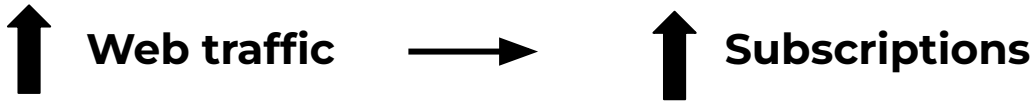
Predicting Recipes Site Traffic

A solution for Tasty Bytes

The Project

Predicting recipes to be shown on web site



Business Goals

Predict which recipes will lead to high traffic

Correctly predict high traffic recipes 80% of the time

Data

Information of each recipe given by Product team

Data

Information of each recipe given by Product team

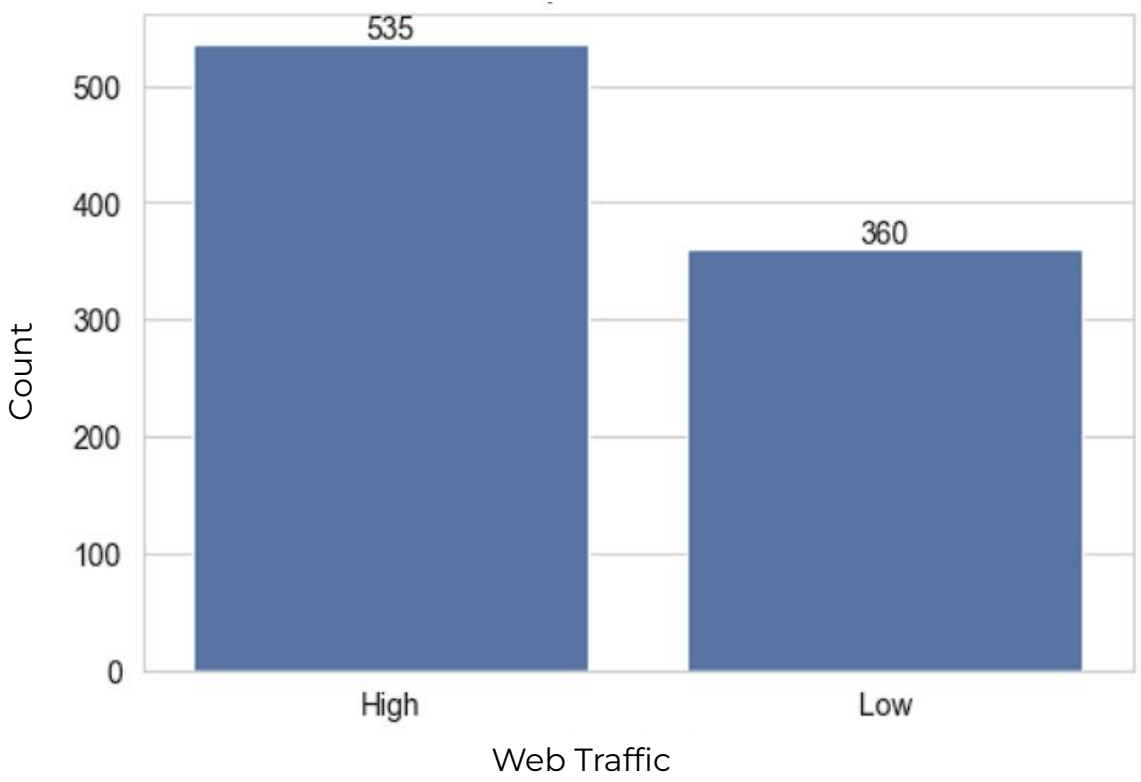
	recipe	calories	carbohydrate	sugar	protein	category	servings	high_traffic
0	1	null	null	null	null	Pork	6	High
1	2	35.48	38.56	0.66	0.92	Potato	4	High
2	3	914.28	42.68	3.09	2.88	Breakfast	1	null
3	4	97.03	30.56	38.63	0.02	Beverages	4	High
4	5	27.05	1.85	0.8	0.53	Beverages	4	null

Data Validation

- 52 rows with missing values (removed)
- Category: “chicken breast” merged into “chicken”
- Servings “4-6 as a snack”: only considered the numeric information

