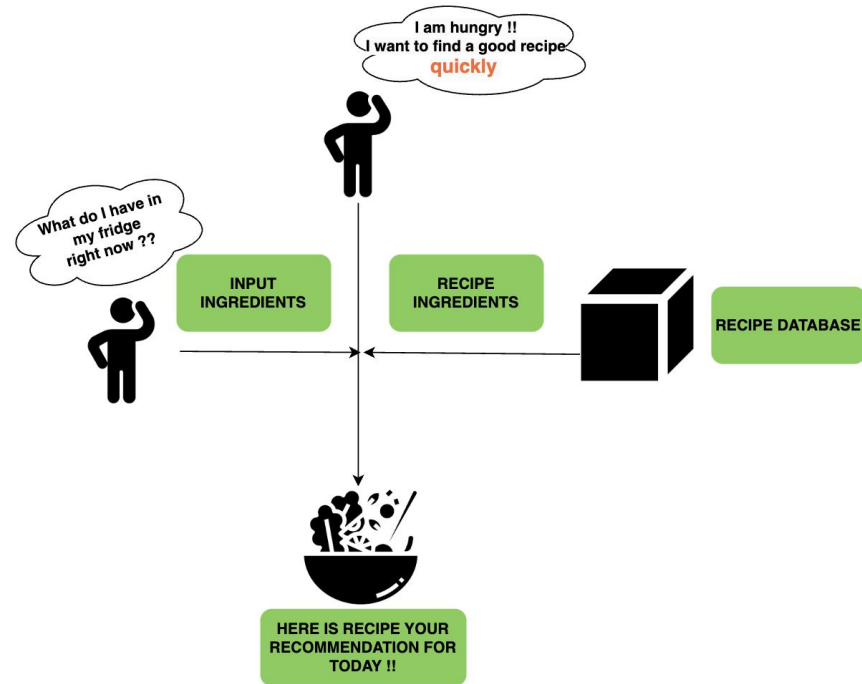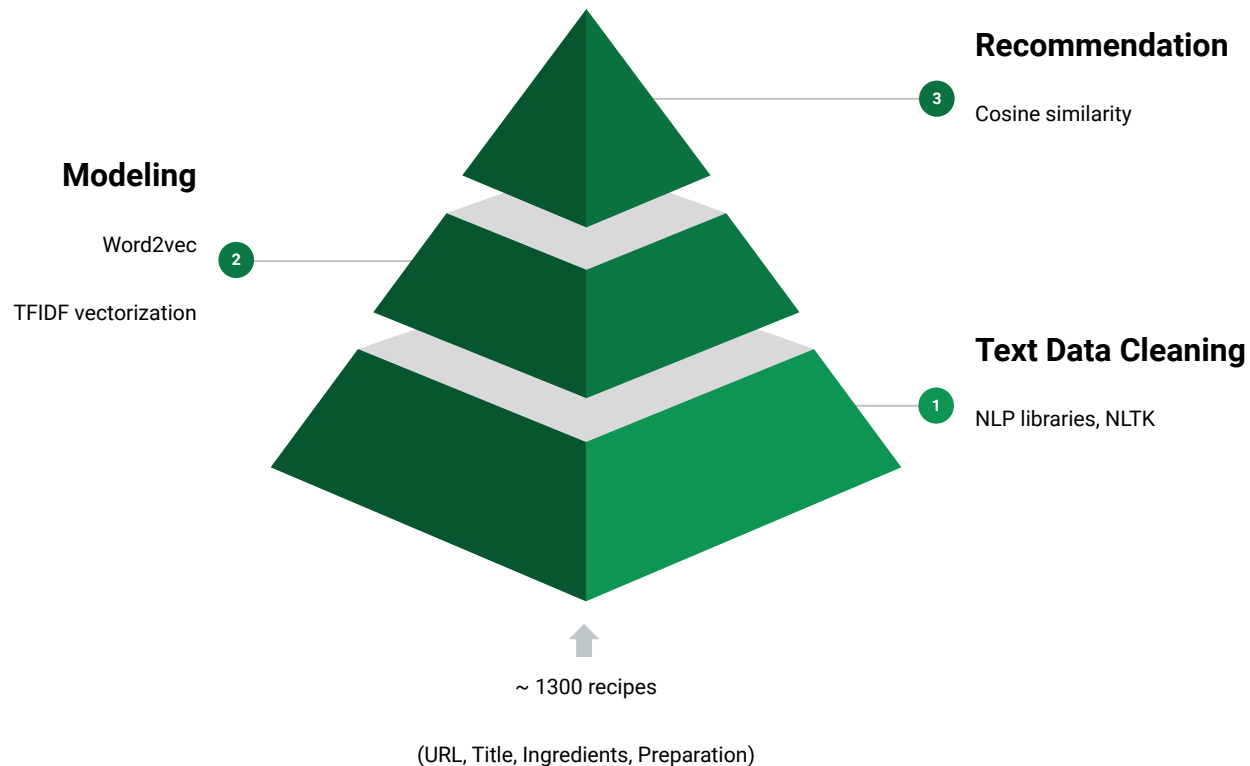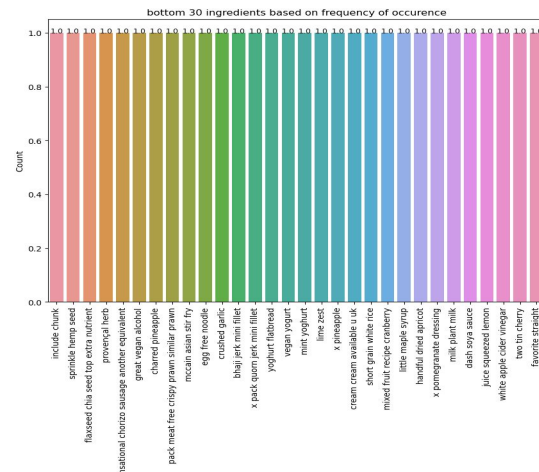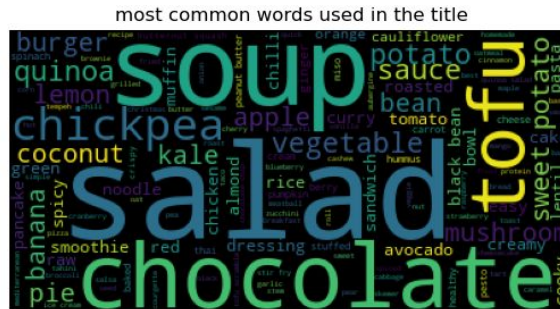# VegVista

