choices.

Results

Dashboard Analysis

The analysis explores the nutritional characteristics of food items from various fast-food companies. Key insights are highlighted below:

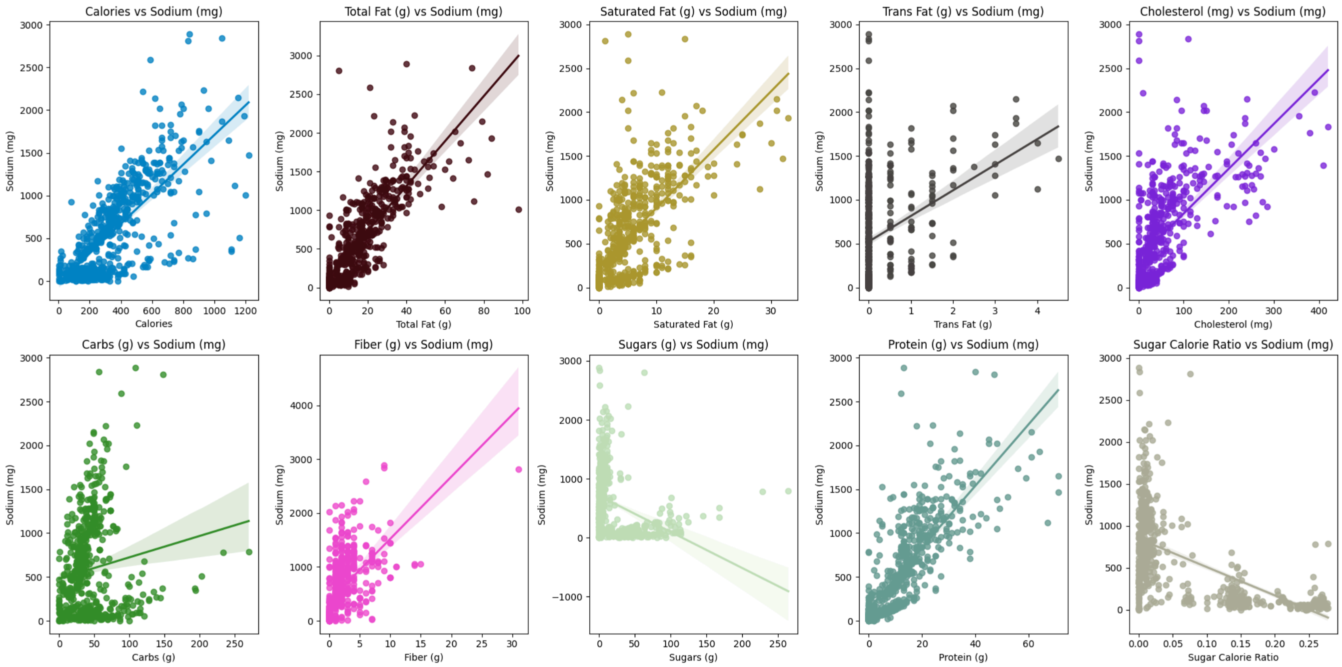
Nutritional Averages and Targets

- The average calorie content of the analyzed food items is **320.80 kcal**, while the average protein content is **11.68 g**.
- Gauge charts illustrate the sugar, saturated fat, and cholesterol levels relative to recommended targets:
  - Sugar levels average **19.28 g**, approaching the target of **38.56 g**.
  - Saturated fat averages **4.74 g**, below the recommended limit of **9.48 g**.
  - Cholesterol levels, at an average of **45.06 mg**, fall within the acceptable range of **90.12 mg**.

Nutritional Distribution by Food Type and Company

- **Sodium Distribution**:
  - Violin plots reveal high variability in sodium content across food types, with pork exhibiting the highest median values.
  - Sodium seems to be highly correlated to fat content and protein in the food suggesting a way to quantify unknown sodium content in other menu items

