

Culinary Matcher: Multilingual Recipe Recommendation System Based on Ingredients

Mt. SAC CISB 63 Midterm Project Fall 2023

Vedavit Shetty

The Culinary Matcher is a sophisticated recipe recommendation tool created to assist users in discovering new culinary delights based on ingredients they have available. Going beyond merely listing out recipes with the specified ingredients, the system employs Natural Language Processing (NLP) techniques such as tokenization, Part-of-Speech (POS) tagging, Named Entity Recognition (NER), and TF-IDF weighting. With the integration of cosine similarity scores, it recommends the most relevant recipes. Moreover, the system boasts a multilingual feature, enabling users to receive recipe details, including titles, ingredients, and instructions, translating it to the language of their choice.

```
In [1]: #ignore warnings:
import warnings
warnings.filterwarnings("ignore")
!pip install textblob
!python -m textblob.download_corpora
```

```
Requirement already satisfied: textblob in /Users/vedavitshetty/anaconda3/lib/python3.11/site-packages (0.17.1)
Requirement already satisfied: nltk>=3.1 in /Users/vedavitshetty/anaconda3/lib/python3.11/site-packages (from textblob) (3.8.1)
Requirement already satisfied: click in /Users/vedavitshetty/anaconda3/lib/python3.11/site-packages (from nltk>=3.1->textblob) (8.0.4)
Requirement already satisfied: joblib in /Users/vedavitshetty/anaconda3/lib/python3.11/site-packages (from nltk>=3.1->textblob) (1.2.0)
Requirement already satisfied: regex>=2021.8.3 in /Users/vedavitshetty/anaconda3/lib/python3.11/site-packages (from nltk>=3.1->textblob) (2022.7.9)
Requirement already satisfied: tqdm in /Users/vedavitshetty/anaconda3/lib/python3.11/site-packages (from nltk>=3.1->textblob) (4.65.0)
[nltk_data] Downloading package brown to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data]   Package brown is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data]   Package averaged_perceptron_tagger is already up-to-date!
[nltk_data] Downloading package conll2000 to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data]   Package conll2000 is already up-to-date!
[nltk_data] Downloading package movie_reviews to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data]   Package movie_reviews is already up-to-date!
Finished.
```

```
In [2]: #!pip install WordCloud

# Import necessary libraries:
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.probability import FreqDist
from nltk.chunk import ne_chunk
import re
from wordcloud import WordCloud
from textblob import TextBlob
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity

# download stopwords, punkt, and averaged_perceptron_tagger, maxent_ne_chunker, and words
nltk.download('stopwords')
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
nltk.download('maxent_ne_chunker')
nltk.download('words')

[nltk_data] Downloading package stopwords to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data] Package averaged_perceptron_tagger is already up-to-
[nltk_data]   date!
[nltk_data] Downloading package maxent_ne_chunker to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data] Package maxent_ne_chunker is already up-to-date!
[nltk_data] Downloading package words to
[nltk_data]   /Users/vedavitshetty/nltk_data...
[nltk_data] Package words is already up-to-date!
```

Out[2]: True

```
In [3]: recipes = pd.read_csv('recipes.csv', index_col=0)
```

```
In [4]: recipes
```

Out [4]:

	Title	Ingredients	Instructions	Image_Name	Cleaned_Ingredients
0	Miso-Butter Roast Chicken With Acorn Squash Pa...	['1 (3½-4-lb.) whole chicken', '2¾ tsp. kosher...	Pat chicken dry with paper towels, season all ...	miso-butter-roast-chicken-acorn-squash-panzanella	['1 (3½-4-lb.) whole chicken', '2¾ tsp. kosher...
1	Crispy Salt and Pepper Potatoes	['2 large egg whites', '1 pound new potatoes (...	Preheat oven to 400°F and line a rimmed baking...	crispy-salt-and-pepper-potatoes-dan-kluger	['2 large egg whites', '1 pound new potatoes (...
2	Thanksgiving Mac and Cheese	['1 cup evaporated milk', '1 cup whole milk', ...	Place a rack in middle of oven; preheat to 400...	thanksgiving-mac-and-cheese-erick-williams	['1 cup evaporated milk', '1 cup whole milk', ...
3	Italian Sausage and Bread Stuffing	['1 (¾- to 1-pound) round Italian loaf, cut in...	Preheat oven to 350°F with rack in middle. Gen...	italian-sausage-and-bread-stuffing-240559	['1 (¾- to 1-pound) round Italian loaf, cut in...
4	Newton's Law	['1 teaspoon dark brown sugar', '1 teaspoon ho...	Stir together brown sugar and hot water in a c...	newtons-law-apple-bourbon-cocktail	['1 teaspoon dark brown sugar', '1 teaspoon ho...
...
13496	Brownie Pudding Cake	['1 cup all-purpose flour', '2/3 cup unsweeten...	Preheat the oven to 350°F. Into a bowl sift to...	brownie-pudding-cake-14408	['1 cup all-purpose flour', '2/3 cup unsweeten...
13497	Israeli Couscous with Roasted Butternut Squash...	['1 preserved lemon', '1 1/2 pound butternut s...	Preheat oven to 475°F.½ lemons and scoop...	israeli-couscous-with-roasted-butternut-squash...	['1 preserved lemon', '1 1/2 pound butternut s...
13498	Rice with Soy-Glazed Bonito Flakes and Sesame...	['Leftover katsuo bushi (dried bonito flakes) ...	If using katsuo bushi flakes from package, moi...	rice-with-soy-glazed-bonito-flakes-and-sesame-...	['Leftover katsuo bushi (dried bonito flakes) ...
13499	Spanakopita	['1 stick (1/2 cup) plus 1 tablespoon unsalted...	Melt 1 tablespoon butter in a 12-inch heavy sk...	spanakopita-107344	['1 stick (1/2 cup) plus 1 tablespoon unsalted...
13500	Mexican Poblano, Spinach, and Black Bean "Lasa...	['12 medium to large fresh poblano chiles (2 1...	Lay 4 chiles on their sides on racks of gas bu...	mexican-poblano-spinach-and-black-bean-lasagne...	['12 medium to large fresh poblano chiles (2 1...

