



Recipe Price Per Serving Estimator

Crissy Bruce

Business Case

The client, Spoonacular, is a recipe website that provides detailed information about recipes but is missing the ability to estimate price per serving for the recipes. Spoonacular has hired Crissy Bruce Consulting to build a model to predict the price per serving for recipes. Spoonacular users will input ingredients, and the model will provide an estimate of the price per serving for the recipe.

Spoonacular - API Process

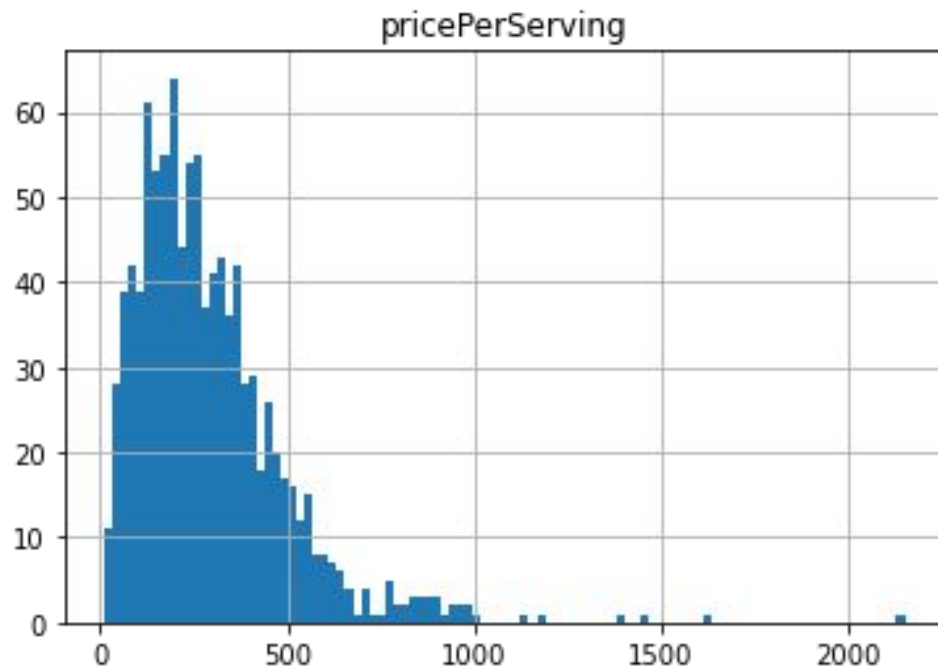
- The recipe data was sourced from <https://spoonacular.com/food-api> via an API call.
 - Requested API key from Spoonacular website
 - Learn about all the requests that are available from the website and trying several before finally learning about the endpoint that will return everything needed
 - Viewed the data via json file to confirm data points
 - Created a function that would pull down multiple requests at a time
 - Converted the data to a dataframe
 - Converted the dataframe to a .csv
-

The Data

- Dataset included recipe name, ingredients, price per serving and more for 1000 different recipes.
 - pricePerServing is in cents and range for all recipes is 13.23(\$0.13)-2149.55(\$21.50)
 - Feature of the dataset is a list of ingredients for all of the recipes
 - The mean price per serving is 293.59977 cents(\$2.94), and the standard deviation is 204.26(\$2.04)
-

EDA

Most prices are between 0 and 500 cents (\$0-\$5)



