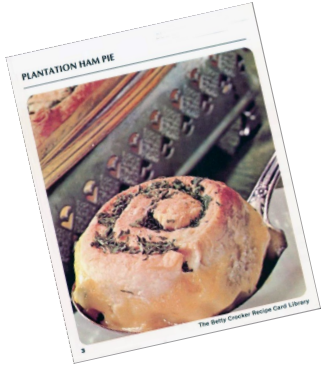
Recipe Analysis with Natural Language Processing

A. Marshall

The Two C's of Modern Cooking

COVID COST



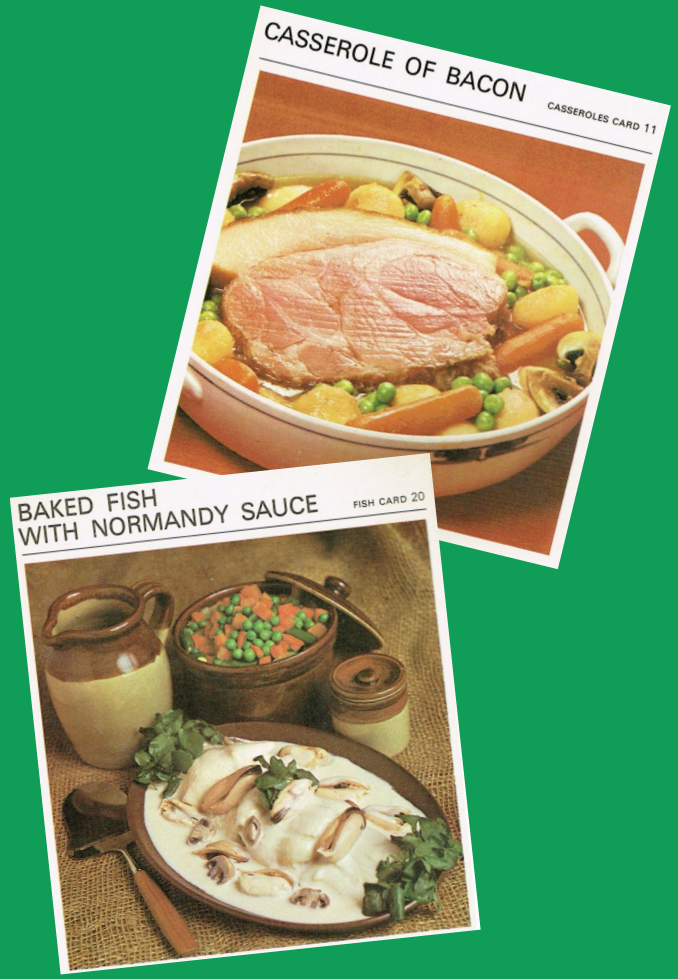
Goal

Websites, publishers, other collections

Easy Dinner Recipes

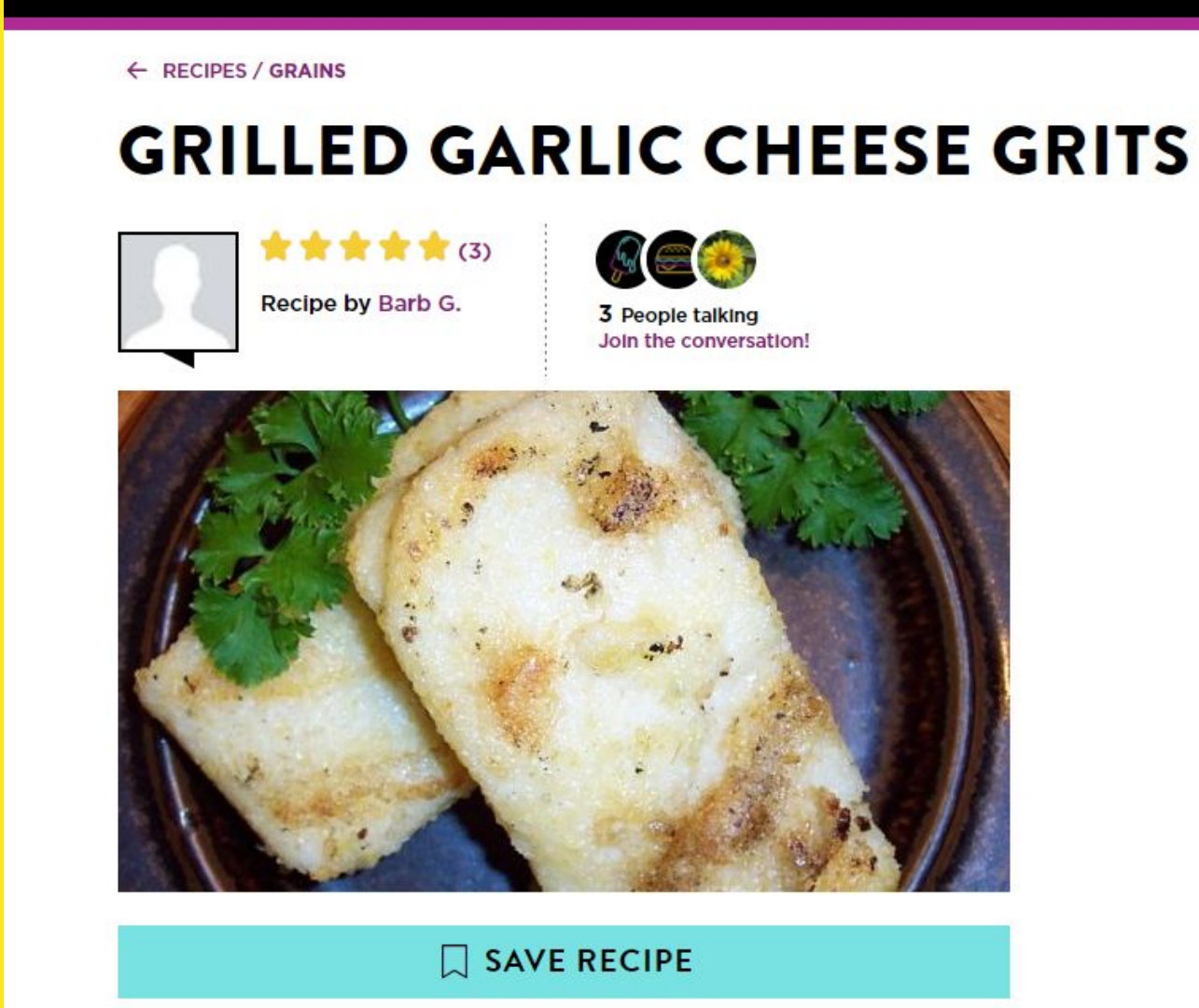