13501 rows × 5 columns

Exploratory Data Analysis (EDA)

Check for missing or null values and handle it

```
In [5]: recipes.isnull().sum()
```

Out [5]:

Title	5
Ingredients	0
Instructions	8
Image_Name	0
Cleaned_Ingredients	0
dtype: int64	

```
In [6]: recipes.dropna(inplace=True)
```

Inspect the Data

```
In [7]: recipes.describe()
```

Out [7]:

	Title	Ingredients	Instructions	Image_Name	Cleaned_Ingredients
count	13493	13493	13493	13493	13493
unique	13302	13471	13464	13464	13471
top	Potato Latkes	[]	Place ingredients in blender in the order list...	#NAME?	['']
freq	5	6	5	30	6

Drop the rows where the ingredients are not listed

```
In [8]: # Filter out rows where Cleaned_Ingredients is ""
recipes = recipes[recipes['Ingredients'] != '[]']
```

```
In [9]: recipes.describe()
```

Out [9]:

	Title	Ingredients	Instructions	Image_Name	Cleaned_Ingredients
count	13487	13487	13487	13487	13487
unique	13296	13470	13458	13458	13470
top	French 75	['1 cube or 1/2 teaspoon sugar', '4 dashes Pey...	Place ingredients in blender in the order list...	#NAME?	['1 cube or 1/2 teaspoon sugar', '4 dashes Pey...
freq	5	4	5	30	4

Data looks good, let's continue to inspect the data

```
In [10]: recipes.head()
```

Out [10]:

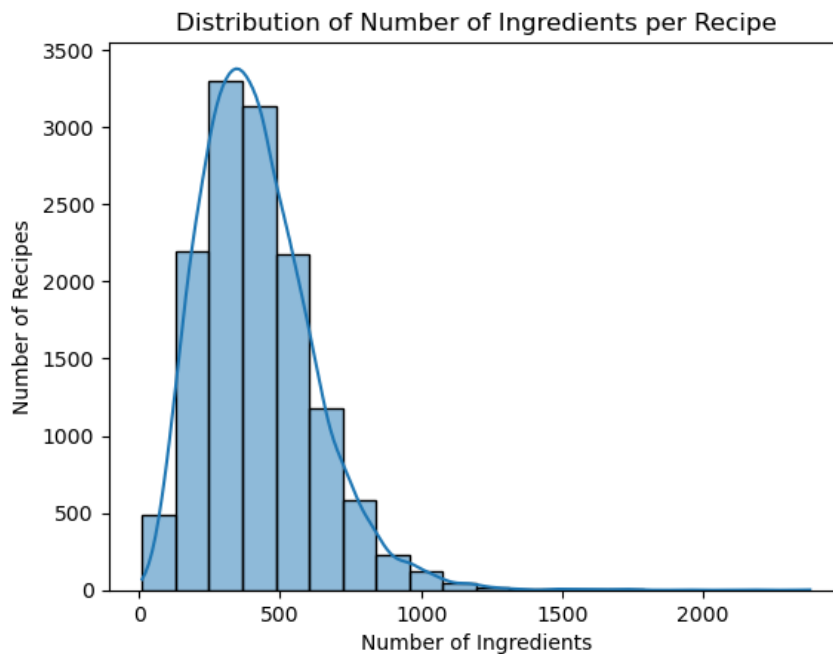
	Title	Ingredients	Instructions	Image_Name	Cleaned_Ingredients
0	Miso-Butter Roast Chicken With Acorn Squash Panzanella	['1 (3½-4-lb.) whole chicken', '2¼ tsp. kosher...	Pat chicken dry with paper towels, season all ...	miso-butter-roast-chicken-acorn-squash-panzanella	['1 (3½-4-lb.) whole chicken', '2¼ tsp. kosher...
1	Crispy Salt and Pepper Potatoes	['2 large egg whites', '1 pound new potatoes (...	Preheat oven to 400°F and line a rimmed baking...	crispy-salt-and-pepper-potatoes-dan-kluger	['2 large egg whites', '1 pound new potatoes (...
2	Thanksgiving Mac and Cheese	['1 cup evaporated milk', '1 cup whole milk', ...	Place a rack in middle of oven; preheat to 400...	thanksgiving-mac-and-cheese-erick-williams	['1 cup evaporated milk', '1 cup whole milk', ...
3	Italian Sausage and Bread Stuffing	['1 (¾- to 1-pound) round Italian loaf, cut in...	Preheat oven to 350°F with rack in middle. Gen...	italian-sausage-and-bread-stuffing-240559	['1 (¾- to 1-pound) round Italian loaf, cut in...
4	Newton's Law	['1 teaspoon dark brown sugar', '1 teaspoon ho...	Stir together brown sugar and hot water in a c...	newtons-law-apple-bourbon-cocktail	['1 teaspoon dark brown sugar', '1 teaspoon ho...

```
In [11]: recipes.info()
```

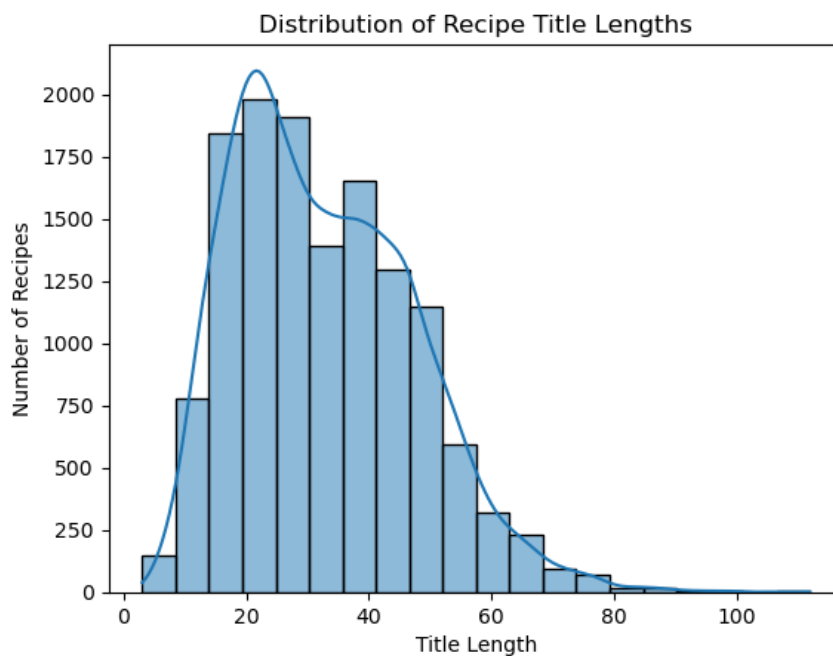
```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 13487 entries, 0 to 13500
Data columns (total 5 columns):
#   Column              Non-Null Count  Dtype
---  -
0   Title               13487 non-null object
1   Ingredients         13487 non-null object
2   Instructions        13487 non-null object
3   Image_Name         13487 non-null object
4   Cleaned_Ingredients 13487 non-null object
dtypes: object(5)
memory usage: 632.2+ KB
```