EDA

Word cloud of the ingredients that appear most often in the recipes



Modeling

Baseline Model: Linear Regression

Test MAE: 845812950644554.8

Train MAE: 2.59

Final Model: Support Vector Regression

Test MAE: 126.33

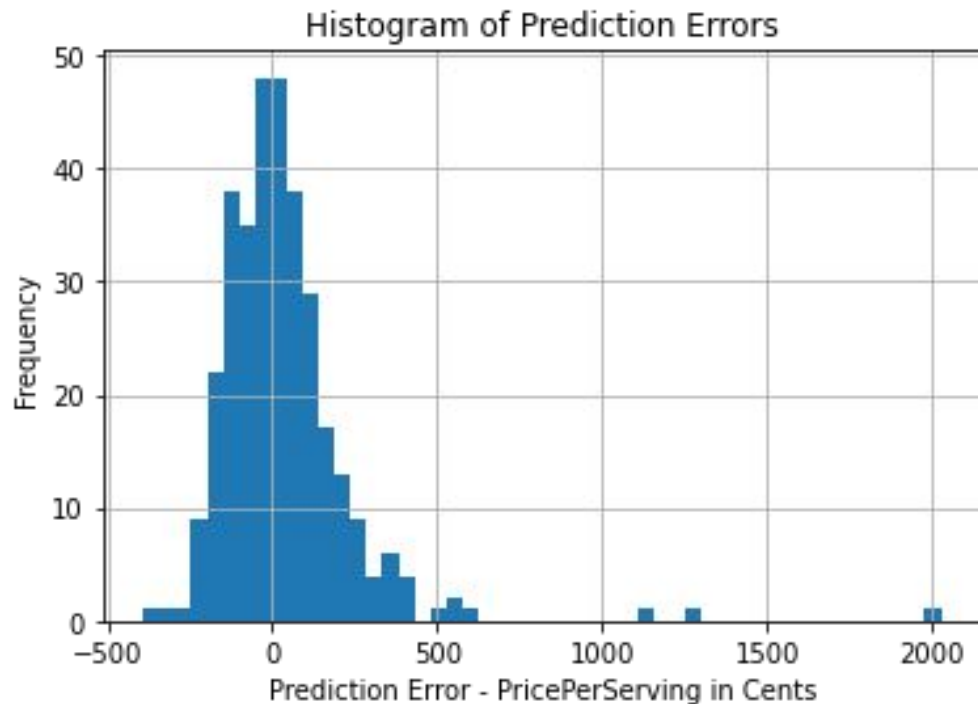
Train MAE: 56.67

The grid search returned an SVD model that is **marginally better** than the original SVR model.

Comment: Average error close to the standard deviation of 204.26 cents (\$2.04)

On average, the model's predictions for price per serving are approximately 126 cents (\$1.26) off on scale of 13.23(\$0.13)-2149.55(\$21.50).

Post Modeling EDA



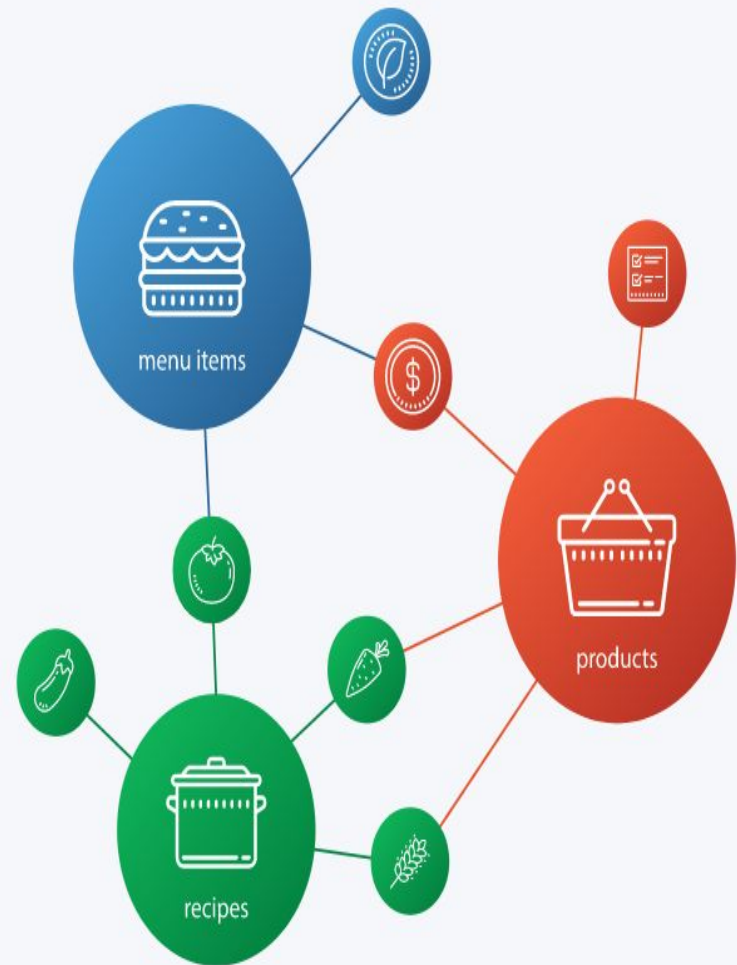
The final model is not a perfect fit based on the RMSE (209.97), but is very close to the standard deviation of the original dataset (204.26 cents).

Conclusion

The final model is not a perfect fit based on the RMSE (209.97), but is very close to the standard deviation of the original dataset (204.26 cents).

Future Steps

- Obtain more recipes as there are some outliers within the original data that could skew the modeling work or remove outliers.
 - Try other models, such as Neural Networks as it contains set of adaptive weights as well as the ability to group the ingredients into different classes.
 - Check for other features that may be relevant at predicting price per serving.
-



Thank you!

Crissy Bruce (crissybruce@gmail.com)

<https://github.com/crissymae/solid-spoon>
