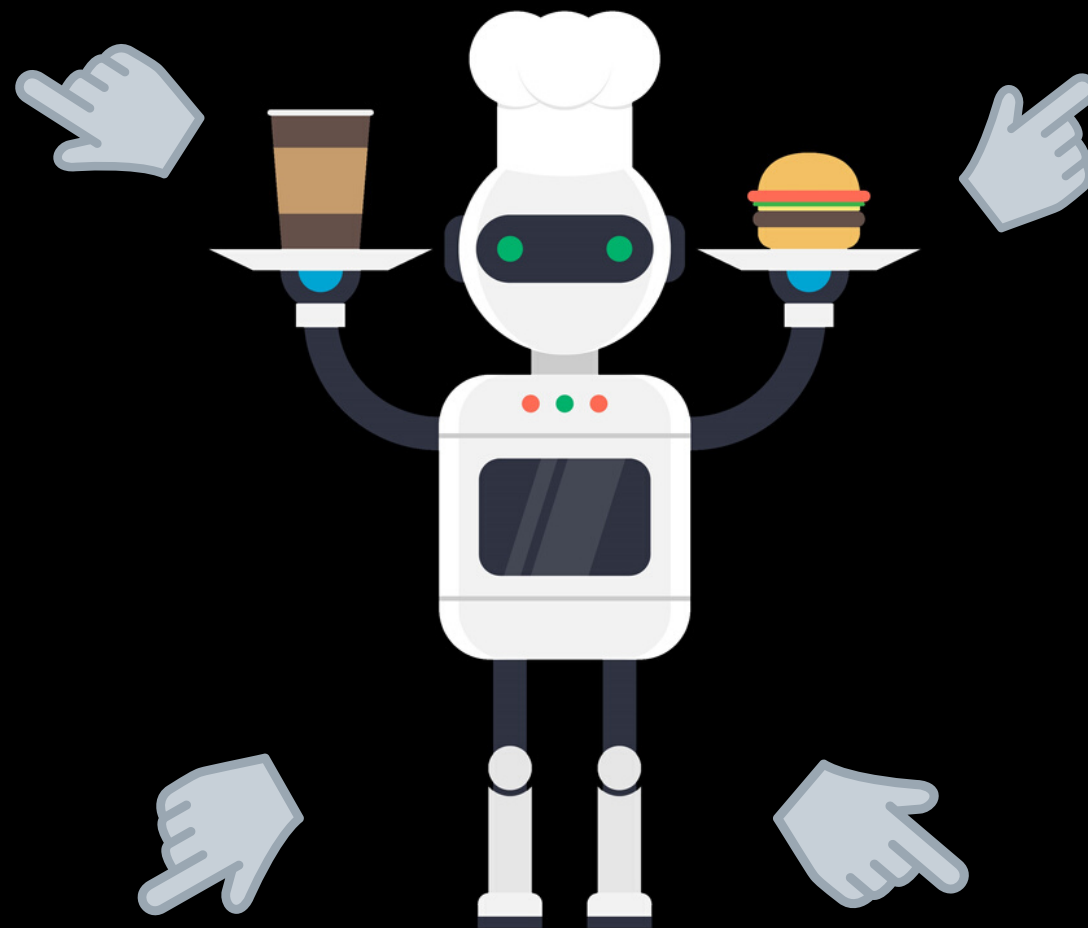# SENTIMENT ANALYSIS OF CUSTOMER REVIEWS

SUMMARIZING THE CUSTOMER REVIEWS TO HELP CUSTOMERS CHOOSE A CASUAL DINING RESTAURANT

Akshay Bansod(A19008)   Pratik Chakravorty(A19023)   Sreetama Das(A19037)

Hitesh Relwani(A19013)   Satendra Nath(A19027)   Vivek Kumar(A19041)

WHY THIS ?

- Risk Reduction

- Search time reduction

- Buyer's remorse

- Group influence

Project Background

Continued

Search..

Barbeque Nation

User Rating

★★★★★
4.2 average based on 234 reviews.

5 star		87
4 star		106
3 star		26
2 star		7
1 star		3

Most Talked Words

Feature Score

0

Continued

The Positives

food

3.0

buffet

3.21

experience

3.2

The Negatives

Continued

Sentiment Distribution of Features
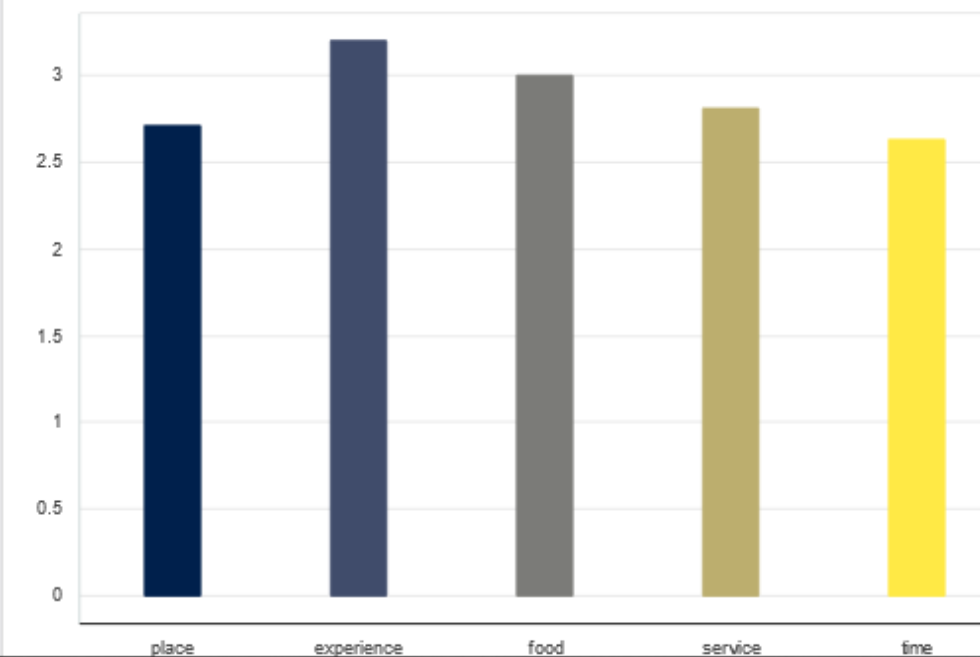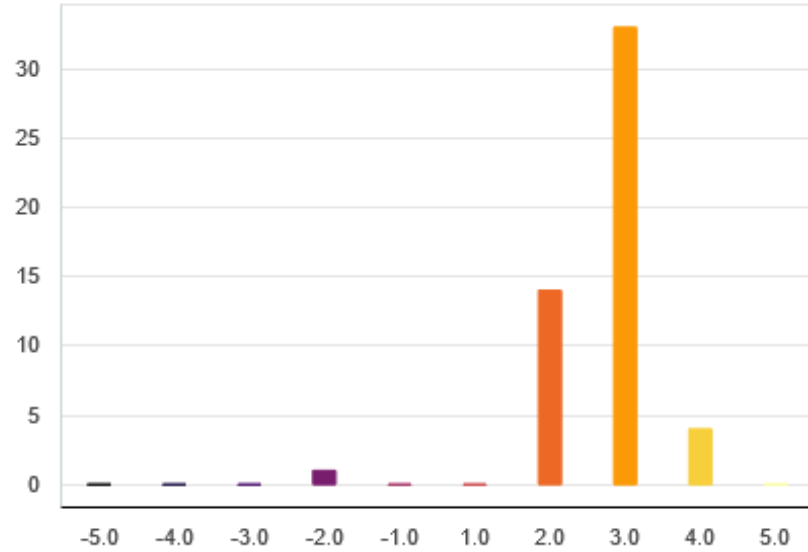
Score Distribution for place