Visualizations

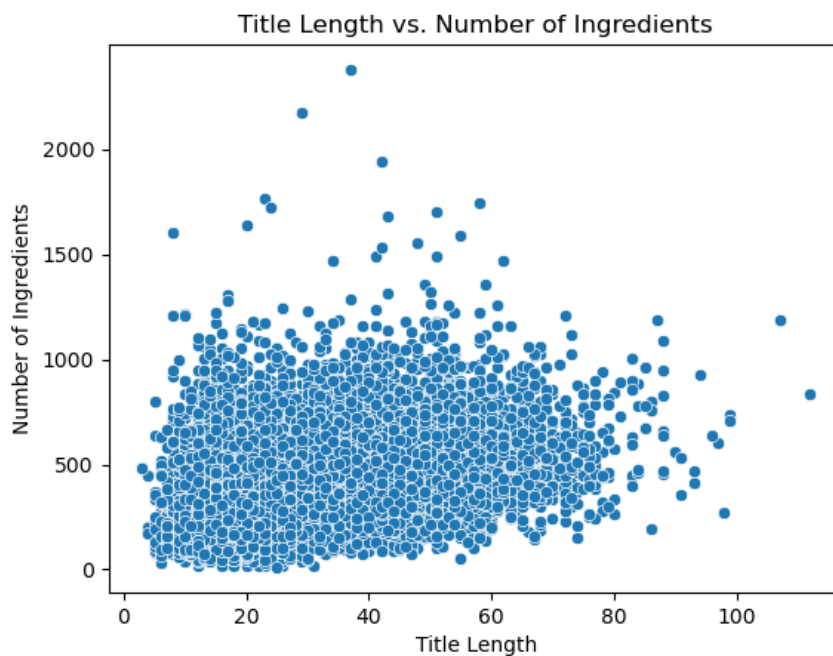
```
In [12]: # Distribution of Number of Ingredients per Recipe
recipes['num_ingredients'] = recipes['Ingredients'].apply(len)
sns.histplot(recipes['num_ingredients'], bins=20, kde=True)
plt.title('Distribution of Number of Ingredients per Recipe')
plt.xlabel('Number of Ingredients')
plt.ylabel('Number of Recipes')
plt.show()
```



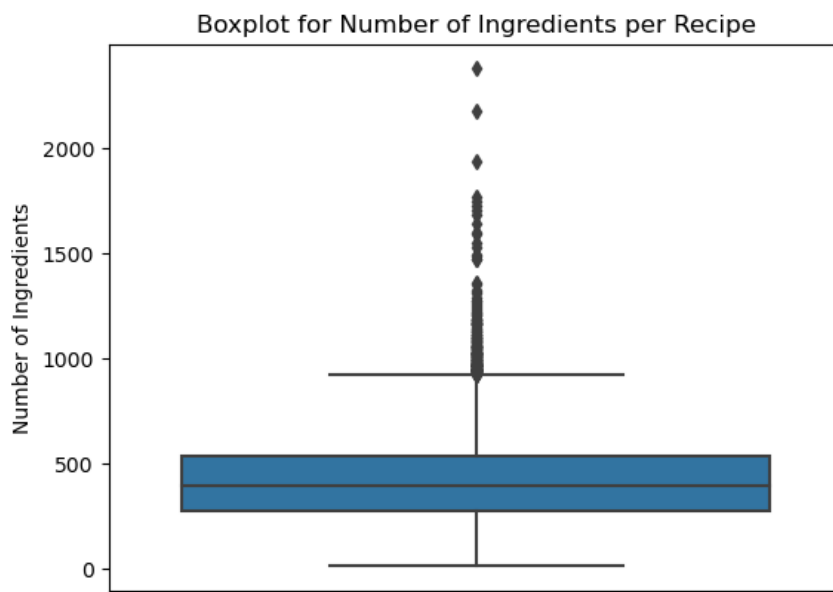
```
In [13]: # Recipe Title Lengths
recipes['title_length'] = recipes['Title'].apply(len)
sns.histplot(recipes['title_length'], bins=20, kde=True)
plt.title('Distribution of Recipe Title Lengths')
plt.xlabel('Title Length')
plt.ylabel('Number of Recipes')
plt.show()
```



```
In [14]: sns.scatterplot(x=recipes['title_length'], y=recipes['num_ingredients'])  
plt.title('Title Length vs. Number of Ingredients')  
plt.xlabel('Title Length')  
plt.ylabel('Number of Ingredients')  
plt.show()
```



```
In [15]: sns.boxplot(y=recipes['num_ingredients'])  
plt.title('Boxplot for Number of Ingredients per Recipe')  
plt.ylabel('Number of Ingredients')  
plt.show()
```



Some analysis:

- 75% of recipes have about 600 or less ingredients with the middle 50% containing about 100 to 600 ingredients
- There's not a strong correlation between recipe title and number of ingredients
- The distribution for recipe title lengths appears to be normally distributed around 20 to 40 characters.

Apply NLP Techniques to preprocess the data

```
op = set(stopwords.words('english'))

recipes['Cleaned_Ingredients'] = recipes['Cleaned_Ingredients'].apply(lambda x: ' '.join([word.lower() for word in x.split() if word not in op]))
recipes.head()
```

	Title	Ingredients	Instructions	Image_Name	Cleaned_Ingredients	num_ingredients	title_length
0	Miso-Butter Roast Chicken With Acorn Squash Panzanella	['1 (3½–4-lb.) whole chicken', '2% tsp. kosher...	Pat chicken dry with paper towels, season all ...	miso-butter-roast-chicken-acorn-squash-panzanella	['1 (3½–4-lb.) whole chicken', '2% tsp. kosher...	828	54
1	Crispy Salt and Pepper Potatoes	['2 large egg whites', '1 pound new potatoes (...	Preheat oven to 400°F and line a rimmed baking...	crispy-salt-and-pepper-potatoes-dan-kluger	['2 large egg whites', '1 pound new potatoes (...	251	31
2	Thanksgiving Mac and Cheese	['1 cup evaporated milk', '1 cup whole milk', ...	Place a rack in middle of oven; preheat to 400...	thanksgiving-mac-and-cheese-erick-williams	['1 cup evaporated milk', '1 cup whole milk', ...	289	27
3	Italian Sausage and Bread Stuffing	['1 (¾- to 1-pound) round Italian loaf, cut in...	Preheat oven to 350°F with rack in middle. Gen...	italian-sausage-and-bread-stuffing-240559	['1 (¾- 1-pound) round italian loaf, cut 1-inc...	558	34
4	Newton's Law	['1 teaspoon dark brown sugar', '1 teaspoon ho...	Stir together brown sugar and hot water in a c...	newtons-law-apple-bourbon-cocktail	['1 teaspoon dark brown sugar', '1 teaspoon ho...	218	12