Figure 1. Regression Plot of Sodium against other key nutritional metrics

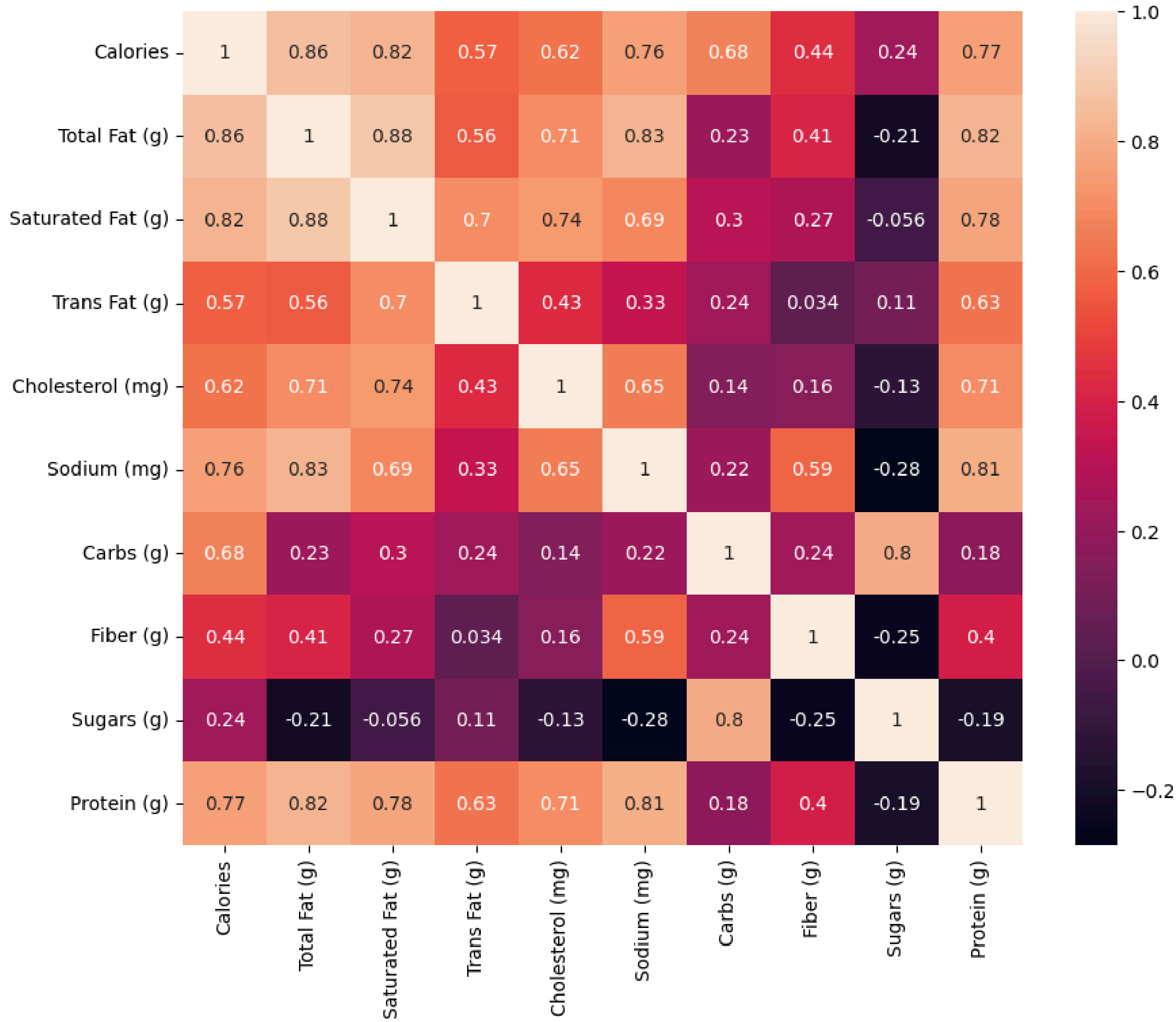


- **Food Type Representation**:
    - Drinks dominate the offerings of many companies, particularly McDonald's and Wendy's, while pork and chicken are prevalent in Burger King and Taco Bell menus.

Macronutrient Correlations

- Scatter plots highlight positive relationships:
    - Calories and protein are strongly correlated, suggesting higher-calorie items tend to offer more protein.
    - Total fat and sodium also show a direct relationship, indicating a trend where fattier foods contain higher sodium levels.