Score Distribution for experience
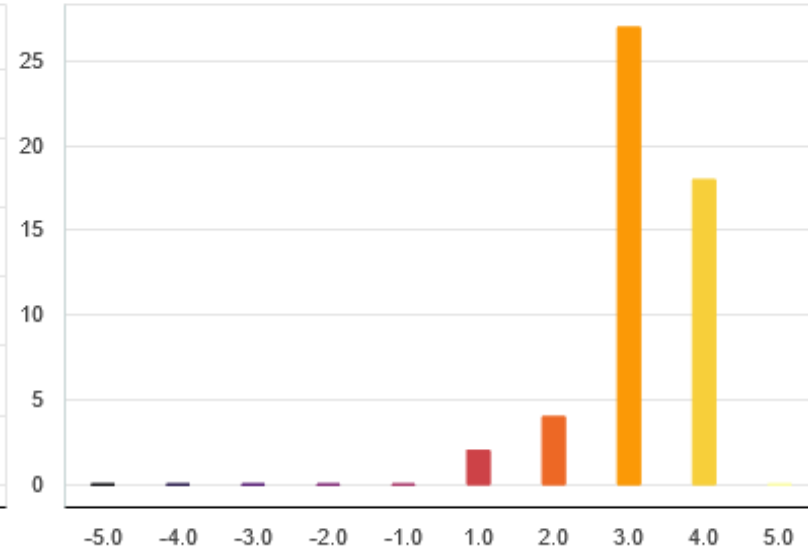
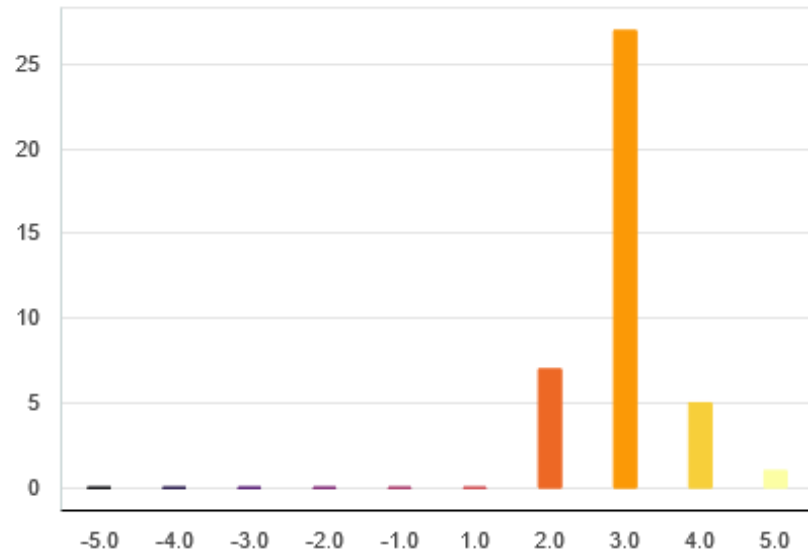### Score Distribution for food

Score Distribution for service

Score Distribution for time

Continued

Home

About Us

Recommended Restaurants ▼

Similar Restaurants

Known For Biryani

Known For Ambience

Known For Service

Search..

87

106

26

7

3

Feature Score

3

2.5

2

1.5

quality

taste

bangalore

outlet
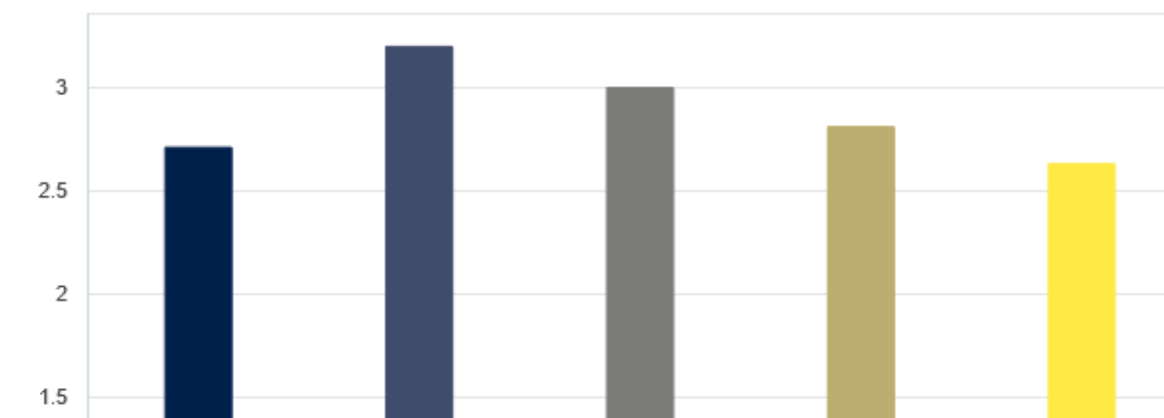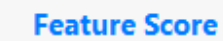
Data Collection