	Title	Ingredients	Instructions	Image_Name	Cleaned_Ingredients	num_ingredients	title_length	Tokenized_Ingredients
0	Miso-Butter Roast Chicken With Acorn Squash Pa...	['1 (3½–4-lb.) whole chicken', '2¾ tsp. kosher...	Pat chicken dry with paper towels, season all ...	miso-butter-roast-chicken-acorn-squash-panzanella	['1 (3½–4-lb.) whole chicken', '2¾ tsp. kosher...	828	54	[[, '1, (, 3½–4-lb., ,,), whole, chicken, '1,...
1	Crispy Salt and Pepper Potatoes	['2 large egg whites', '1 pound new potatoes (...	Preheat oven to 400°F and line a rimmed baking...	crispy-salt-and-pepper-potatoes-dan-kluger	['2 large egg whites', '1 pound new potatoes (...	251	31	[[, '1, 2, large, egg, whites, '1, ,, '1, poun...
2	Thanksgiving Mac and Cheese	['1 cup evaporated milk', '1 cup whole milk', ...	Place a rack in middle of oven; preheat to 400...	thanksgiving-mac-and-cheese-erick-williams	['1 cup evaporated milk', '1 cup whole milk', ...	289	27	[[, '1, 1, cup, evaporated, milk, '1, ,, '1, 1, c...
3	Italian Sausage and Bread Stuffing	['1 (¾- to 1-pound) round Italian loaf, cut in...	Preheat oven to 350°F with rack in middle. Gen...	italian-sausage-and-bread-stuffing-240559	['1 (¾- 1-pound) round italian loaf, cut 1-inc...	558	34	[[, '1, 1, (, ¾-, 1-pound,), round, italian, 1...
4	Newton's Law	['1 teaspoon dark brown sugar', '1 teaspoon ho...	Stir together brown sugar and hot water in a c...	newtons-law-apple-bourbon-cocktail	['1 teaspoon dark brown sugar', '1 teaspoon ho...	218	12	[[, '1, 1, teaspoon, dark, brown, sugar, '1, ,, ...

7/20

```
In [18]: recipes['POS_Tagged_Ingredients'] = recipes['Tokenized_Ingredients'].apply(nltk.pos_tag)
         recipes.head()
```

Out[18]:

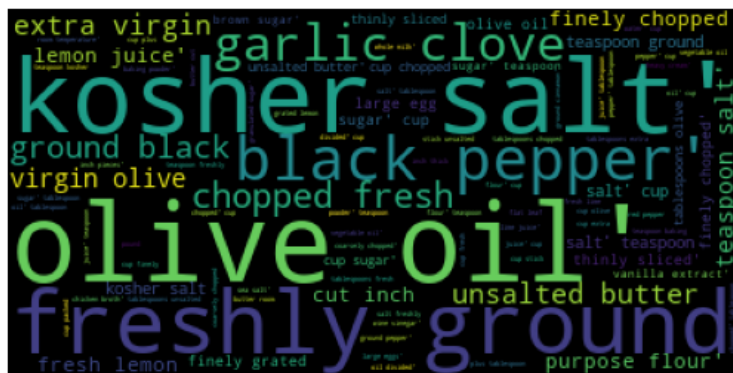
	Title	Ingredients	Instructions	Image_Name	Cleaned_Ingredients	num_ingredients	title_length	Tokenized_Ingredients	POS_Tagger
0	Miso-Butter Roast Chicken With Acorn Squash Pa...	['1 (3½-4-lb.) whole chicken', '2¼ tsp. kosher...	Pat chicken dry with paper towels, season all ...	miso-butter-roast-chicken-acorn-squash-panzanella	['1 (3½-4-lb.) whole chicken', '2¼ tsp. kosher...	828	54	[[, ' , 1, (, 3½-4-lb., .,), whole, chicken, ' , ...	[[([, NN), (' , ''
1	Crispy Salt and Pepper Potatoes	['2 large egg whites', '1 pound new potatoes (...	Preheat oven to 400°F and line a rimmed baking...	crispy-salt-and-pepper-potatoes-kluger	['2 large egg whites', '1 pound new potatoes (...	251	31	[[, ' , 2, large, egg, whites, ' , ,, ' , 1, poun...	[[([, NN) (larg
2	Thanksgiving Mac and Cheese	['1 cup evaporated milk', '1 cup whole milk', ...	Place a rack in middle of oven; preheat to 400...	thanksgiving-mac-and-cheese-erick-williams	['1 cup evaporated milk', '1 cup whole milk', ...	289	27	[[, ' , 1, cup, evaporated, milk, ' , ,, ' , 1, c...	[[([, NN), (' , '') N
3	Italian Sausage and Bread Stuffing	['1 (¾- to 1-pound) round Italian loaf, cut in...	Preheat oven to 350°F with rack in middle. Gen...	italian-sausage-and-bread-stuffing-240559	['1 (¾- 1-pound) round italian loaf, cut 1-inc...	558	34	[[, ' , 1, (, ¾-, 1-pound,), round, italian, l...	[[([, NN), (' , ''
4	Newton's Law	['1 teaspoon dark brown sugar', '1 teaspoon ho...	Stir together brown sugar and hot water in a c...	newtons-law-apple-bourbon-cocktail	['1 teaspoon dark brown sugar', '1 teaspoon ho...	218	12	[[, ' , 1, teaspoon, dark, brown, sugar, ' , ,, ...	[[([, NN) (teasp

Word Cloud

```
In [19]: from wordcloud import WordCloud

text = ' '.join(recipes['Cleaned_Ingredients'])
wordcloud = WordCloud().generate(text)

plt.imshow(wordcloud)
plt.axis('off')
plt.show()
```



Frequency Distribution


```
In [20]: all_words = ' '.join(recipes['Cleaned_Ingredients']).split()
fdist = FreqDist(all_words)
fdist.most_common(10)
```

```
Out[20]: [('1', 37955),
('cup', 28116),
('2', 21701),
('teaspoon', 15987),
('tablespoons', 15543),
('1/2', 13440),
('fresh', 10109),
('cups', 10075),
('1/4', 9645),
('1/2', 9173)]
```

