Recipe Site Traffic

Fawzy Almatary

Data Scientist

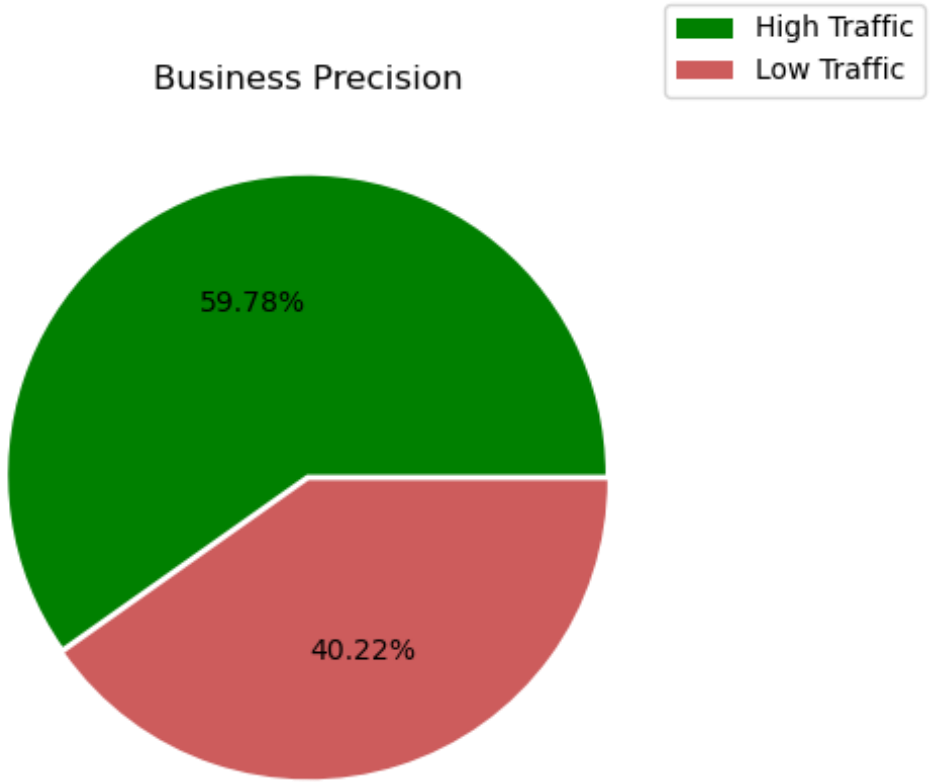
Background and Business Goals:

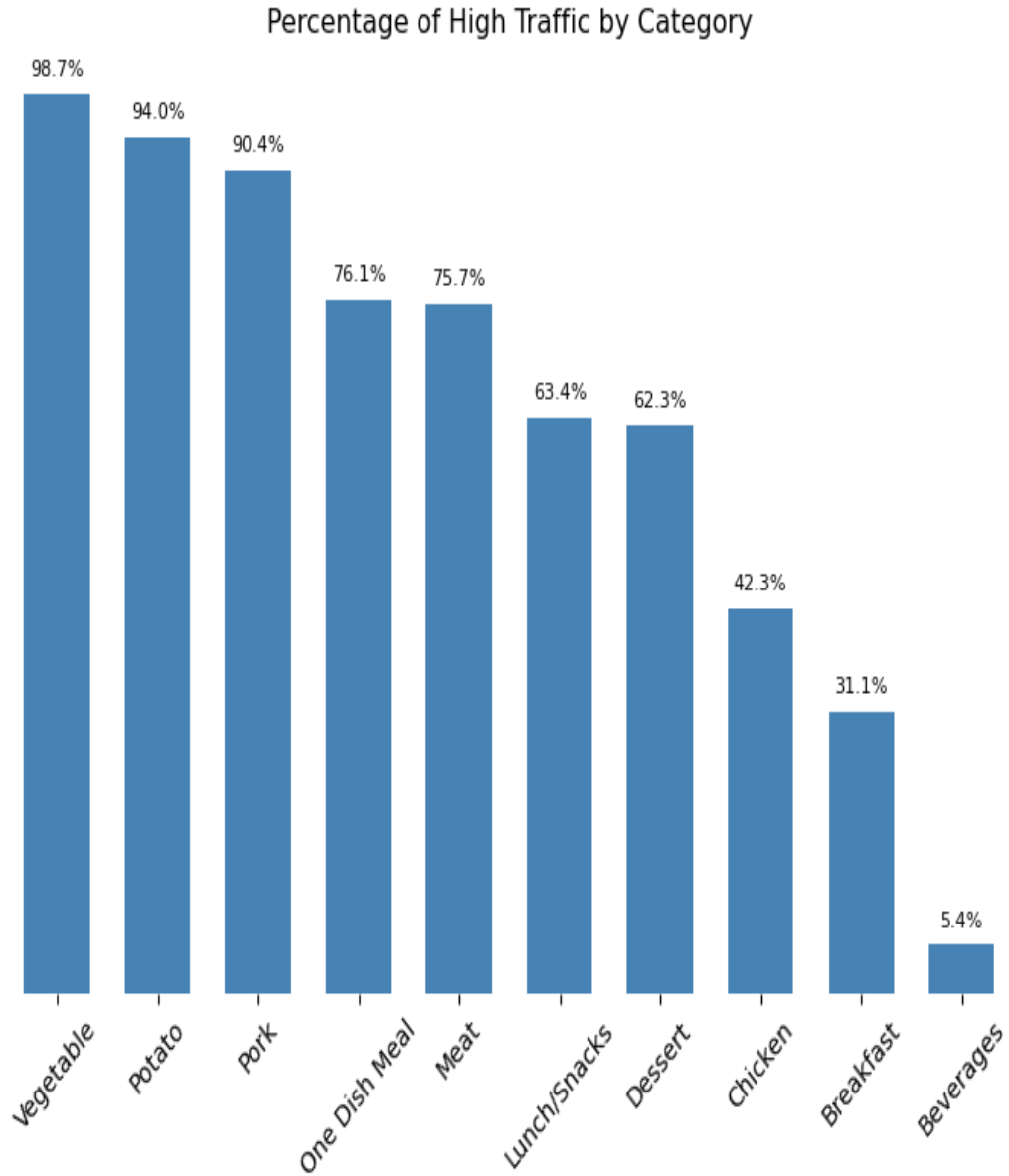
- Tasty Bites is a subscription-based meal planning and ingredient delivery service that provides healthy and balanced meals to customers, regardless of their budget.

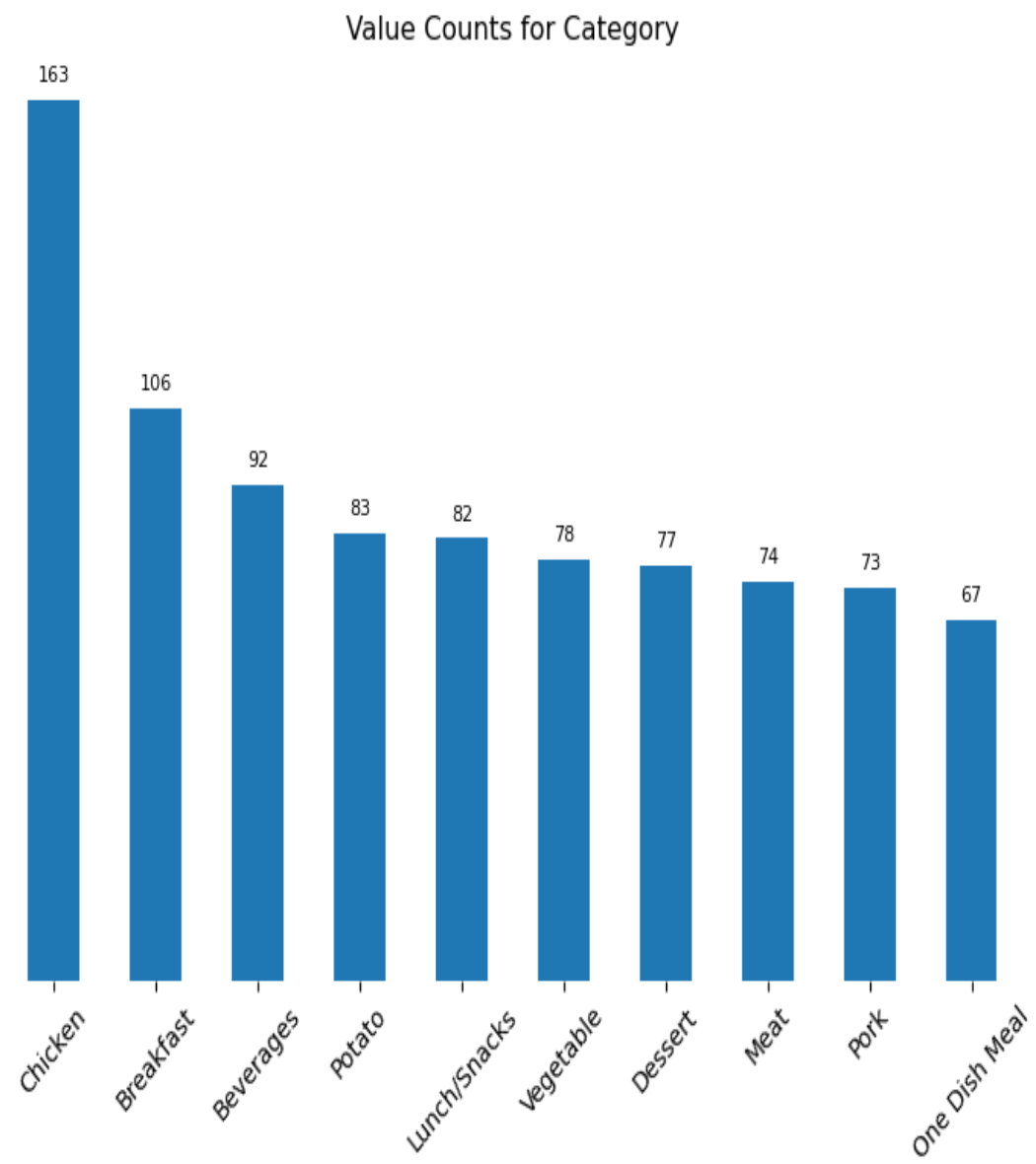
- The Product Manager of Tasty Bytes wants us to predict which recipes will be popular and lead to high traffic of the website 80% of the time and minimize the chance of showing unpopular recipes.