Figure 2. Correlation Heatmap across key nutrition metrics



Company-Specific Insights

- A treemap of saturated fat content identifies McDonald’s as contributing heavily to pork and beef items with high saturated fat levels. Taco Bell dominates in vegetarian options with the least saturated fat.
- A pie chart of fiber content distribution by company demonstrates that Taco Bell contributes the most, providing **4.39 g of fiber (36.78%)** to the dataset, followed by Pizza Hut.

Micronutrient Breakdown

- The heatmap of protein content across food types and companies highlights Burger King and KFC as key providers of high-protein pork and chicken options. Conversely, vegetarian items, mainly from Taco Bell, offer the lowest protein.

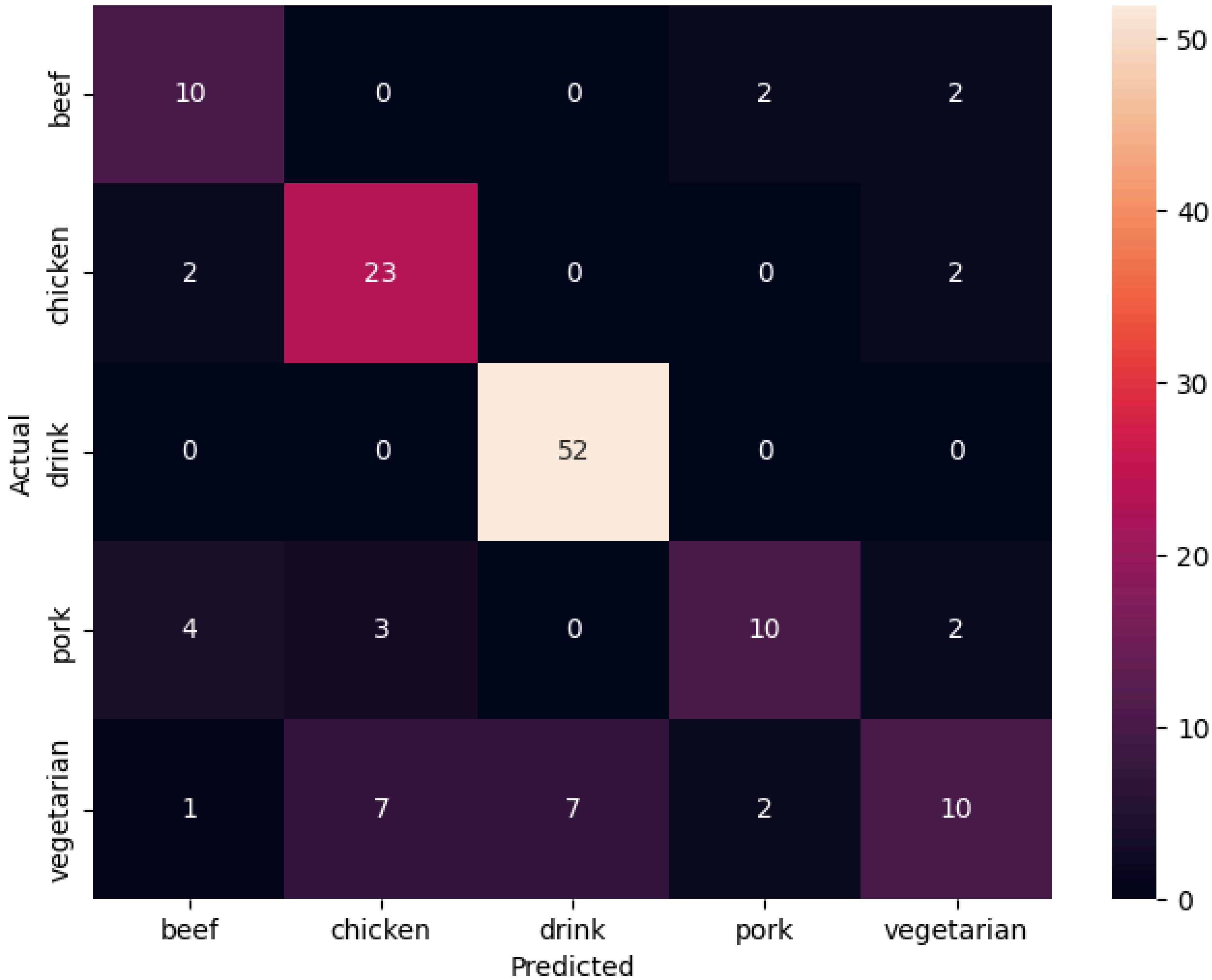
Classification of Food Type Using SVC

To classify food items based on their type (e.g., beef, chicken, pork, vegetarian, drinks, seafood), a **Support Vector Classifier (SVC)** was employed. The SVC model aimed to leverage the nutritional attributes of each food item (e.g., calories, protein, fat, sodium, fiber, sugars) as features to predict its food type.