- The dataset used is "Zomato Bangalore Restaurants"
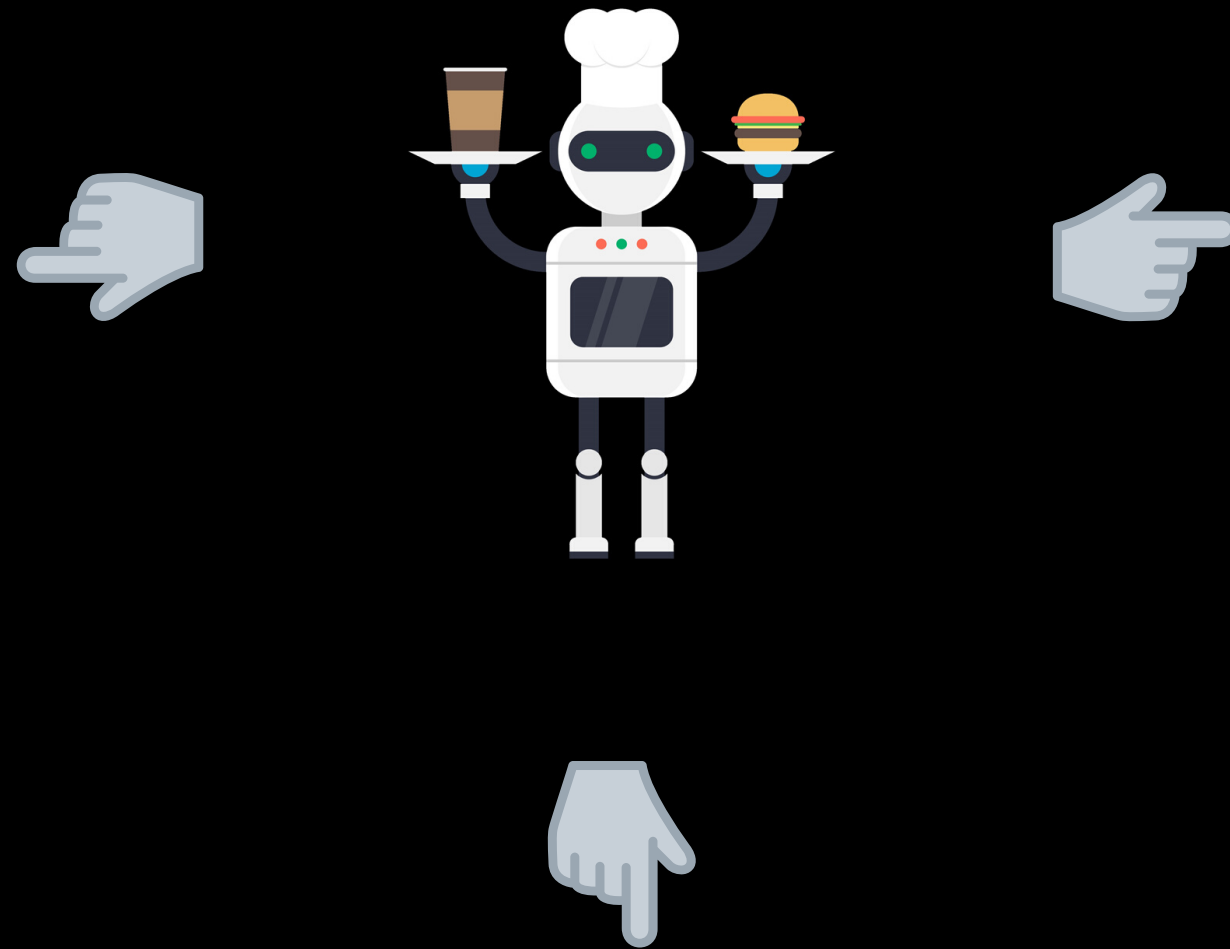
- Columns used were Restaurant Names, rating and the reviews

- The casual dining segment has around 40K rows satisfying our requirements

Process

Symspell 👉 Lemmatization 👉 Deep Segment 👉 Feature Extraction

👉 Dependency Parsing 👉 Sentiment Score

Topic Modelling

Symspell

- Compound splitting and decompounding
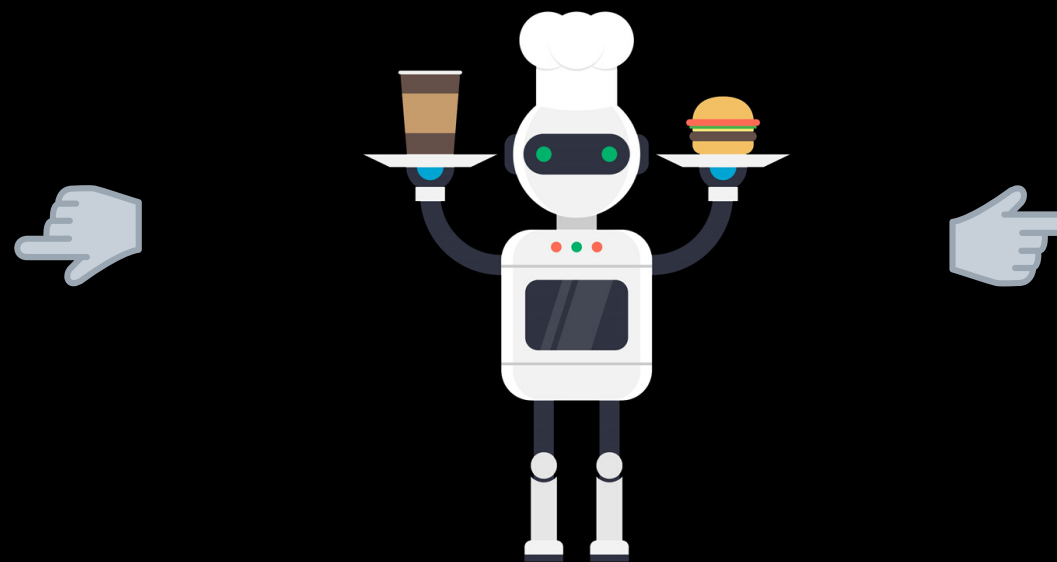
- Automatic spelling correction

```python
review = 'The food was gr8 but specially the biryni was amazing , the staff was courtious and ambience was wonderful'
slang_corr = slang_trans(review)
corr_string = spell_checker_eng_word(slang_corr)
print("Incorrect English:",review)
print("Corrected English:",corr_string)
```

```
Incorrect English: The food was gr8 but specially the biryni was amazing , the staff was courtious and ambience was wonderful
Corrected English: the food was great but specially the biryani was amazing , the staff was courteous and ambience was wonderful
```

Lemmatization

- It involves resolving words to their dictionary form
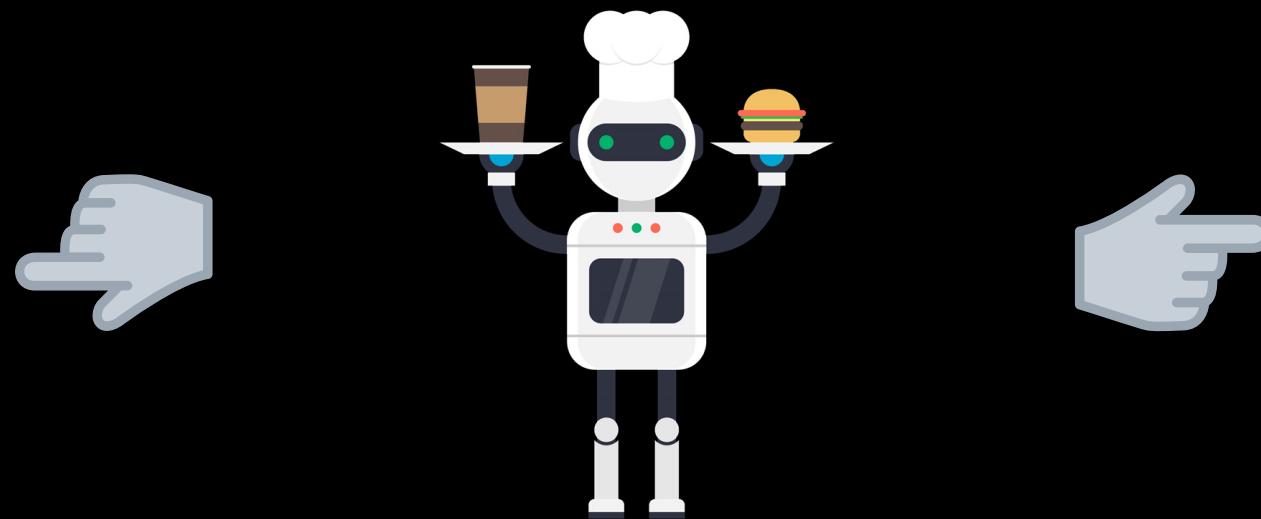
- To get the best results, we have to feed the part of speech tags to the lemmatizer

```python
lemmatize_sentence(sentence):
    nltk_tagged = nltk.pos_tag(nltk.word_tokenize(sentence))
    wn_tagged = map(lambda x: (x[0], nltk2wn_tag(x[1])), nltk_tagged)
    res_words = []
    for word, tag in wn_tagged:
        if tag is None:
            res_words.append(word)
        else:
            res_words.append(lemmatizer.lemmatize(word, tag))
    return " ".join(res_words)
```

Deep Segment

- Performs sentence segmentation on the unpunctuated text
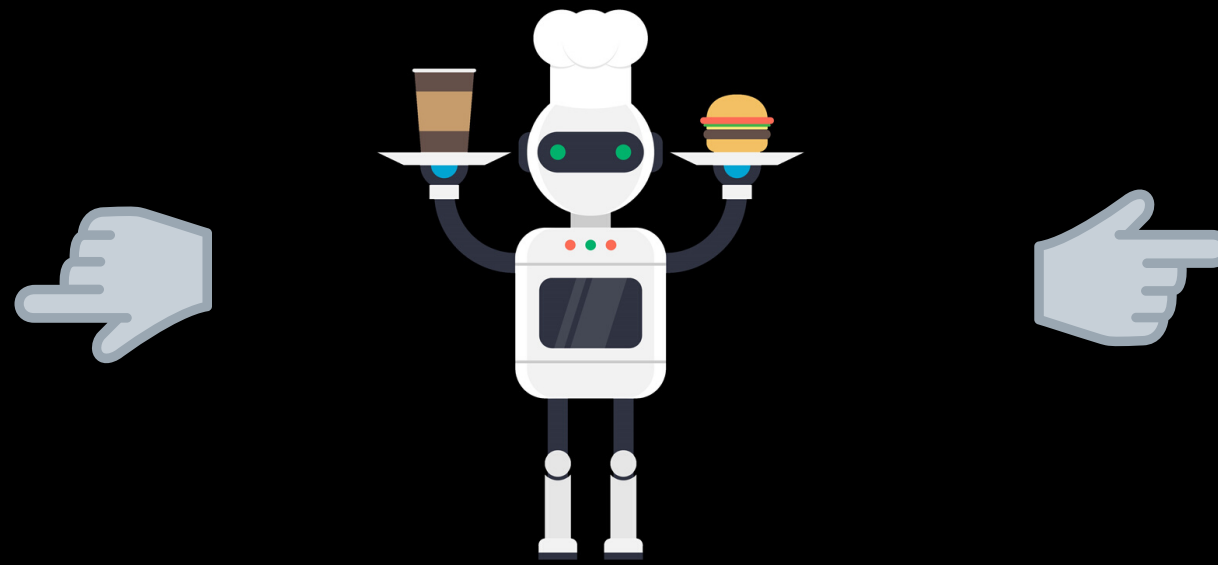
- Uses seq2seq for punctuation correction at a sentence level

```
Input : a very nice and good place with best buffet in town The staff be very humble and polite
Output: ['a very nice and good place with best buffet in town', 'The staff be very humble and polite']
```

Feature Extraction

- Feature Extraction i.e Chunking computes the root of the phrase

```
Input : the food is good and the biryani is awesome
Noun Chunks : [the food, the biryani]
Cleaned Chunks : ['food', 'biryani']
```

- Extraction of top noun chunks and their associated frequencies

```
[('main course', 6),
 ('starters', 5),
 ('options', 4),
 ('staff', 4),
 ('desserts', 4),
 ('food', 4),
 ('quality', 3),
```