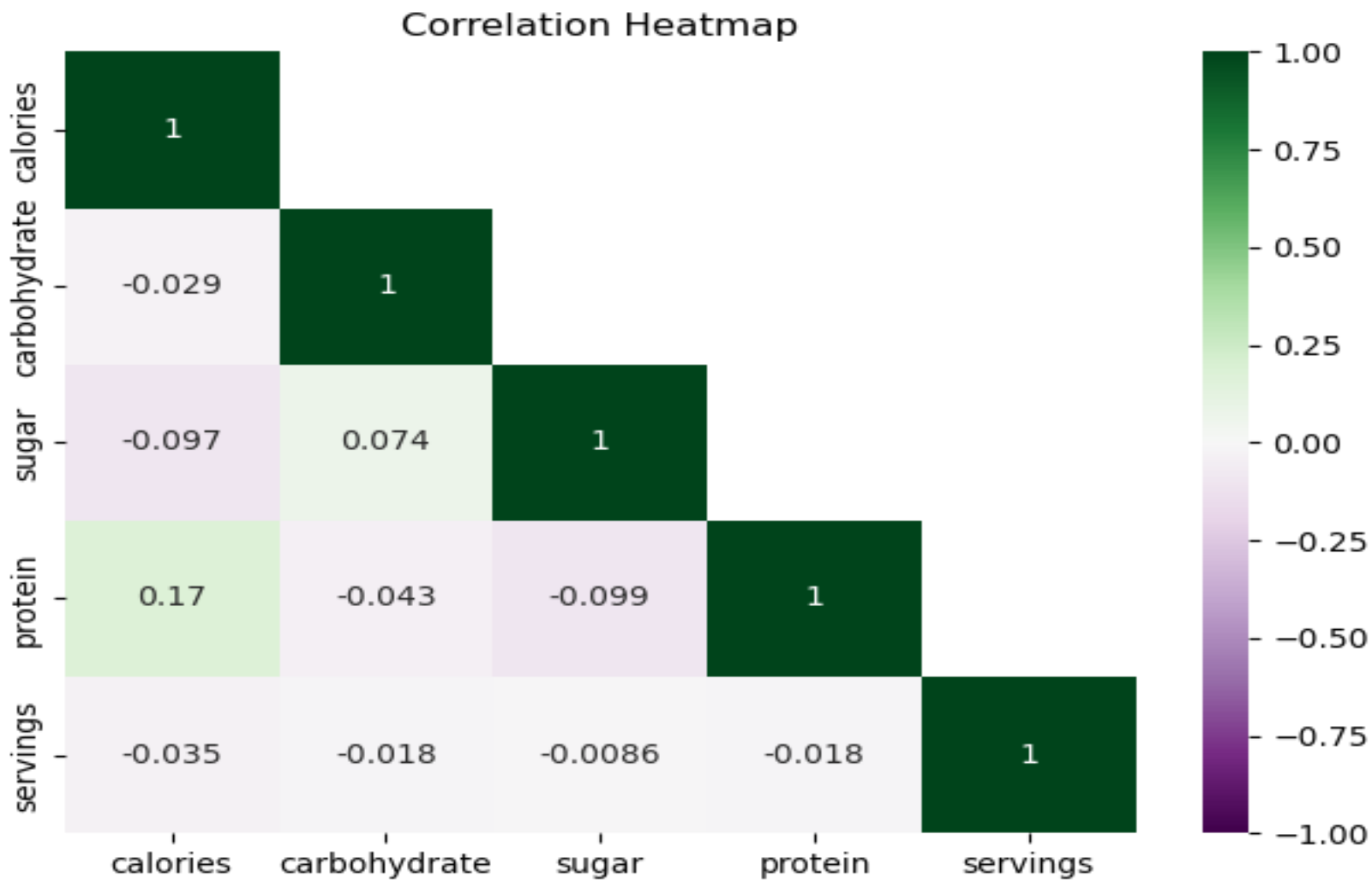
- Increase website traffic will boost subscriptions, which will drive higher profits.