NER (Named Entity Recognition)

```
In [21]: # NER (Named Entity Recognition) Function
# This function identifies notable entities within the given text.
# Specifically, it searches for Geopolitical entities (GPE), Personal names (PERSON), and Organizations
def named_entity_recognition(text):
    tree = ne_chunk(nltk.pos_tag(word_tokenize(text)))
    named_entities = []
    for subtree in tree.subtrees():
        if subtree.label() in ['GPE', 'PERSON', 'ORGANIZATION']:
            entity = " ".join([word for word, tag in subtree.leaves()])
            named_entities.append(entity)
    return named_entities

# Get the NER for the first 100 recipes
Named_Entities = recipes['Instructions'].head(100).apply(named_entity_recognition)
Named_Entities
```

```
Out[21]: 0    [Pat, Cut, Sprinkle, Set, Place, Pat, Place, T...
1                                     []
2                                [Cook, Sprinkle, Bake]
3                                [Brown, Whisk, Cooks]
4                                [Shake, Strain]
...
95    [Purée, Season, Mash, Cook, Toss, Sprinkle, Se...
96    [Cook, Set, Scrape, Pulse, Chill, Scrape, Spri...
97                                     [Carefully]
98    [Pat, Cook, Turn, Whisk, Dollop, Bake]
99    [Whisk, Arrange, Cook, Tear]
Name: Instructions, Length: 100, dtype: object
```

Translation

```
In [22]: # Translation Function
# This function translates the given text into the specified target language using the TextBlob library
# If the translation fails for any reason, it returns the original text.
def translate_text(text, target_lang='es'):
    try:
        blob = TextBlob(text)
        return blob.translate(from_lang='en', to=target_lang).string
    except:
        return text

# Applying the translation function to the first 100 titles in the recipes dataframe, translating to Spanish
spanish_recipe_title = recipes['Title'].head(100).apply(lambda x: translate_text(x, 'es'))
spanish_recipe_title
```

```
Out[22]: 0    Pollo asado miso-mordido con squash bellota pa...
1    Patatas crujientes de sal y pimienta
2    Acción de Gracias Mac and Cheese
3    Salchicha italiana y relleno de pan
4    Ley de Newton
...
95    Tazones de maíz y garbanzos con miso tahini
96    Tarta de crema de fruta de piedra
97    Plátanos Hornados con crema y queso (plátano h...
98    Tamale pastel con tomate y maíz frescos
99    Panzanella en el camino carbonizado con vinagr...
Name: Title, Length: 100, dtype: object
```

```
In [23]: # Applying the translation function to the first 100 instructions in the recipes dataframe
spanish_recipe_instructions = recipes['Instructions'].head(100).apply(lambda x: translate_text(x, 'es'))
spanish_recipe_instructions
```

```
Out[23]: 0    Pase el pollo seco con toallas de papel, sazon...
1    Precaliente el horno a 400 ° F y línea una ban...
2    Coloque una rejilla en el medio del horno; Pre...
3    Precaliente el horno a 350 ° F con rejilla en ...
4    Revuelva el azúcar morena y el agua caliente e...
...
95   Purée jalapeños, jengibre, ajo, cilantro, jugo...
96   Precaliente un horno a 375 ° F. Cocine la mant...
97   Precaliente el horno a 400 ° F.\nColoque los p...
98   Precaliente el horno a 450 ° F. Calienta una s...
99   Batir ajo, salmuera, mostaza y sal en un tazón...
Name: Instructions, Length: 100, dtype: object
```

Subjectivity

```
In [24]: # Subjectivity Analysis Function
# This function returns the subjectivity score of the given text.
# The score ranges from 0 (objective) to 1 (subjective).
def get_subjectivity(text):
    return TextBlob(text).sentiment.subjectivity

recipes['Instructions_Subjectivity'] = recipes['Instructions'].apply(get_subjectivity)
recipes.head()
```

1	and Pepper Potatoes	whites', '1 pound new potatoes (...)	400°F and line a rimmed baking...	and-pepper-potatoes-dan-kluger	'1 pound new potatoes (...)	251	31	[[, ', 2, large, egg, whites, ', ', ', 1, poun...	[[[, NN], ('', 'lar
2	Thanksgiving Mac and Cheese	['1 cup evaporated milk', '1 cup whole milk', ...	Place a rack in middle of oven; preheat to 400...	thanksgiving-mac-and-cheese-erick-williams	['1 cup evaporated milk', '1 cup whole milk', ...	289	27	[[, ', 1, cup, evaporated, milk, ', ', ', 1, c...	[[[, NN], ('', ')
3	Italian Sausage and Bread Stuffing	['1 (¼- to 1-pound) round Italian loaf, cut in...	Preheat oven to 350°F with rack in middle. Gen...	italian-sausage-and-bread-stuffing-240559	['1 (¼- 1-pound) round italian loaf, cut 1-inc...	558	34	[[, ', 1, (, ¼-, 1-pound,), round, italian, l...	[[[, NN], ('', ')
4	Newton's Law	['1 teaspoon dark brown sugar', '1 teaspoon ho...	Stir together brown sugar and hot water in a c...	newtons-law-apple-bourbon-cocktail	['1 teaspoon dark brown sugar', '1 teaspoon ho...	218	12	[[, ', 1, teaspoon, dark, brown, sugar, ', ', ...	[[[, NN], ('teasp

TF-IDF Weighting & Matrix Creation

```
In [25]: vectorizer = TfidfVectorizer(tokenizer=lambda x: x, lowercase=False, preprocessor=lambda x: x)
matrix = vectorizer.fit_transform(recipes['Tokenized_Ingredients'])
```

Cosine Similarity

```
In [26]: def compute_cosine_similarity(user_tokens):
user_vector = vectorizer.transform([user_tokens])
cosine_scores = cosine_similarity(user_vector, matrix)
recipes['cosine_score'] = cosine_scores[0]
return recipes
```

Recipe Recommendation Function

```

In [27]: def recommend_recipes(user_ingredients):
# Convert user ingredients to tokens
user_tokens = [ingredient.lower().strip() for ingredient in user_ingredients.split(',')]

# Compute cosine similarity scores for recipes
compute_cosine_similarity(user_tokens)

# Calculate count of exact ingredient matches
recipes['match_count'] = recipes['Tokenized_Ingredients'].apply(lambda x: sum(1 for token in user_t

# Return top 5 recipes with highest match count and then highest cosine similarity scores
top_recipes = recipes.sort_values(by=['match_count', 'cosine_score'], ascending=[False, False]).head(5)
return top_recipes[['Title', 'Ingredients', 'Instructions']]

In [28]: def display_recommendations(user_input):
# Get the recommended recipes
recommended = recommend_recipes(user_input)

# Print each recommended recipe in a formatted way
for index, row in recommended.iterrows():
    print("-" * 80) # Print a separator
    print(f"RECIPE ID: {index}")
    print(f"TITLE: {row['Title']}\n")

    # Display Ingredients
    print("INGREDIENTS:")
    ingredients = row['Ingredients']
    # Check if ingredients are a list or a string
    if isinstance(ingredients, list):
        for ingredient in ingredients:
            print(f"- {ingredient}")
    else:
        print(ingredients)
    print()

    # Display Instructions
    print("INSTRUCTIONS:\n", row['Instructions'], "\n")

print("-" * 80) # Print a separator at the end