Dependency Parsing

- To get proper association of the noun/noun phrases with the adverbs and/or adjectives(whether it be positiveor negative)/ or verb
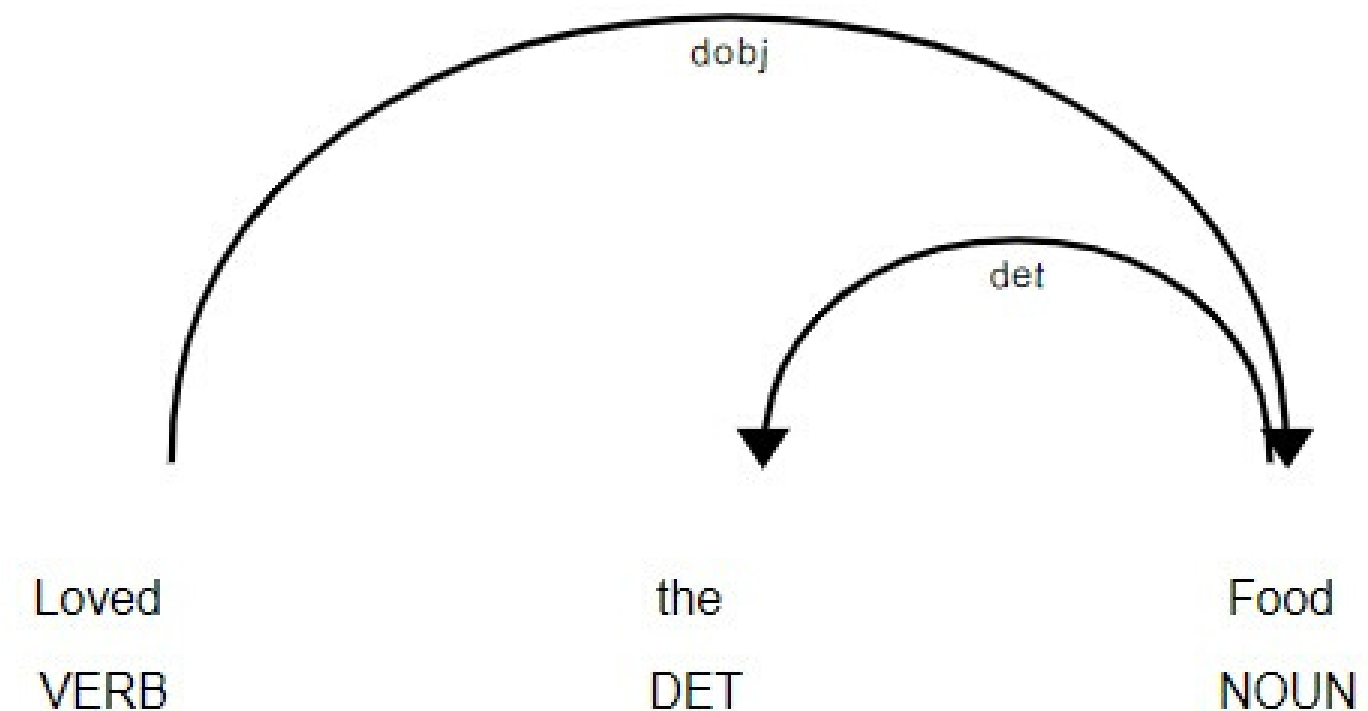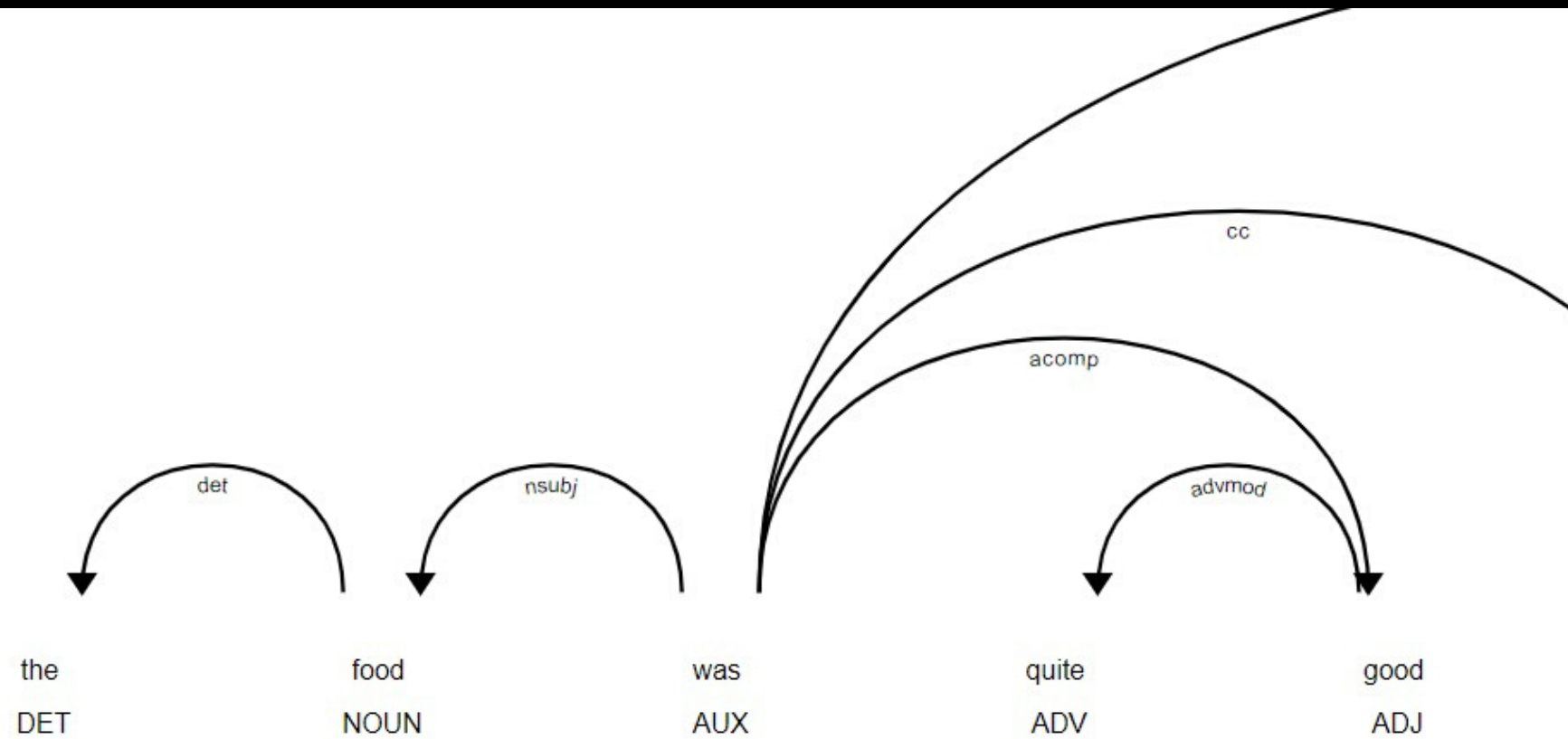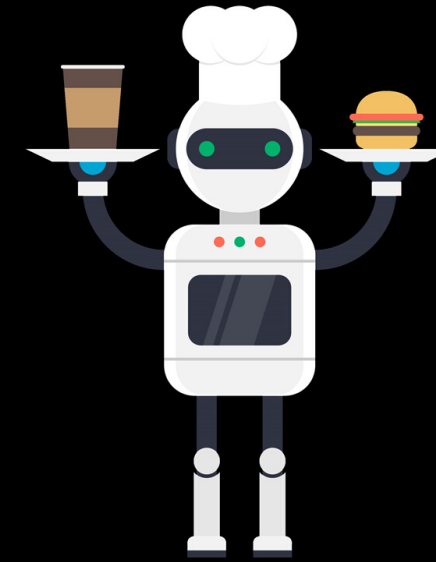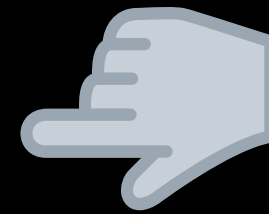
The dictionary which has its key as the features and values as
the terms describing the features.

```
{'staff': [' courteous', ' extremely polite', ' very polite', 'friendly'],
 'main course': ['avoided',
  'not acceptable',
  ' disappointing',
  'not that good'],
 'options': ['courteous', 'consider', 'nice', 'many'],
 'service': ['courteous',
  ' very courteous',
  'Absolutely loved',
  ' upto',
  ' disappointing'],
```