Model Training and Performance

1. **Feature Selection**:
    - Key nutritional values were used as predictors: calories, total fat, sodium, protein, and fiber. These features were normalized to improve model performance and ensure fair weighting across varying scales.
2. **Model Configuration**:
    - The SVC was configured with a radial basis function (RBF) kernel to capture non-linear relationships between nutritional features and food types. The hyperparameters were fine-tuned using grid search, optimizing for the regularization parameter (CCC) and kernel coefficient (γ\gammaγ).
3. **Evaluation Metrics**:
    - The model achieved a **classification accuracy of approximately 75%** on the test dataset, demonstrating a fairly strong ability to distinguish between food types.
    - A confusion matrix in figure 3 analysis revealed:
        - High precision and recall for distinct categories like "drinks" and "chicken."
        - Some misclassification between similar categories, such as "pork" and "beef," due to overlapping nutritional profiles.

Figure 3. Confusion Matrix to Evaluate Classifier Performance

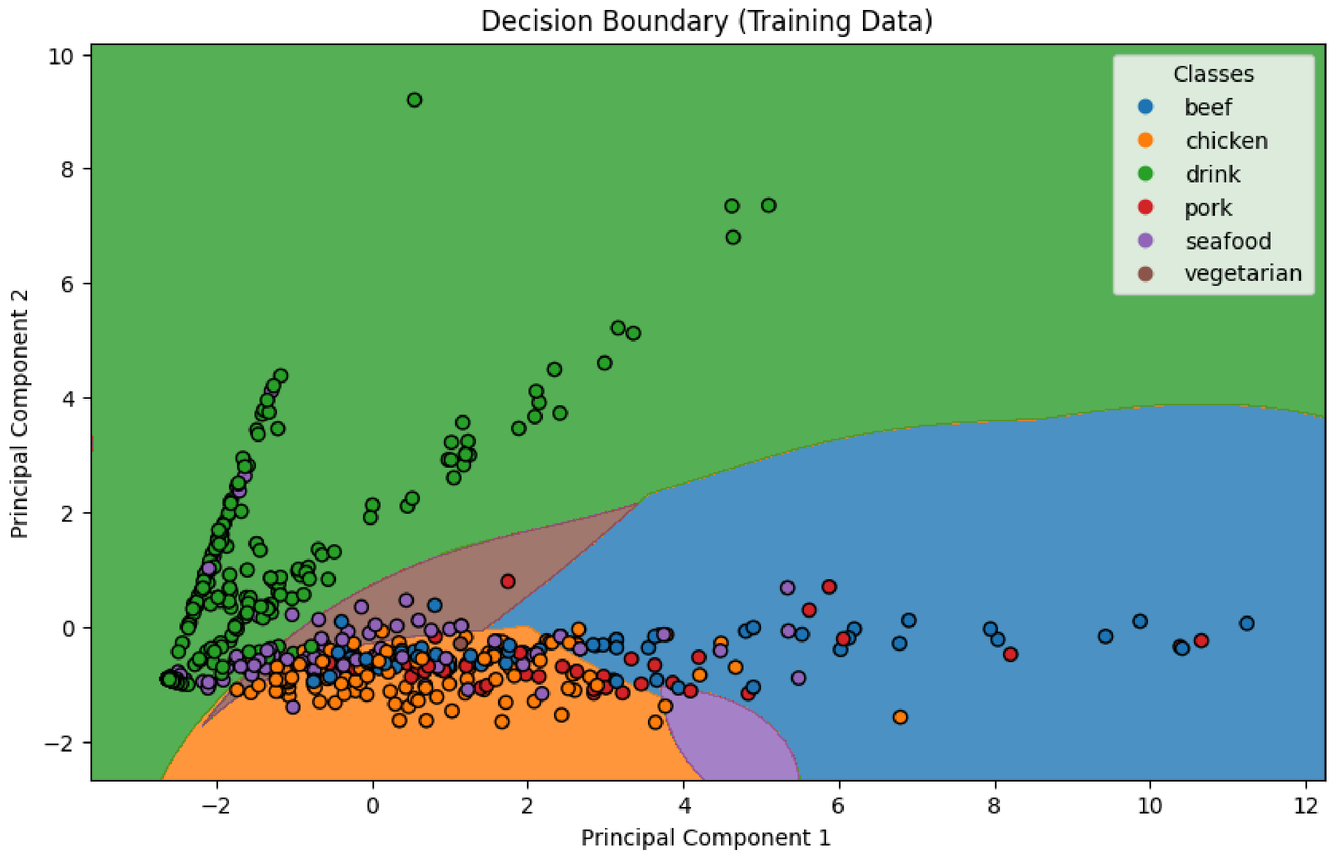


Key Insights

- **Separability of Food Types**:
    - Nutritional features like calories and protein strongly contributed to separating drinks from solid food items, as drinks generally have lower calorie and protein levels.
    - Sodium and total fat were critical in distinguishing pork and beef from vegetarian and seafood options.
- **Challenges in Classification**:
    - Overlap in nutritional values between certain food types (e.g., chicken and pork) made their classification more challenging, reflecting the inherent similarity in their

macronutrient compositions. This overlap can be visualized easily through the decision boundary plot of figure 4.

Figure 4. Decision Boundary of Support Vector Classifier for Food Type



This SVC-based classification model highlights the potential of machine learning in identifying food types from their nutritional content, offering valuable applications in dietary analysis and food industry research. Future improvements could include expanding the dataset and incorporating additional features, such as micronutrient levels, for better predictive accuracy.