~ a happy vegan ~

Anupa Jayakody

PROJECT OVERVIEW & IMPACT

DATASET & PRE PROCESSING

Recommendation

3 Cosine similarity

Modeling

Word2vec 2

TFIDF vectorization

Text Data Cleaning

1 NLP libraries, NLTK

~ 1300 recipes

(URL, Title, Ingredients, Preparation)
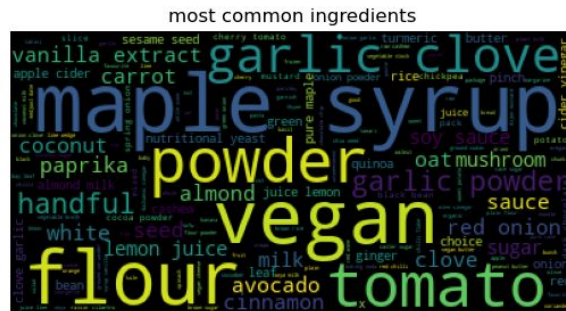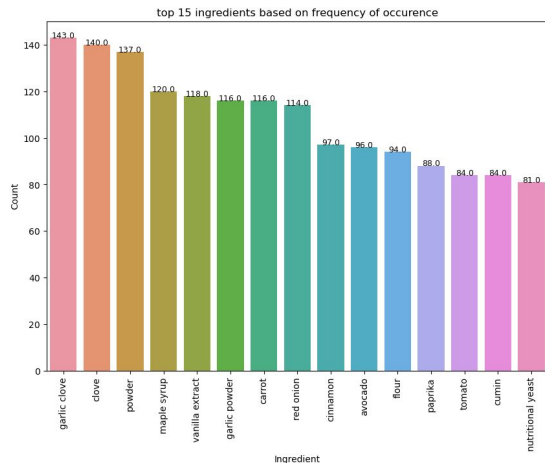
EDA & FINDINGS

EDA & FINDINGS

Average length of recipes is between 5-8

Most no.of recipes titles are of length 3 words and this has a resemblance of the length of the ingredients too.