Consistency, Speed

Vet submissions, SEO, organization

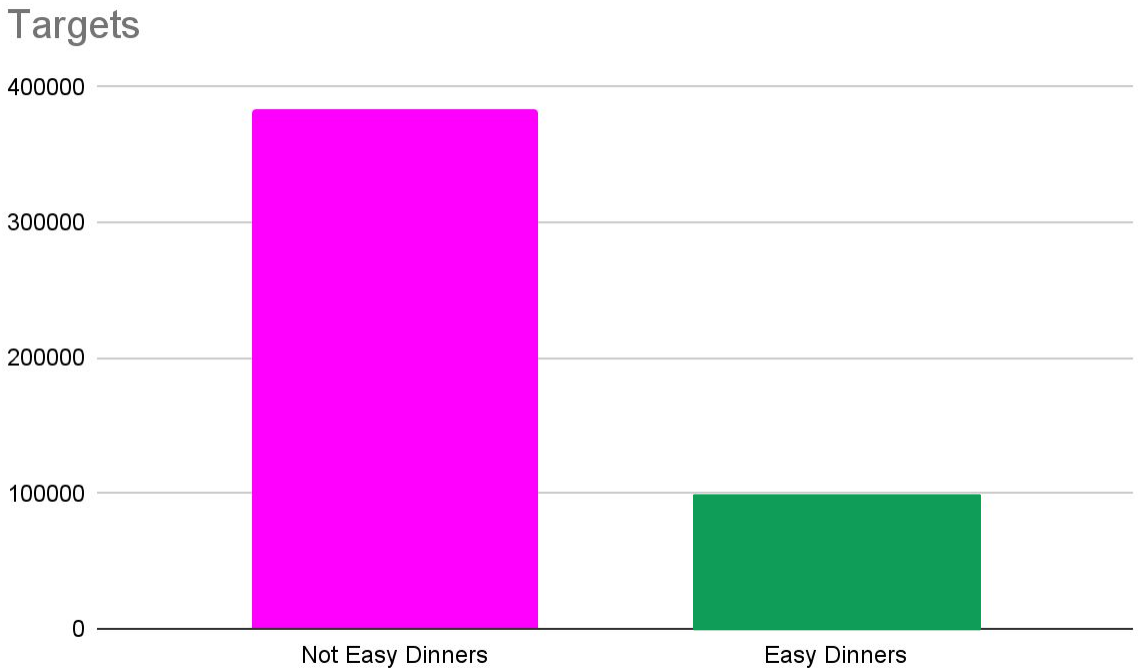


Data

- Food.com
- ~500,000 recipes
- Metadata:
  - Tags
  - Search Terms



Target Data

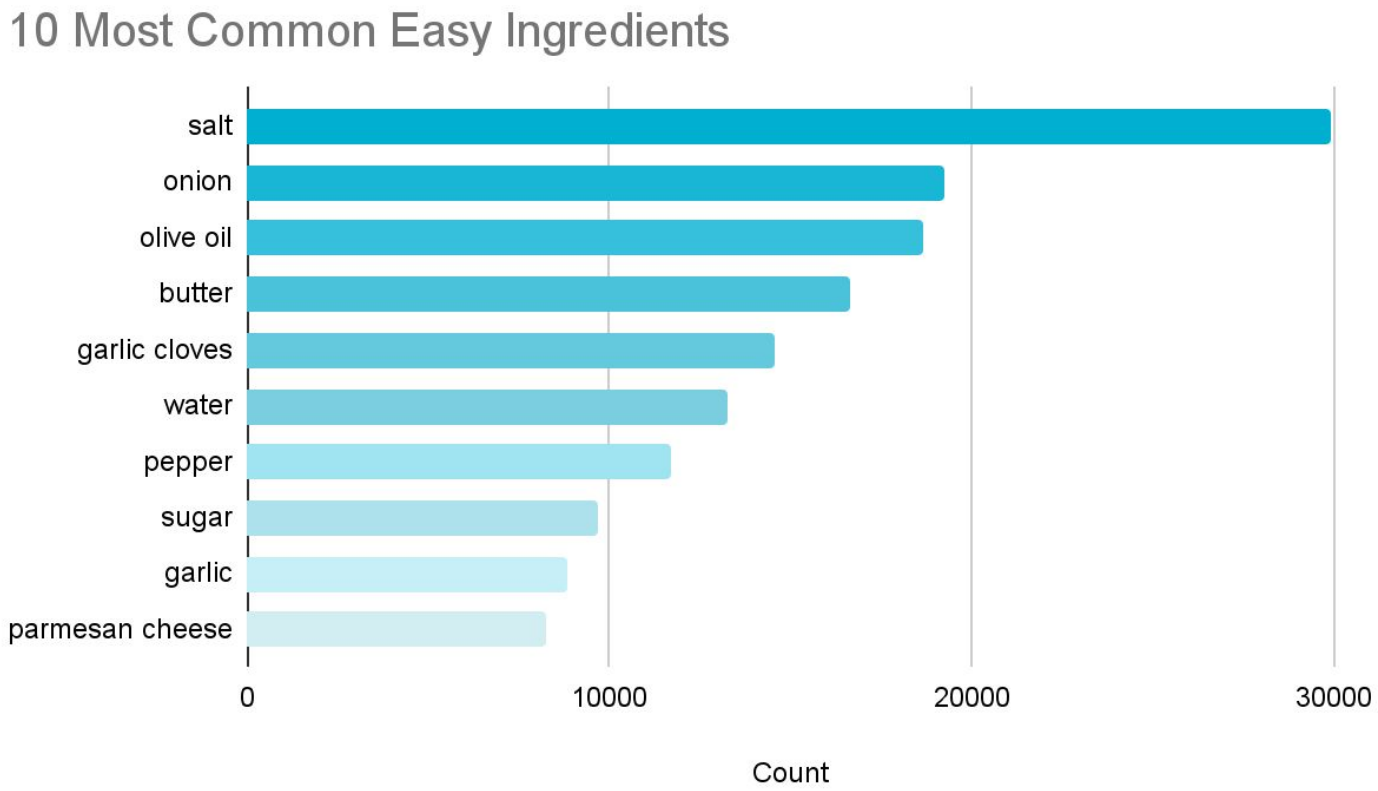


Search Term: dinner

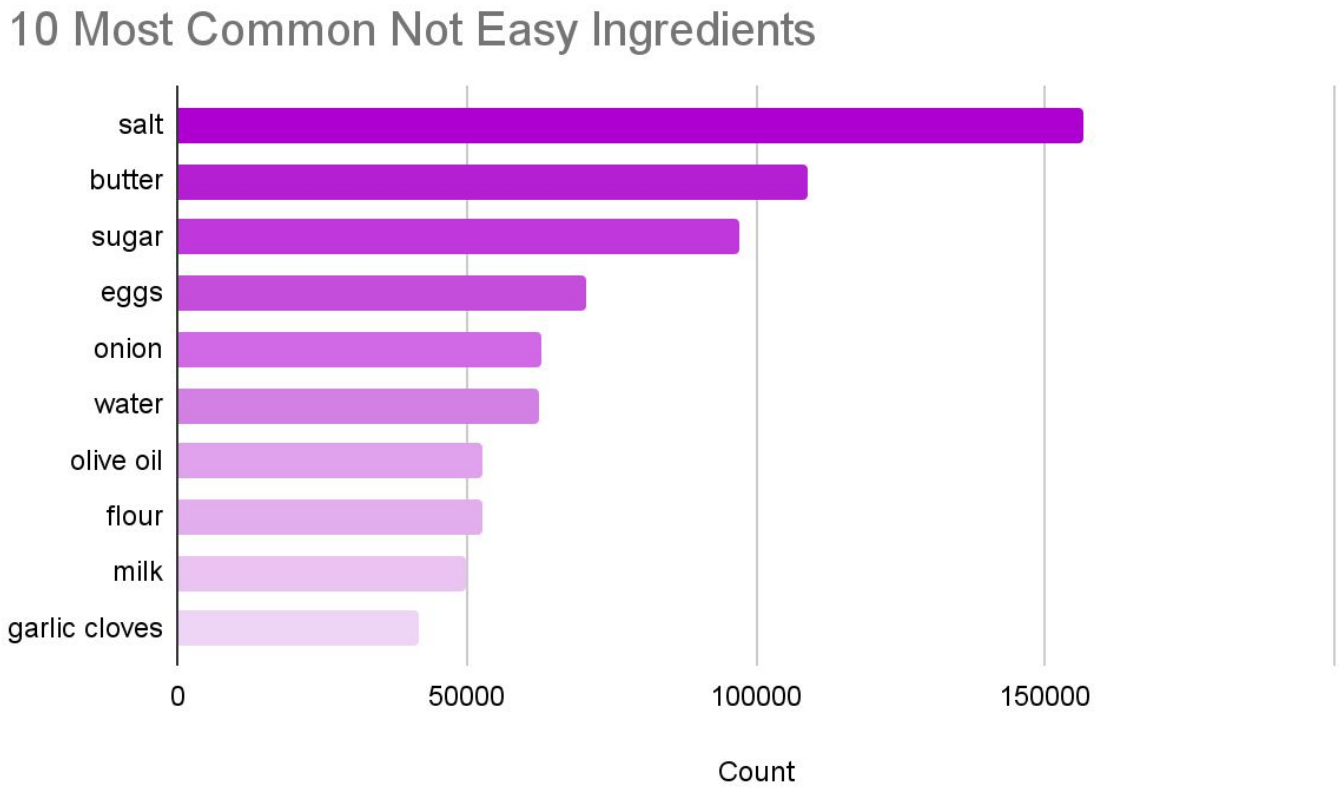
Tags:

- easy
- 30-minutes-or-less
- 15-minutes-or-less
- 3-steps-or-less
- beginner-cook

EDA

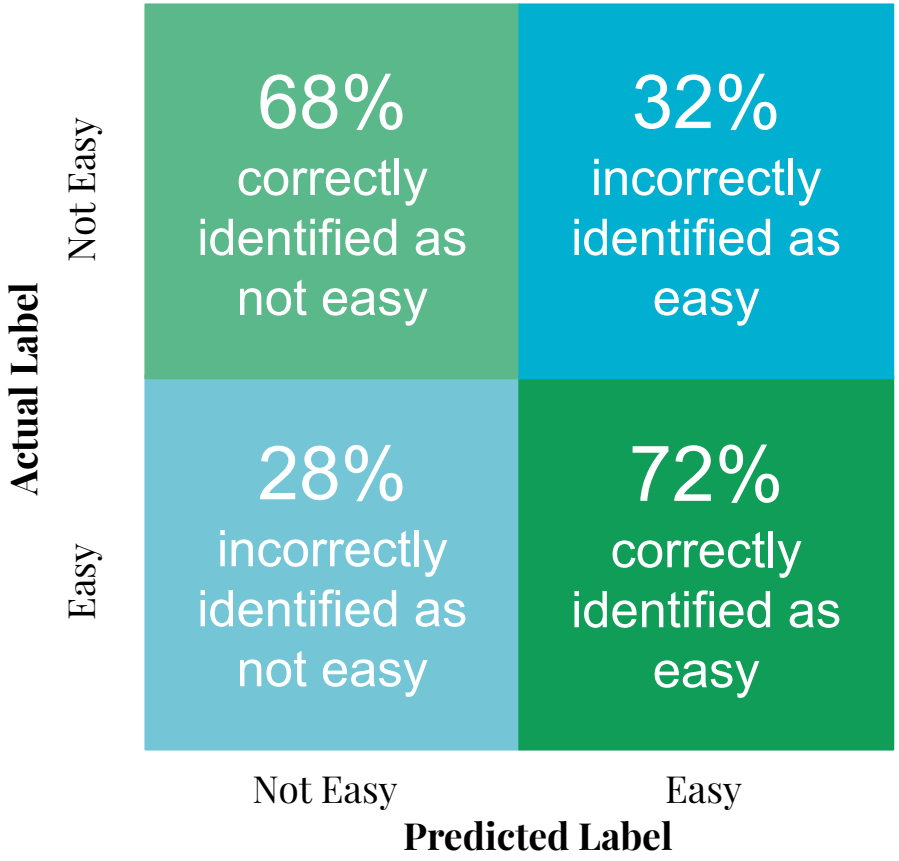

10 Most Common Easy Ingredients
salt
onion
olive oil
butter
garlic cloves
water
pepper
sugar
garlic
parmesan cheese
0
10000
20000
30000
Count