```

Test the Refined Recommendation System

```
In [29]: user_input = input("Enter the ingredients you have (comma-separated): ")  
display_recommendations(user_input)
```

Enter the ingredients you have (comma-separated): chicken, onion

RECIPE ID: 8074

TITLE: Dr. Lee's Red Wine Chicken Stew

INGREDIENTS:

['1 tbsp olive oil', '2 cloves of garlic, mashed', '1/2 cup chopped tomatoes', '2 boneless chicken breasts', '1 cup sliced onion', '1 cup sliced mushrooms', '1 cup full-bodied red wine', '1 cup chicken stock (1 cup water plus 1/2 chicken or vegetable bouillon cube)', '1/4 tsp thyme']

INSTRUCTIONS:

In a Dutch oven at medium heat, combine olive oil and garlic and sauté briefly. Then add tomatoes and cook for five minutes. Add chicken breasts, onion, mushrooms, wine, chicken stock, and thyme. Lower heat and cook for 40 to 50 minutes until chicken breasts are done.

RECIPE ID: 8235

TITLE: Green Posole with Chicken

INGREDIENTS:

['2 tablespoons vegetable oil', '1 medium onion, finely chopped (about 1 cup)', '1 cup bottled green salsa or 1 recipe Mexican Tomatillo Salsa', '4 cups canned chicken broth or Chicken Stock', '1 rotisserie chicken, skin and bones discarded and meat shredded', 'Two 15-ounce cans white hominy, rinsed and drained', 'Kosher salt and freshly milled black pepper']

INSTRUCTIONS:

Heat the oil in a large saucepan over high heat until hot. Reduce the heat to medium, add the onion, and cook, stirring occasionally, until softened, about 5 minutes. Add the salsa and cook, stirring, for 5 minutes.

Add the chicken broth and simmer, partially covered, for 10 minutes. Add the chicken and hominy and simmer until heated through. Add salt and pepper to taste. To serve, ladle into bowls and let everyone garnish their own portion.

Finely chopped onion, chopped avocado, sliced radishes, chopped cucumbers, shredded napa cabbage, and tortilla chips.

RECIPE ID: 2946

TITLE: Easy Green Curry with Chicken, Bell Pepper, and Sugar Snap Peas

INGREDIENTS:

['1 tablespoon virgin coconut or canola oil', '1/4 cup green curry paste', '1 1/2 teaspoons freshly grated ginger', '1/2 teaspoon finely grated lime zest', '1 medium onion, sliced into 1/4-inch strips', '1 teaspoon kosher salt', '1 red bell pepper, sliced into 1/4-inch strips', '1 cup homemade chicken stock or low-sodium chicken broth', '4 cups 1/2-inch cubed cooked chicken (from 1 [2 1/2-pound] rotisserie chicken)', '1/2 pound sugar snap peas, trimmed (about 2 1/2 cups)', '1 (14-ounce) can coconut milk', '1/4 cup basil leaves', '1 tablespoon fresh lime juice', 'Cooked rice or rice noodles and lime wedges (for serving)']

INSTRUCTIONS:

Heat oil in a medium pot over medium. Cook curry paste, ginger, and lime zest, stirring, until heated through, about 1 minute. Add onion and salt and cook, stirring, until onion is soft, about 3 minutes. Add bell pepper and stir to combine. Add chicken stock and bring to a boil, then reduce heat to low and simmer until pepper is crisp-tender, about 3 minutes. Add chicken, snap peas, and coconut milk and cook over medium, being careful not to boil, until warmed through, about 5 minutes. Stir in basil and lime juice. Serve with rice or rice noodles and lime wedges alongside.

RECIPE ID: 10567

TITLE: Mini Chicken Pot Pies with Bacon and Marjoram

INGREDIENTS:

['5 applewood-smoked bacon slices', '1 1/2 cups chopped onion', '12 ounces peeled whole baby carrots (about 2 1/2 cups)', '1 8-ounce package trimmed haricots verts or other slender green beans, halved crosswise', '4 teaspoons chopped fresh marjoram', '1 3/4 cups low-salt chicken broth', '2/3 cup plus 1 tablespoon crème fraîche', '3 cups coarsely shredded chicken from 1 small purchased roasted chicken (skin removed)', '1 sheet frozen puff pastry (half of 17.3-ounce package), thawed']

INSTRUCTIONS:

Preheat oven to 450°F. Cook bacon in heavy large skillet over medium heat until crisp. Drain on paper towels. Chop bacon. Add onion to drippings in skillet; sauté until tender and golden, about 8 minutes. Add next 3 ingredients; stir 1 minute. Add broth; bring to boil over high heat. Reduce heat to medium high and boil until vegetables are almost tender and some liquid is reduced, about 8 minutes. Stir in 2/3 cup crème fraîche, chicken, and bacon. Bring to simmer. Season with pepper. Divide among four 2-cup soufflé dishes.

Unfold puff pastry onto work surface; roll out to 12-inch square. Cut into 4 equal squares. Top filling in soufflé dishes with pastry; fold edges down onto rims of dish. Brush top of crusts (not edges) with remaining 1 tablespoon crème fraîche. Cut small X in center of crusts; pierce all over with fork. B

ake until crusts are golden brown and filling is heated through, about 22 minutes.

RECIPE ID: 1193

TITLE: Instant Pot Chicken Stock

INGREDIENTS:

['2 lb. chicken parts', '1 lb. ground chicken', '1 medium onion, chopped', '1 medium carrot, peeled, c
hopped', '1 celery stalk, chopped', '1 sprig parsley', 'An electric pressure cooker or Instant Pot']

INSTRUCTIONS:

Combine chicken parts, ground chicken, onion, carrot, celery, parsley, and 4 cups water in cooker ins
ert. Lock on lid, making sure steam-release valve is in the proper sealed position. Select "Manual" an
d program for 1 hour at high pressure (it will take about 20 minutes for the pressure to build before
cooking automatically begins).