MODELING & EVALUATION

Word embeddings for each ingredient list for the recipes —>

Recommendations based similarity between the input and database embeddings

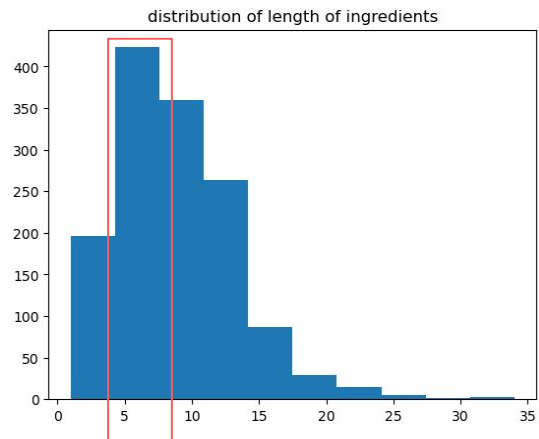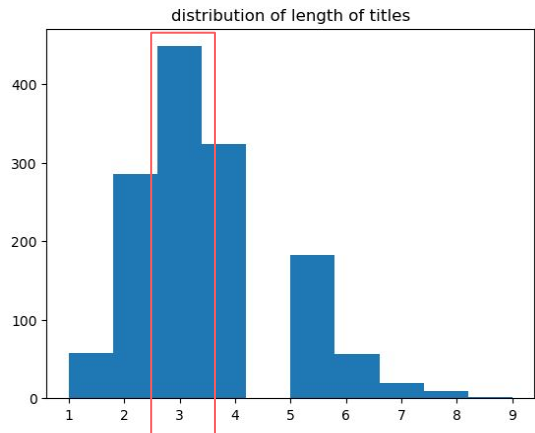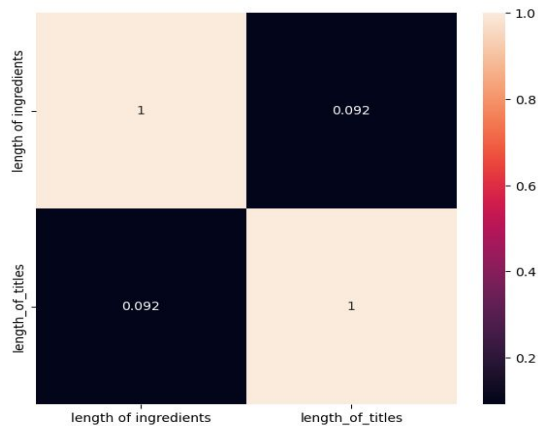Parts of speech tagging in creating the embeddings vs individual tokens of words

"onion" embedding vector

"sauce" embedding vector

"fish" embedding vector

"chicken" embedding vector

"chicken" embedding vector

"chickpea" embedding vector

```
phrases_model.wv.most_similar('onion', topn=10)
✓ 0.1s

[('olive_oil', 0.9998865723609924),
 ('garlic_clove', 0.999873161315918),
 ('black_bean', 0.9997977018356323),
 ('tomato', 0.9997972846031189),
 ('red_pepper', 0.999787449836731),
 ('flake', 0.9997850656509399),
 ('thyme', 0.9997355341911316),
 ('clove', 0.9997222423553467),
 ('onion_powder', 0.9997029304504395),
 ('red_onion', 0.99969881772995)]
```

```
model_1.wv.most_similar('onion', topn=10)
✓ 0.0s

[('red', 0.9987379908561707),
 ('clove', 0.9974899888038635),
 ('pepper', 0.9966903328895569),
 ('bell', 0.9964130520820618),
 ('garlic', 0.9948572516441345),
 ('broth', 0.9943221211433411),
 ('tomato', 0.9901957511901855),
 ('cumin', 0.9878203272819519),
 ('oregano', 0.9874938726425171),
 ('yellow', 0.9865705966949463)]
```

MODELING & EVALUATION


```
input_ingredients= 'tofu', 'onion', 'spinach'

recommendation_working(input_ingredients)
```

	title	ingredients	score
0	Tofu Chorizo and Lemon Cauliflower Rice	For the tofu:,400g tofu, well pressed for at l...	1.0000000000000002
1	Chickpea and Sweet Potato Tikka Masala with Gr...	2 handfuls of baby spinach,1 green chilli, fin...	1.0000000000000002
2	Ciambotta (Vegetable Stew)	1 tablespoon oil or water, for sautéing,1 larg...	0.9999999999999996
3	Vanilla Date Smoothie	Scale,1x2x3x ,1 cup unsweetened almond milk (u...	0.9999999999999996
4	APPLE, BEET, CARROT & KALE SALAD	Units,USM ,Scale,1x2x3x ,1 large apple (green ...	0.9999999999999996

```
input_ingredients= 'tofu'

recommendation_working(input_ingredients)
```

	title	ingredients	score
0	Super Vegetable Salad	28 cherry tomatoes,3 florets of broccoli,12 gr...	1.0
1	Sweet Porridge with Nuts, Seeds and Dried Fruit	3 tablespoons of porridge oats,1/2 mug of wate...	1.0
2	Chocolate and Maple Ice Cream	480ml almond milk ,140g pecans, soaked for 2 h...	1.0
3	Cottage Pie Cobbler	· 2 tsp olive oil,· 1 red onio...	1.0
4	Wicked Easy Black Pepper + Sriracha Tofu	Oil for the pan,1 block firm tofu, cubed,1 Tbs...	1.0

Challenges

1. High cosine similarity results
2. Different results when the same ingredient added alone vs when combined with more

MODELING EXT STEPS

1. If no data found for the input ingredients, what??

2. Recommendations based on the title and ingredients both ??

3. Possibility to predict the time taken to make the recipe??

PRODUCTIZATION

4. Creating an application interface to the user

What do I have in my fridge right now ??

Tofu, Potatoes, Onion

1. Turkish Tofu & Spinach Börek
2. Tikka Smoked Tofu Skewers with Apple Slaw
3. Recipe 3
4. Recipe 4
5. Recipe 5