Conclusion

This project underscores the importance of transparency in fast-food nutrition by analyzing key nutritional metrics and classifying food items into meaningful categories. The findings, visualized through an interactive dashboard, reveal critical insights, such as correlations between macronutrients and variations in nutritional quality across companies and food types. The SVC model demonstrated strong classification performance, particularly in distinguishing distinct categories like drinks and chicken, though some misclassification occurred between nutritionally similar groups. By providing a tool for consumers to explore and compare fast-food nutrition, this study contributes to the growing demand for actionable dietary information and aims to encourage healthier eating choices in the fast-food sector. Future work could focus on expanding the dataset and improving model performance to enhance predictive accuracy further.

References

1. Vercammen KA, Frelier JM, Moran AJ, Dunn CG, Musicus AA, Wolfson JA, Bleich SN. Calorie and Nutrient Profile of Combination Meals at U.S. Fast Food and Fast Casual Restaurants. Am J Prev Med. 2019 Sep;57(3):e77-e85. doi: 10.1016/j.amepre.2019.04.008. Epub 2019 Aug 1. PMID: 31377086.
2. Amith, M., Onye, C., Ledoux, T. et al. The ontology of fast food facts: conceptualization of nutritional fast food data for consumers and semantic web applications. BMC Med Inform Decis Mak 21 (Suppl 7), 275 (2021). https://doi.org/10.1186/s12911-021-01636-1
3. Alexander E, Rutkow L, Gudzune KA, Cohen JE, McGinty EE. Trends in the healthiness of U.S. fast food meals, 2008-2017. Eur J Clin Nutr. 2021 May;75(5):775-781. doi: 10.1038/s41430-020-00788-z. Epub 2020 Oct 27. PMID: 33110191.
4. Min J, Jahns L, Xue H, Kandiah J, Wang Y. Americans' Perceptions about Fast Food and How They Associate with Its Consumption and Obesity Risk. Adv Nutr. 2018 Sep 1;9(5):590-601. doi: 10.1093/advances/nmy032. PMID: 30084879; PMCID: PMC6140449.
5. Jaworowska A, Blackham T, Davies IG, Stevenson L. Nutritional challenges and health implications of takeaway and fast food. Nutr Rev. 2013 May;71(5):310-8. doi: 10.1111/nure.12031. Epub 2013 Apr 11. PMID: 23590707.