As soon as the time has elapsed, let pressure release naturally (this will take about 30 minutes). Str
ain stock through a fine-mesh sieve into a large glass measuring cup or airtight container. Let cool,
then skim fat off top before using.

```
In [30]: def display_recommendations_in_different_language(user_input, language='Spanish', language_code='es'):
# Get the recommended recipes
recommended = recommend_recipes(user_input)

# Print each recommended recipe in a formatted way
for index, row in recommended.iterrows():
    print("-" * 80) # Print a separator
    print(f"RECIPE ID: {index}")

    # Translate and display Title in Spanish
    translated_title = translate_text(row['Title'], language_code)
    print(f"TITLE (in {language}): {translated_title}\n")

    # Translate and display Ingredients in Spanish
    print(f"INGREDIENTS (in {language}):")
    ingredients = row['Ingredients']
    # Check if ingredients are a list or a string
    if isinstance(ingredients, list):
        for ingredient in ingredients:
            translated_ingredient = translate_text(ingredient, language_code)
            print(f"- {translated_ingredient}")
    else:
        translated_ingredients = translate_text(ingredients, language_code)
        print(translated_ingredients)
    print()

    # Translate and display Instructions
    translated_instructions = translate_text(row['Instructions'], language_code)
    print(f"INSTRUCTIONS (in {language}): \n", translated_instructions, "\n")

    print("-" * 80) # Print a separator at the end
```

```
In [31]: # Display ingredients in Spanish (default)
user_input = input("Enter the ingredients you have (comma-separated): ")
display_recommendations_in_different_language(user_input)
```


tén; Saltee hasta que estén tiernos y dorados, unos 8 minutos. Agregue los siguientes 3 ingredientes; Revuelva 1 minuto. Agregar caldo; Lleva a ebullición a fuego alto. Reduzca el fuego a mediano y hierva hasta que las verduras estén casi tiernas y se reduce algo de líquido, aproximadamente 8 minutos. Agregue 2/3 taza de crema fresca, pollo y tocino. Trae a fuego lento. Sazone con pimienta. Divide entre cuatro platos de soufflé de 2 tazas.

Desplegar hojaldre en la superficie de trabajo; Salir a un cuadrado de 12 pulgadas. Cortar en 4 cuadrados iguales. Top relleno en platos de soufflé con pasteles; Dobla los bordes hacia abajo en las llantas del plato. Cepille la parte superior de las costras (no los bordes) con 1 cucharada restante crème Fraîche. Cortar pequeño x en el centro de las costras; Pierce por todas partes con tenedor. Hornee hasta que las costras estén doradas y el relleno se calienta, aproximadamente 22 minutos.

RECIPE ID: 1193

TITLE (in Spanish): Caldo de pollo de olla instantánea

INGREDIENTS (in Spanish):

[illegible]

INSTRUCTIONS (in Spanish):

Combine piezas de pollo, pollo molido, cebolla, zanahoria, apio, perejil y 4 tazas de agua en el inserto de la olla. Bloquee la tapa, asegurándose de que la válvula de liberación de vapor esté en la posición sellada adecuada. Seleccione "Manual" y programa durante 1 hora a alta presión (la presión tardará aproximadamente 20 minutos antes de que la cocción comience automáticamente).

Tan pronto como haya transcurrido el tiempo, deje que la liberación de presión naturalmente (esto llevará unos 30 minutos). Colar Stock a través de un tamiz de malla fina en una taza de medición de vidrio grande o un recipiente hermético. Deje enfriar, luego escote la grasa de la parte superior antes de usarlo.

```
In [32]: # Display ingredients in German
user_input = input("Enter the ingredients you have (comma-separated): ")
display_recommendations_in_different_language(user_input, 'German', 'de')
```

Enter the ingredients you have (comma-separated): chicken, onion

RECIPE ID: 8074

TITLE (in German): Dr. Lees Rotweinhuhnneintopf

INGREDIENTS (in German):

['1 EL Olivenöl', '2 Nelken Knoblauch, Püree', '1/2 Tasse gehackte Tomaten', '2 Hähnchenbrust mit Knochen ohne Knochen', '1 Tasse geschnittene Zwiebel', '1 Tasse geschnittene Pilze', '1 Tasse Vollmundiger Rotwein', '1 Tasse Hühnerbrühe (1 Tasse Wasser plus 1/2 Hühnchen oder Gemüse-Bouillon-Würfel)', '1/4 TL Thymian']

INSTRUCTIONS (in German):

In einem niederländischen Ofen bei mittlerer Hitze Olivenöl und Knoblauch kombinieren und kurz anbraten. Dann Tomaten hinzufügen und fünf Minuten kochen. Fügen Sie Hühnerbrust, Zwiebeln, Pilze, Wein, Hühnerbrühe und Thymian hinzu. Senken Sie die Hitze und kochen Sie 40 bis 50 Minuten, bis Hühnerbrust durchgeführt werden.

RECIPE ID: 8235

TITLE (in German): Grüne Posole mit Hühnchen

INGREDIENTS (in German):

['2 Esslöffel Gemüseöl', '1 mittlere Zwiebel, fein gehackt (ca. 1 Tasse)', '1 Tasse Flaschen grüne Salsa oder 1 Rezept Mexikanische Tomatillo -Salsa', 4 Tasse Hühnerbrühe in Konserven oder Hühnchen, '1 Rotisserie Huhn, Haut und Knochen weggeworfen und Fleisch zerkleinert', 'zwei 15-Unzen-Dosen weiße Hominy, gespült und abgelassenes koscheres Salz und frisch gemahlener schwarzer Pfeffer']

INSTRUCTIONS (in German):

Erhitzen Sie das Öl in einem großen Topf bei starker Hitze, bis sie heiß sind. Reduzieren Sie die Hitze auf Medium, fügen Sie die Zwiebel hinzu und kochen Sie unter gelegentlichem Rühren etwa 5 Minuten lang, bis sie weich sind. Fügen Sie die Salsa hinzu und kochen Sie sie unter Rühren 5 Minuten lang. Fügen Sie die Hühnerbrühe und köcheln Sie 10 Minuten lang teilweise abgedeckt. Fügen Sie das Huhn und die Hominy hinzu und köcheln Sie, bis er erhitzt wird. Fügen Sie zum Geschmack Salz und Pfeffer hinzu. Zum Servieren in Schalen schöpfen und alle ihre eigene Portion garnieren lassen. Fein gehackte Zwiebel, gehackte Avocado, Splitterradieschen, gehackte Gurken, zerkleinertes Napa-Kohl und Tortilla-Chips.