Key Findings

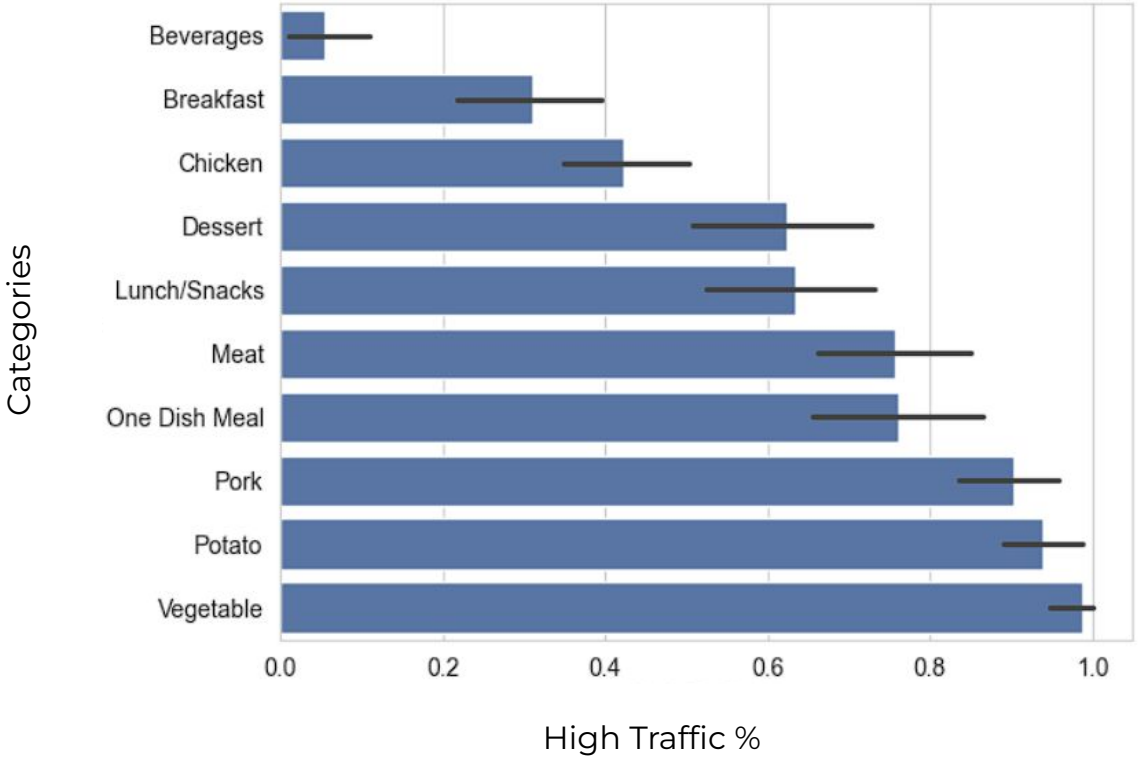
Recipe's Web Traffic



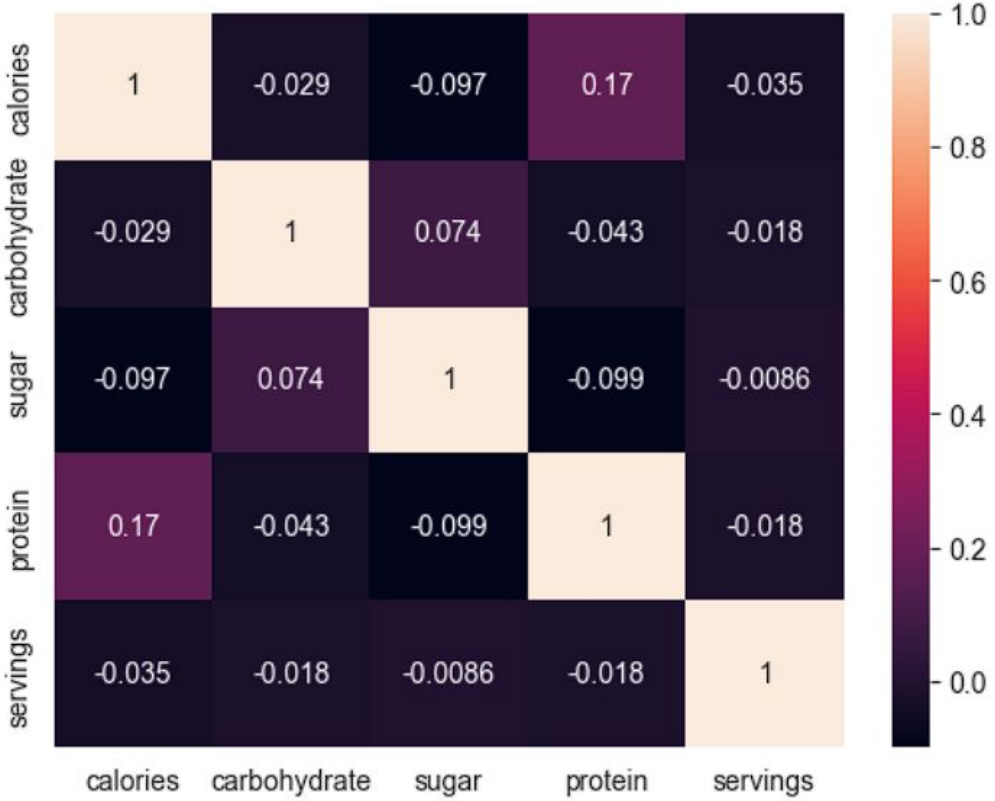
High: 59.8%

Low: 40.2%

High Traffic Probability per Categories



Variables Correlation



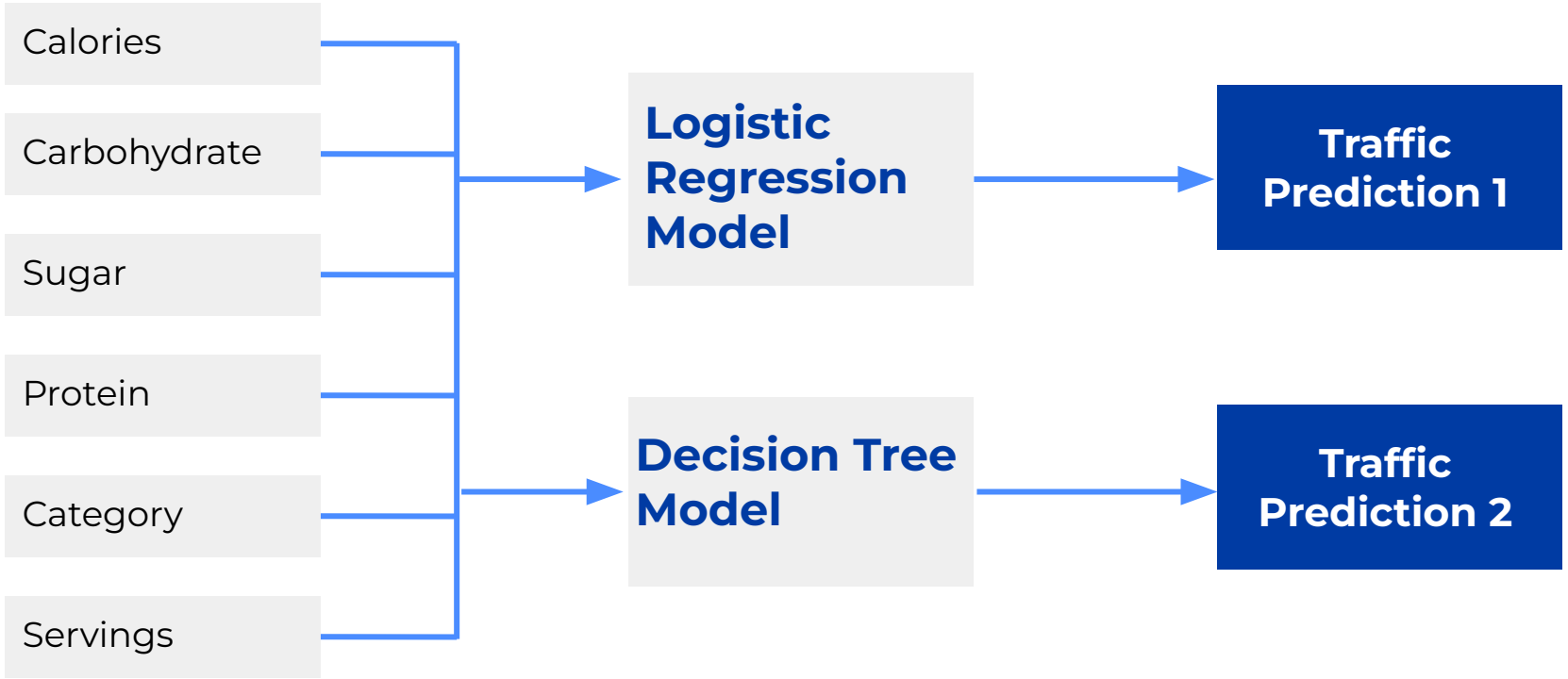
Weak relationship between variables

0: no correlation

1: perfect linear correlation

Outcomes

Models


Calories
Carbohydrate
Sugar
Protein
Category
Servings
Logistic Regression Model
Decision Tree Model
Traffic Prediction 1
Traffic Prediction 2

Model Metric

Precision

It measures how reliable the model is when it says something is positive (i.e. high traffic).

Range: 0 to 1. The greater, the better.

Model Metric

Model Name	Precision Score
Logistic Regression	0.819
Decision Tree	0.814

Model Metric

Model Name	Precision Score
Logistic Regression	0.819
Decision Tree	0.814

Business Metric

KPI

1. Selection of recipes with at least 80% chances of high traffic.
2. From this selection, compare the predicted class (high or non) against the real one.
3. Get the score of correct predictions

The higher the percentage, the better the model performs.

Business Metric

KPI

Model Name	KPI %
Logistic Regression	97.3%
Decision Tree	89.7%

Business Metric

KPI

Model Name	KPI %
Logistic Regression	97.3%
Decision Tree	89.7%

Recommendations

Recommendations

- Apply A/B testing on website during next month and collect more data.
- Deploy Logistic Regression model into production. Compare with estimated KPI (97.3%)
- Provide missing information (cost per serving, ingredients) in order to increase the model precision.

Thanks

Questions?

GabrielMazzotta@hotmail.com