- From the business perspective, all recipes were initially predicted to generate high traffic.
- Current business precision, percentage recipes actually leading to high traffic, is only "59.78%".




Percentage of High Traffic by Category
98.7%
94.0%
90.4%
76.1%
75.7%
63.4%
62.3%
42.3%
31.1%
5.4%
Vegetable
Potato
Pork
One Dish Meal
Meat
Lunch/Snacks
Dessert
Chicken
Breakfast
Beverages

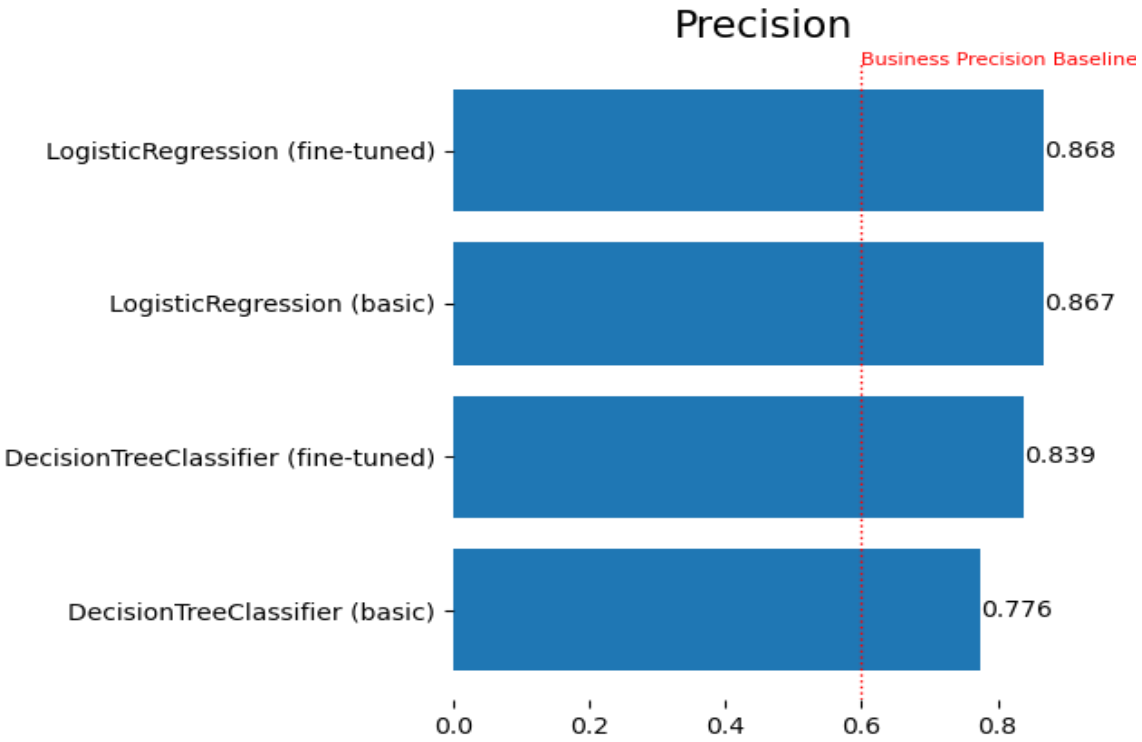

Value Counts for Category
163
106
92
83
82
78
77
74
73
67
Chicken
Breakfast
Beverages
Potato
Lunch/Snacks
Vegetable
Dessert
Meat
Pork
One Dish Meal

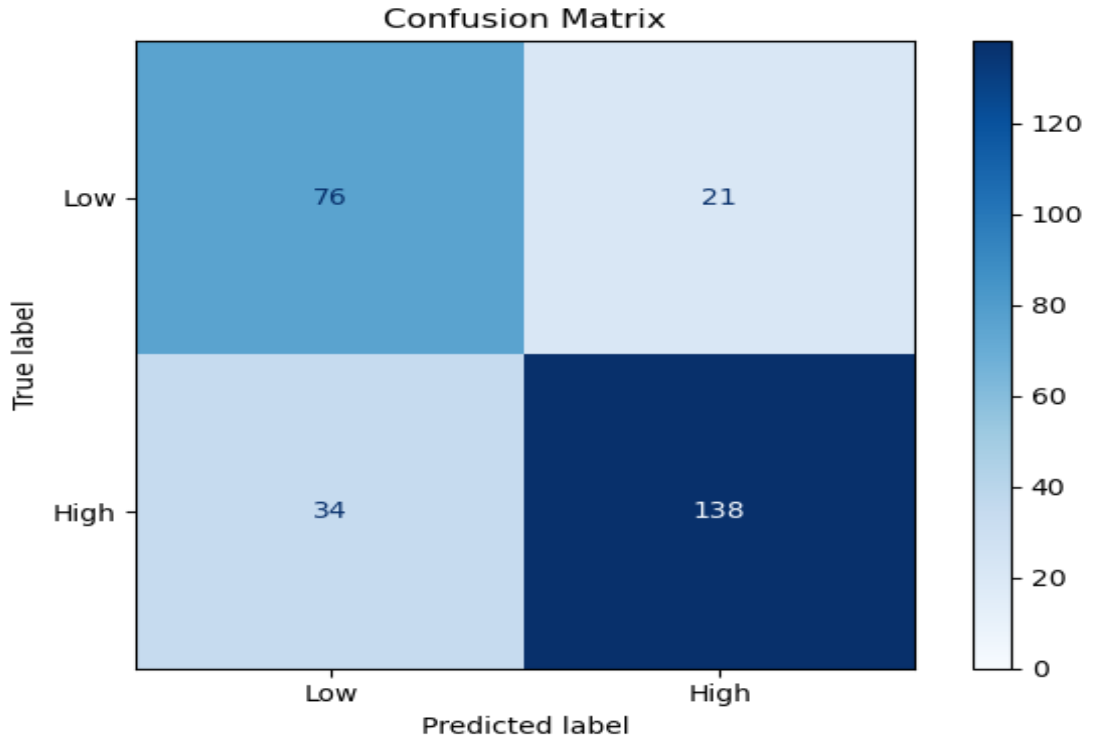

Correlation Heatmap
calories
carbohydrate
sugar
protein
servings
1
-0.029
1
-0.097
0.074
1
0.17
-0.043
-0.099
1
-0.035
-0.018
-0.0086
-0.018
1
calories
carbohydrate
sugar
protein
servings
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

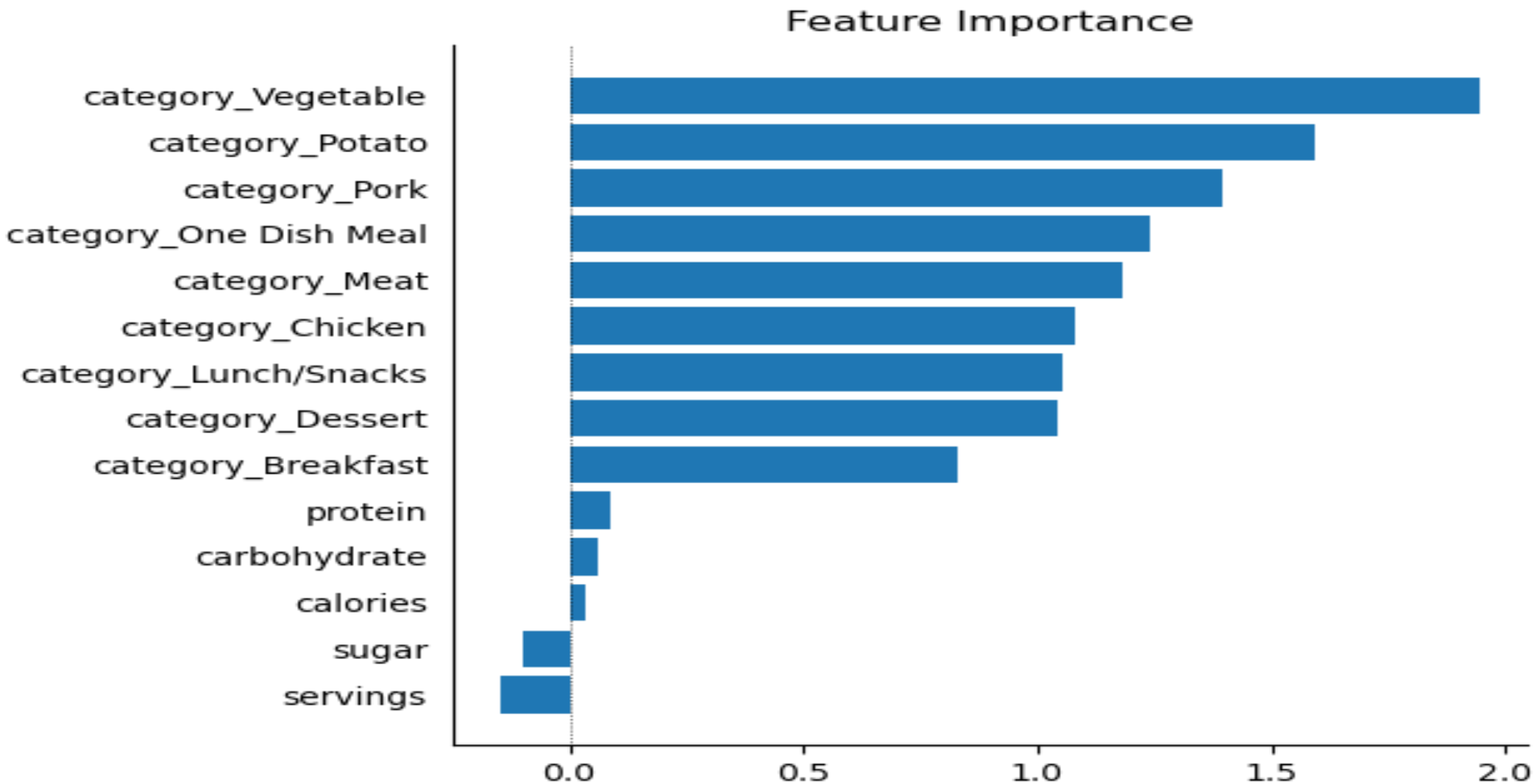
Business Metrics and Model Development:

- ✓ **Precision** metric was chosen as the KPI.
- ✓ Initially, only 60% of the predictions were accurate. The goal is to increase this number to at least 80%.
- ✓ To address this, a model is developed and fine-tuned to determine the likelihood of a recipe generating high traffic.
- ✓ Two models were developed: a Decision Tree Model and a Logistic Regression model.

Model	Precision	Accuracy	F1 Score
LogisticRegression (fine-tuned)	0.867925	0.795539	0.833837
LogisticRegression (basic)	0.867089	0.791822	0.830303
DecisionTreeClassifier (fine-tuned)	0.838926	0.736059	0.778816
DecisionTreeClassifier (basic)	0.775641	0.680297	0.737805






Feature Importance
category_Vegetable
category_Potato
category_Pork
category_One Dish Meal
category_Meat
category_Chicken
category_Lunch/Snacks
category_Dessert
category_Breakfast
protein
carbohydrate
calories
sugar
servings
0.0
0.5
1.0
1.5
2.0

Recommendations:

- ✓ To increase website traffic and boost subscriptions for higher profits, implement the Logistic Regression model, which achieved a precision of 0.86, representing a **43% improvement** over the baseline precision of 0.6.
- ✓ Prioritize key features such as the **vegetable, pork, and potato categories** to boost traffic levels based on exploratory analysis and feature importance analysis.
- ✓ Expand the dataset by incorporating **additional features** like "Time to Make," "Cost per Serving," and "Ingredients" to enhance the model's predictions.
- ✓ Regularly update and retrain the model using fresh data for continued effectiveness.

Thank You