Sentiment Score

- The AFINN lexicon has been used for sentiment analysis
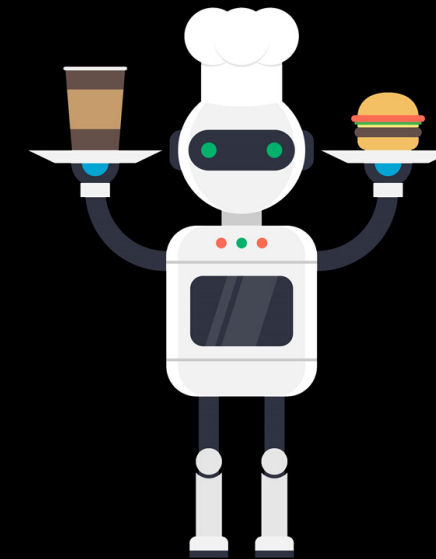
Restaurant	Word	Mean	Freq	-5.0	-4.0	-3.0	-2.0	-1.0	1.0	2.0	3.0	4.0	5.0
Barbeque Nation	place	2.75	80	0	0	1	0	1	2	13	56	7	0
Barbeque Nation	taste	1.22	18	0	0	1	5	0	0	3	7	2	0
Barbeque Nation	order	2	3	0	0	0	0	0	1	1	1	0	0
Barbeque Nation	main course	0.6	5	0	0	0	2	0	1	0	2	0	0
Barbeque Nation	service	1.09	47	0	0	7	9	0	0	3	28	0	0
Barbeque Nation	ambiance	2.5	16	0	0	0	0	0	1	6	9	0	0
Barbeque Nation	restaurant	2.15	13	0	0	1	0	0	0	5	7	0	0

Topic Modelling

- It is the process of identifying topics in a set of documents


```
┌──────────┐     ┌────────────────────┐     ┌──────────┐     ┌────────────────────┐
│ reviews  │ ──▶ │ positive (4 & 5-stars) │ ──▶ │ tokenize │ ──▶ │ topic modeling (+) │ ──┐
└──────────┘     └────────────────────┘     └──────────┘     └────────────────────┘   │
                                                                                        │
        ┌────────────────────────────────┐              ┌──────────────────────────┐   │
        │ distribution score for each topic for │ ◀───── │ topics for positive reviews │ ◀─┘
        │ each restaurant                 │              └──────────────────────────┘
        └────────────────────────────────┘
            │
        ┌──────────────────────────┐         ┌──────────────────────┐
        │ clustering on the basis of │ ──────▶ │ recommendations      │
        │ distribution score        │         └──────────────────────┘
        └──────────────────────────┘
```
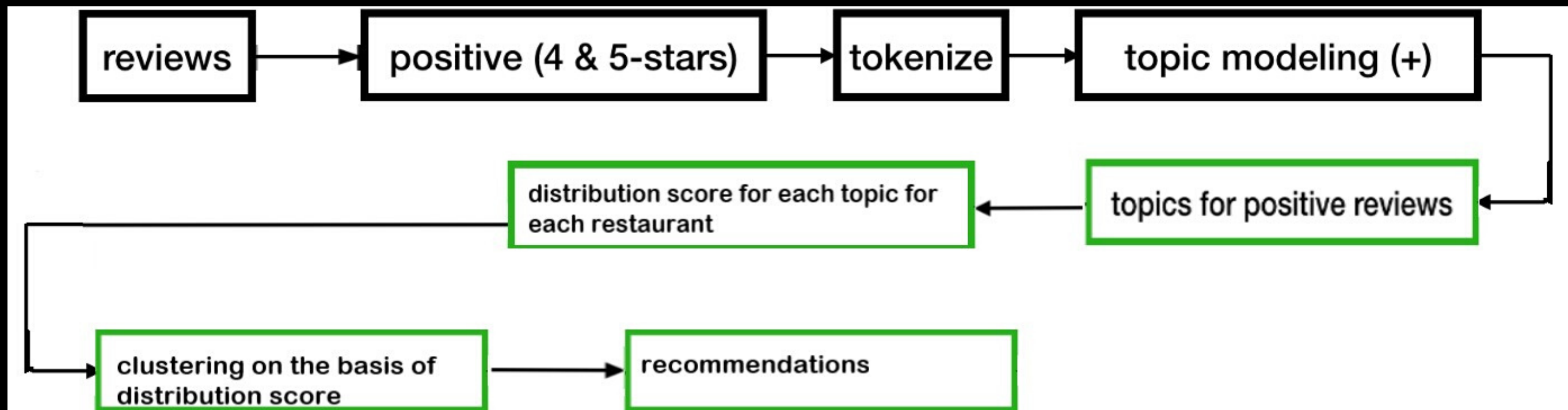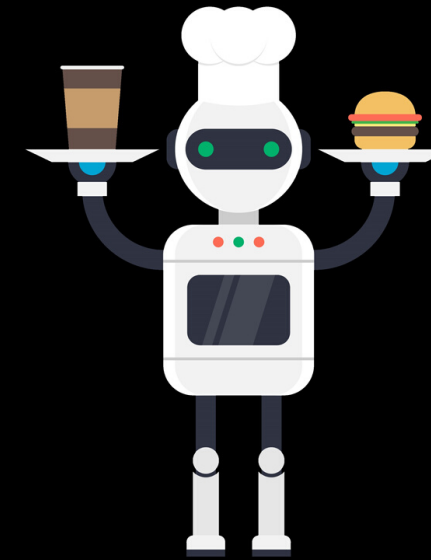
Continued

```
Topics for POSITIVE reviews
-----------------------------------------
Topic 0:    3.467*chicken, 1.218*biriyani, 1.136*order, 0.871*rice, 0.720*tasty
Topic 1:    3.364*biryani, 0.393*mutton, 0.317*chicken, 0.275*best, 0.264*dum
Topic 2:    3.883*nice, 0.155*ambience, 0.143*tasty, 0.114*friends, 0.107*staff
Topic 3:    1.842*service, 1.540*great, 1.281*ambience, 0.932*staff, 0.884*visit
Topic 4:    2.723*taste, 1.912*awesome, 0.737*quality, 0.668*quantity, 0.521*delivery
Topic 5:    2.275*veg, 1.258*non, 0.534*starters, 0.365*buffet, 0.324*main
```
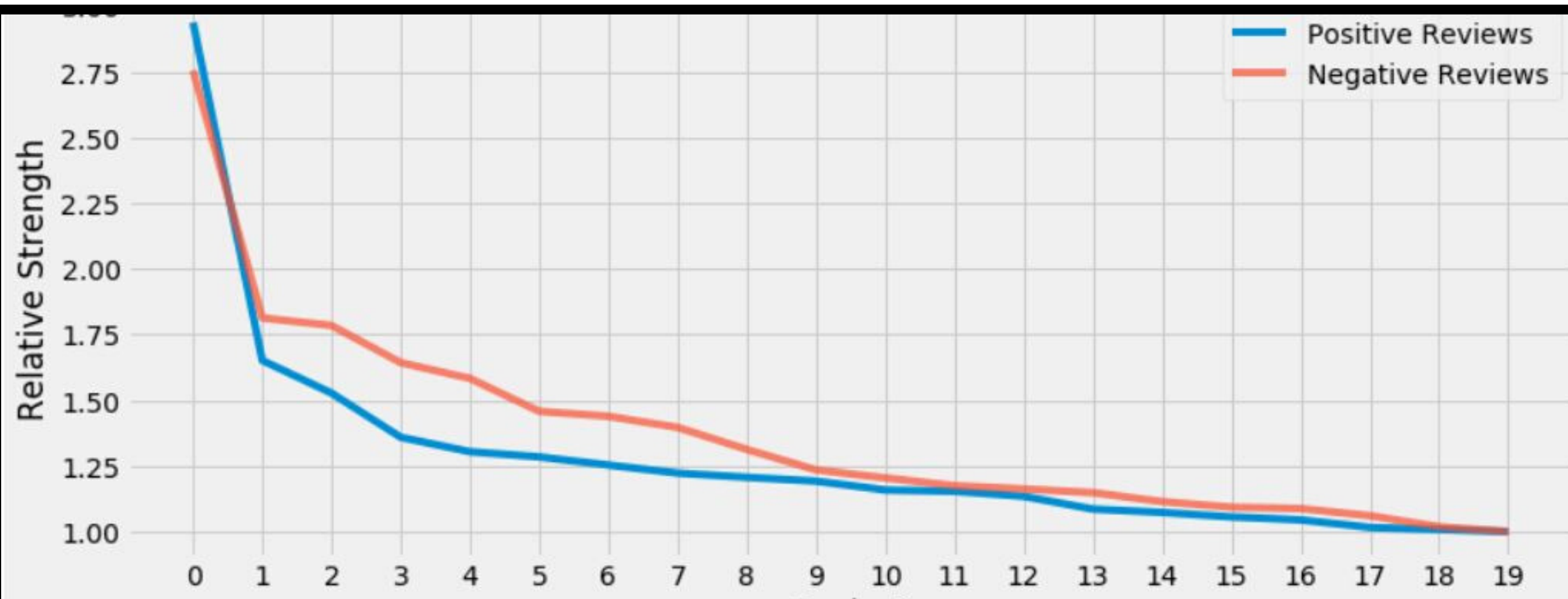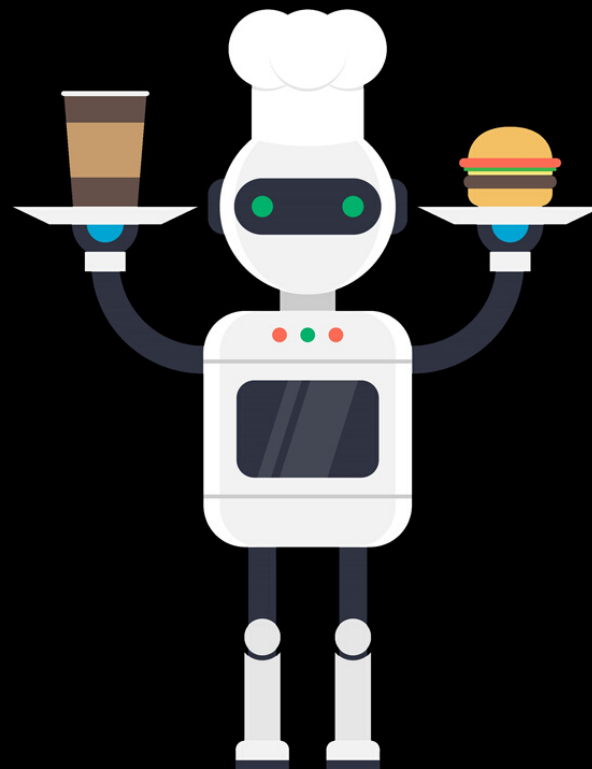
Conclusions

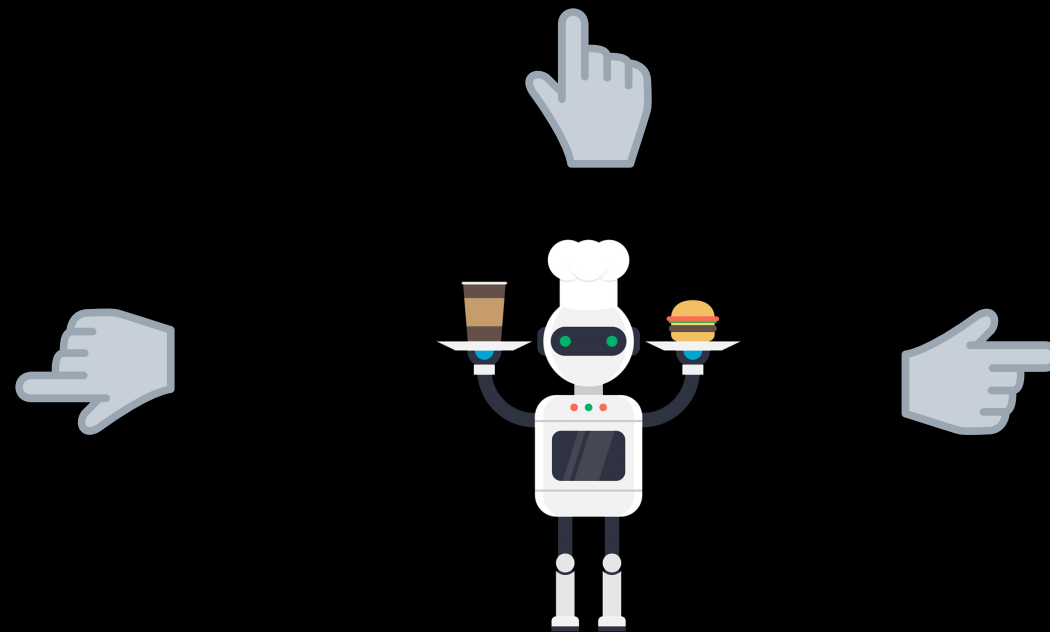- We have proposed a technique to analyze the customer reviews of casual dining restaurants

- The UI would enable customers to search similar restaurants based on different attributes such as Similar Restaurants, Ambience and Service

Future Scope

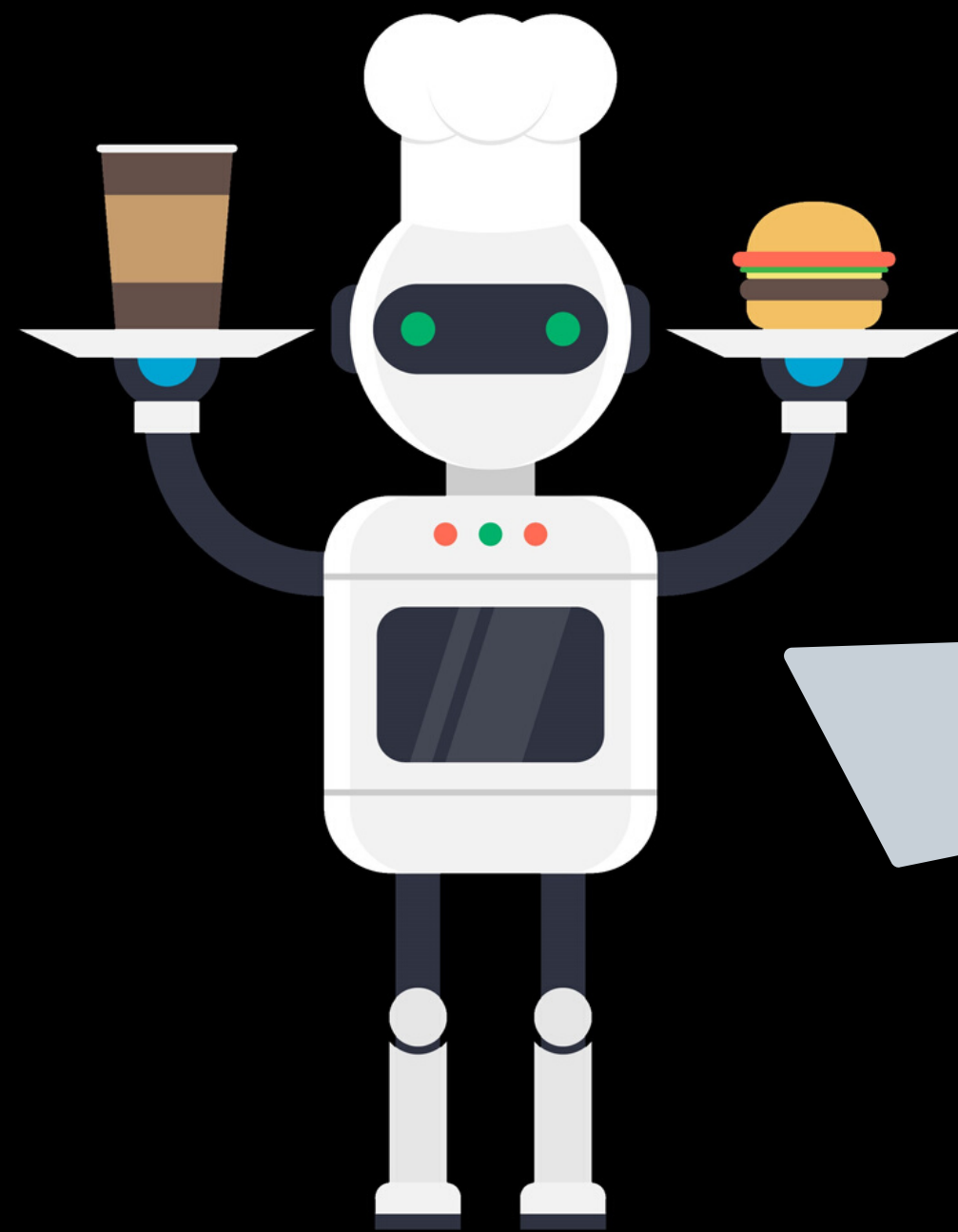- Zomato dataset ➡️ Local systems storage ➡️ Preprocessing ➡️

Analysis == Make the above process dynamic

- Customize the product to be business/owner specific

- Increase the search criteria beyond what we have

Thank You

References

https://github.com/wolfgarbe/SymSpell

https://towardsdatascience.com/stemming-lemmatization-what-ba782b7c0bd8

https://medium.com/@praneethbedapudi/deepcorrection-1-sentence-segmentation-of-unpunctuated-text-a1dbc0db4e98

https://nlpforhackers.io/complete-guide-to-spacy/