EDA

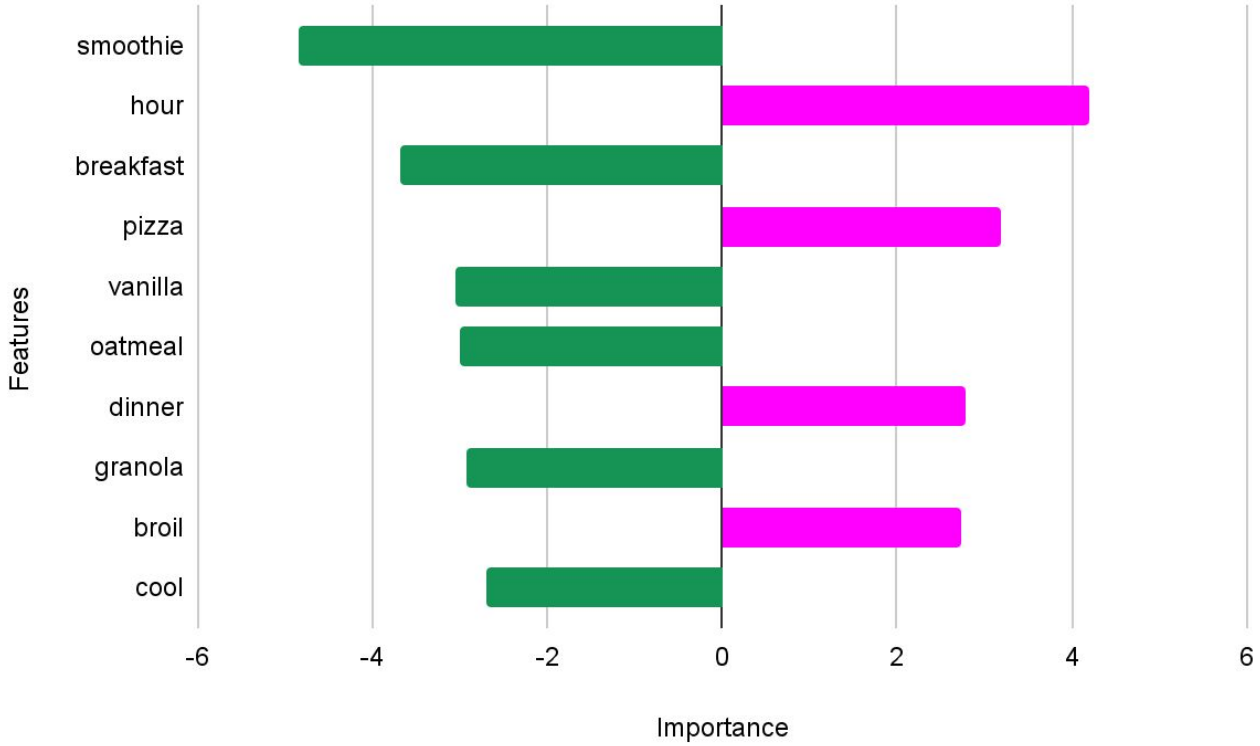

10 Most Common Not Easy Ingredients
salt
butter
sugar
eggs
onion
water
olive oil
flour
milk
garlic cloves
0
50000
100000
150000
Count

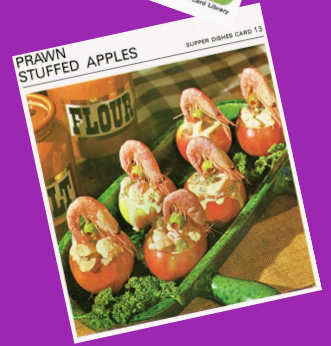
Model

- Logistic Regression
- 70% accuracy



Feature Importance





Next Steps

Increase Model Targets

Incorporate Ratings Scores

Further Data Cleaning

THANK YOU

Images courtesy of Betty Crocker via the Gallery of Regrettable Foods

Project Details: https://github.com/AJMarshall-25/NLP_Recipe_Analysis
Contact: alex.marshall842@outlook.com