RECIPE ID: 2946

TITLE (in German): Einfach grünes Curry mit Hühnchen, Paprika und Zuckerschnapperbsen

INGREDIENTS (in German):

['1 Esslöffel jungfräuliche Kokosnuss oder Rapsöl', '1/4 Tasse grüne Curry-Paste', '1 1/2 Teelöffel frisch geriebener Ingwer', '1/2 Teelöffel fein gerieben 1/4-Zoll-Streifen', '1 Teelöffel koscheres Salz', '1 roter Paprika, in 1/4-Zoll-Streifen geschnitten', '1 Tasse hausgemachte Hühnerbrühe oder Niedrig-sodium-Hühnerbrühe', 4 Tassen 1///4 Tassen 2-Zoll-Kochhuhn (von 1 [2 1/2-Pfund] Rotisserie-Hühnchen), '1/2 Pfund Zucker-Schnappbauern, beschnitten (ca. 2 1/2 Tassen)', '1 (14-Unbezirk) können Kokosmilch', '1/4 Tasse Basilikumblätter', '1 Esslöffel frischer Limettensaft', 'gekochte Reis oder Reinsnudeln und Limettenkeile (zum Servieren)']

INSTRUCTIONS (in German):

Öl in einem mittleren Topf über Medium erhitzen. Curry-Paste, Ingwer und Limettenschale unter Rühren, bis er durch den Erhitzen aufgeheizt ist, ca. 1 Minute kochen. Fügen Sie Zwiebel und Salz hinzu und kochen Sie unter Rühren, bis die Zwiebel weich ist, ca. 3 Minuten. Paprika hinzufügen und rühren, um zu kombinieren. Fügen Sie Hühnerbrühe hinzu und kochen Sie sie zum Kochen. Reduzieren Sie dann die Hitze auf niedrig und köcheln Sie, bis Pfeffer ca. 3 Minuten knusprig ist. Fügen Sie Hühnchen, Schnappböhnchen und Kokosmilch hinzu und kochen Sie über Medium, und achten Sie darauf, dass Sie nicht kochen, bis sie sich durchgewärmt haben, etwa 5 Minuten. Basilikum und Limettensaft einrühren. Mit Reis- oder Reinsnudeln und Kalkschnitzen nebeneinander servieren.

RECIPE ID: 10567

TITLE (in German): Mini Chicken Pot Pies mit Speck und Marjoram

INGREDIENTS (in German):

['5 Applewood-Raucher-Speckscheiben', '1 1/2 Tassen gehackte Zwiebel', '12 Unzen geschälte ganze Babykartoffeln (ca. 2 1/2 Tassen)', '1 8-Unzen-Paket, die Haricots Verts oder andere schräge Grün geschnittene Paket begegnet Bohnen, halbiert kreuzweise', '4 Teelöffel gehackt frische Marjoram', '1 3/4 Tassen Low-Salt-Hühnchenbrühe', '2/3 Tasse plus 1 Esslöffel Crème Fraîche', '3 Tassen grob zerkleinertes Hühnchen von 1 kleinen gekauften gerösteten Gebraten Huhn (Haut entfernt)', '1 Blatt gefrorener Blätterteig (Hälfte von 17,3 Unzen Paket), aufgetaut']

INSTRUCTIONS (in German):

Ofen auf 450F vorheizen. Speck in schwerer großer Pfanne bei mittlerer Hitze bis knusprig kochen. Auf Papiertüchern abtropfen lassen. Speck hacken. Fügen Sie Dripeln in der Pfanne Zwiebel hinzu; bis zart und golden anbraten, ca. 8 Minuten. Fügen Sie die nächsten 3 Zutaten hinzu; 1 Minute umrühren. Brühe hinzufügen; Über hohe Hitze zum Kochen bringen. Die Hitze auf mittleres reduzieren und kochen, bis das

Gemüse nahezu zart ist und die Flüssigkeit ca. 8 Minuten reduziert ist. 2/3 Tasse Crème Fraîche, Hühner und Speck einrühren. Zum Kochen bringen. Pfeffern. Teilen Sie sich auf vier 2-Tassen-Soufflé-Gerichte.

Blätterteig auf die Arbeitsfläche entfalten; Rollen Sie auf das 12-Zoll-Quadrat aus. In 4 gleiche Quadrate geschnitten. Top –Füllung in Soufflé –Gerichten mit Gebäck; Die Kanten auf Schalenfeldern falten. Pinsel der Krusten (nicht Kanten) mit restlichen 1 Esslöffel Crème Fraîche. Schneiden Sie kleines X in der Mitte der Krusten; Überall mit Gabel durchdringen. Backen Sie, bis die Krusten goldbraun sind und die Füllung durch etwa 22 Minuten erhitzt wird.

RECIPE ID: 1193

TITLE (in German): Instant Pot Hähnchenbrühe

INGREDIENTS (in German):

['2 lb. Hühnerteile', '1 lb. gemahlenes Hähnchen', '1 mittlere Zwiebel, gehackt', '1 mittlere Karotte, geschält, gehackt', '1 Selleriestiel, gehackt', '1 Sprig –Petersilie', ',', ',', ',', ',', ',', ',', '1 Selleriestiel, '1 Sprig –Petersilie', ',', ',', ',', '1 Selleriestiel Ein elektrischer Schnellkochtopf oder Instant Pot ']

INSTRUCTIONS (in German):

Kombinieren Sie Hühnerteile, gemahlene Hähnchen, Zwiebeln, Karotten, Sellerie, Petersilie und 4 Tassen Wasser im Herdeinsatz. Sperren Sie den Deckel und achten Sie darauf, dass sich das Dampfveröffentlichungsventil in der richtigen versiegelten Position befindet. Wählen Sie „Handbuch“ und programmieren Sie 1 Stunde lang mit hohem Druck (es dauert ungefähr 20 Minuten, bis der Druck zum Aufbau automatisch beginnt).

Sobald die Zeit verstrichen ist, lassen Sie den Druck auf natürliche Weise freigegeben (dies wird ungefähr 30 Minuten dauern). Lagerbestand durch ein feines Mesh-Sieb in einen großen Glasmessbecher oder luftdicht. Lassen Sie abkühlen und dann vor dem Gebrauch Fett ab.

Summary

The development phases and features of the project are:

- Exploratory Data Analysis (EDA): This phase involved cleaning and visualizing the data to understand its structure and distribution.
- Natural Language Processing (NLP): Implementation of various NLP techniques like tokenization, POS tagging, and NER to preprocess and analyze the recipe data.
- Visualization: Created histograms, scatter plots, and word clouds to offer visual insights into the dataset.
- Translation: Integrated a translation function to offer multilingual support, with initial support for Spanish.
- Recommendation System: Utilized TF-IDF and cosine similarity to match user-input ingredients with recipes and rank them by relevance.
- Interactive Testing: The user can input ingredients they possess and obtain top recipes either in many languages (as demonstrated with English, Spanish, and German)

Conclusion

The Culinary Matcher successfully bridges the gap between users' available ingredients and potential gourmet creations they can whip up. Its emphasis on relevance and multilingual support ensures a broad audience appeal, giving users not only the chance to try out new recipes but also to learn and appreciate them in a language they're comfortable with. The system's modular design means future expansions could include support for additional languages, user profiles with preferences, and integration with online grocery stores for ingredient purchase.

In []: