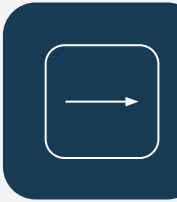


Predicting High Traffic Recipes

Tasty Bites

German Paul

24.11.23



German Paul

Head of Data Science

Tasty Bites



Problem Statement

- Classify high traffic recipes
- Reach an accuracy over 80%

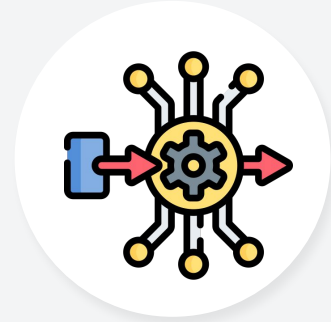
Steps to reach the goal



Data Validation



Data Visualization



Model creation

Data Validation

1. Data Validation

2. Data Formatting

3. Data Transformation



Data Validation

Check for unique values in every column



`df.column.nunique()`

Check for distribution



`df.column.value_counts()`



Data Formatting

4 as a snack



4

Nutrient NaN



Mean values



Data Transformation

x_data



Standardized x_data

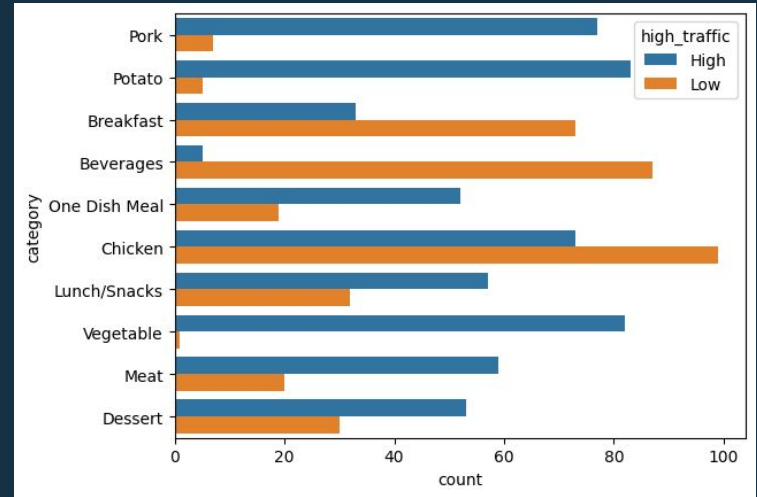
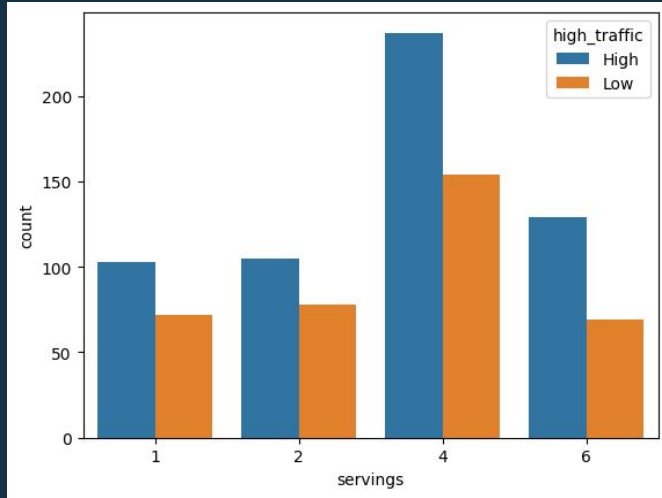
Categorical Category column



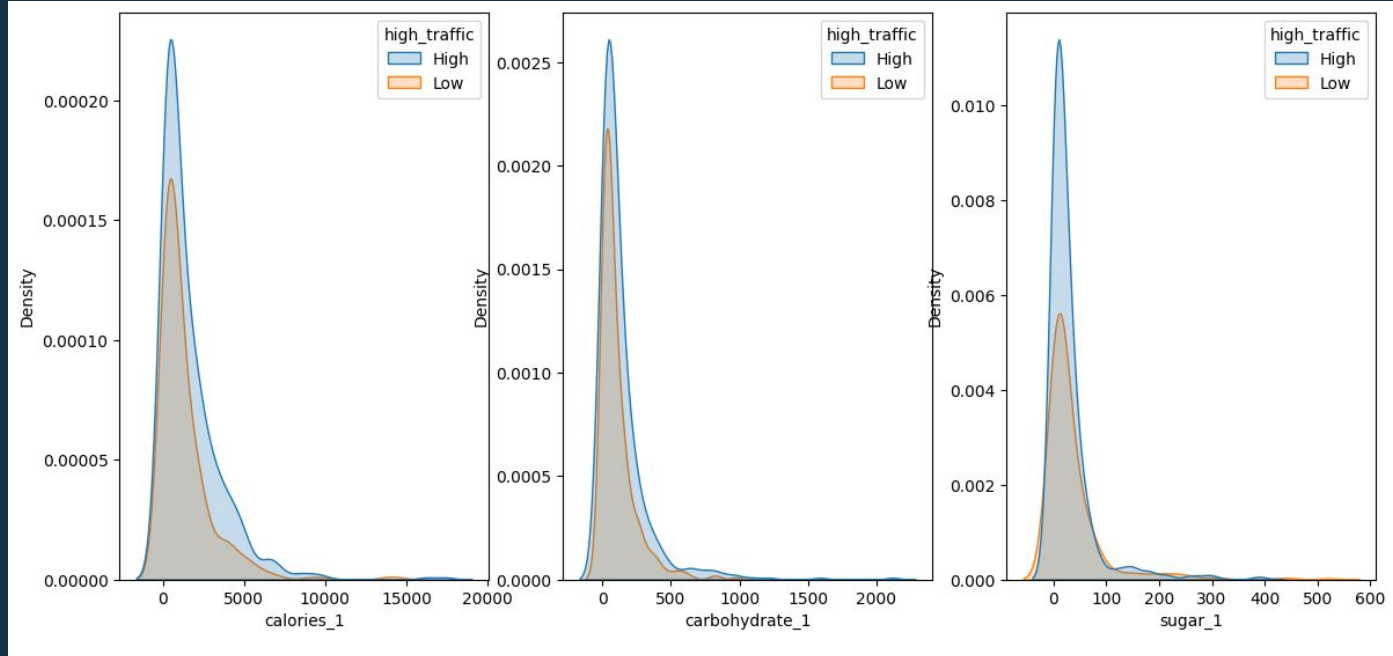
Label encoded category column



Data Visualization



Data Visualization



Model Creation

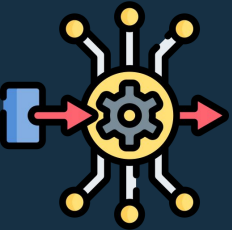
Logistic Regression

Accuracy of 76%

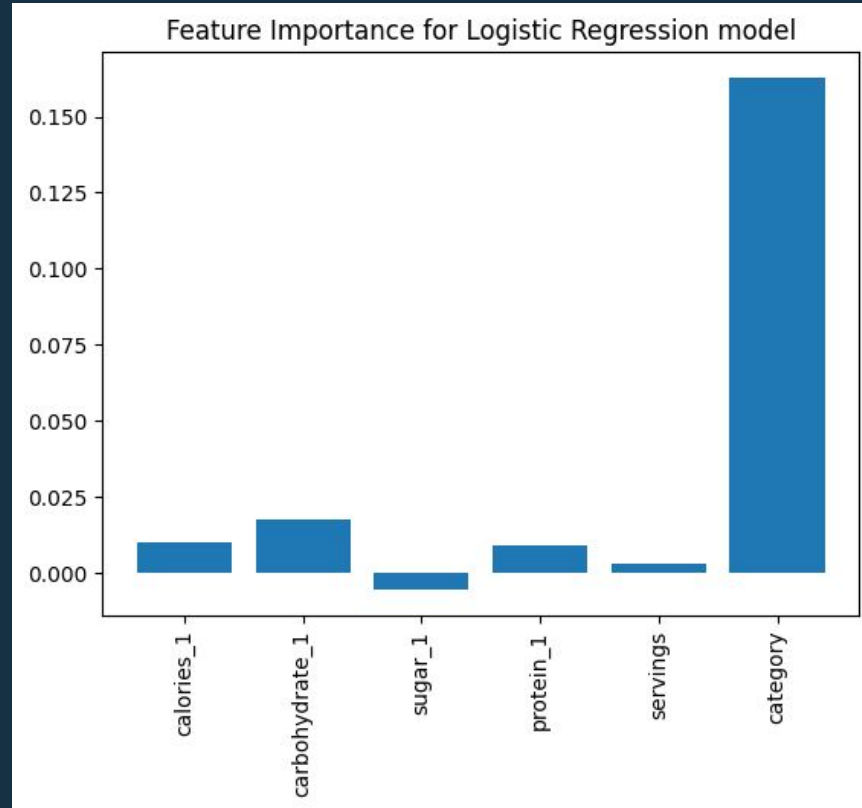
Linear SVC

Accuracy of 74%

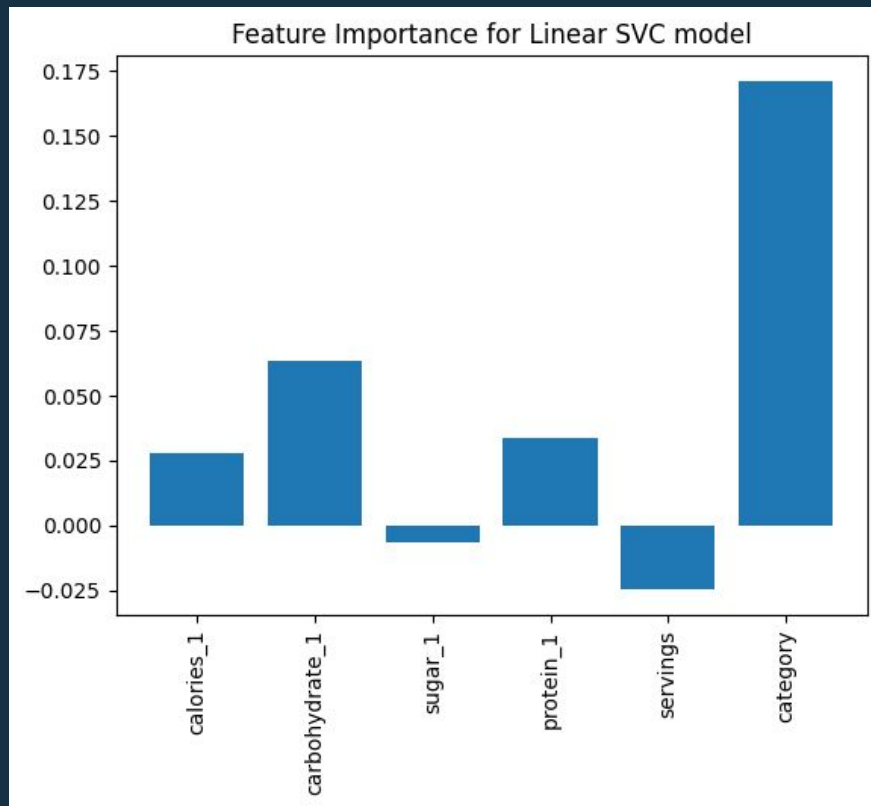
Use of GridSearchCV to
find best parameters



Logistic Regression



Linear SVM

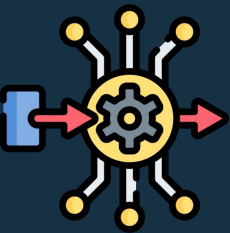


Improving the Model

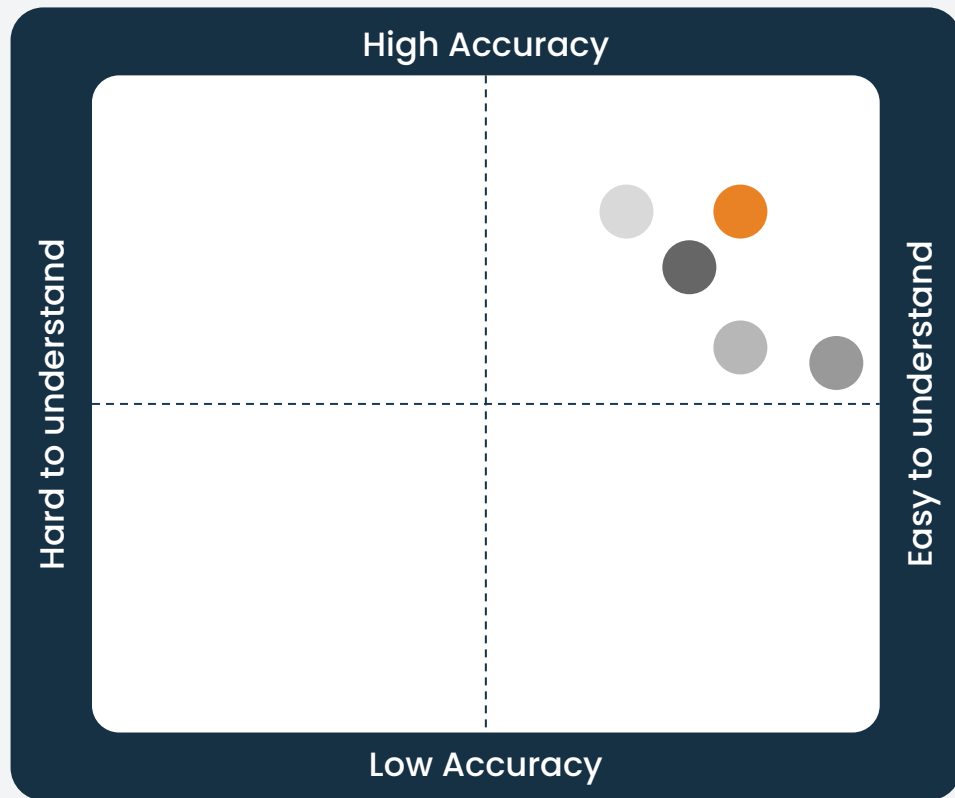


The model can only be as good as the data provided

Getting more data



MODEL ANALYSIS



Logistic Regression



Linear SVM



Always choosing High Traffic



Naive Bayes



XGBoost Classifier

Solution

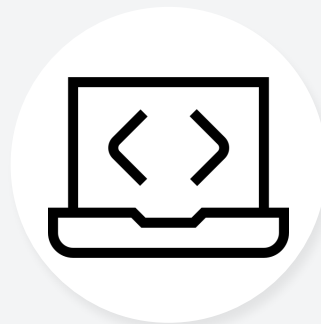
Tasty Bites should implement the Logistic Regression Model in production



Make customers happy



Open-Source



Efficient algorithm - easy
to retrain

RISK

- Cybersecurity breaches and data leaks.
- Economic downturns impacting IT budgets.
- Emerging competitors disrupting the market.

CHALLENGES

- Attracting and retaining top cybersecurity talent.
- Scaling operations to meet growing demand.
- Adapting to evolving technology and threat landscape.

