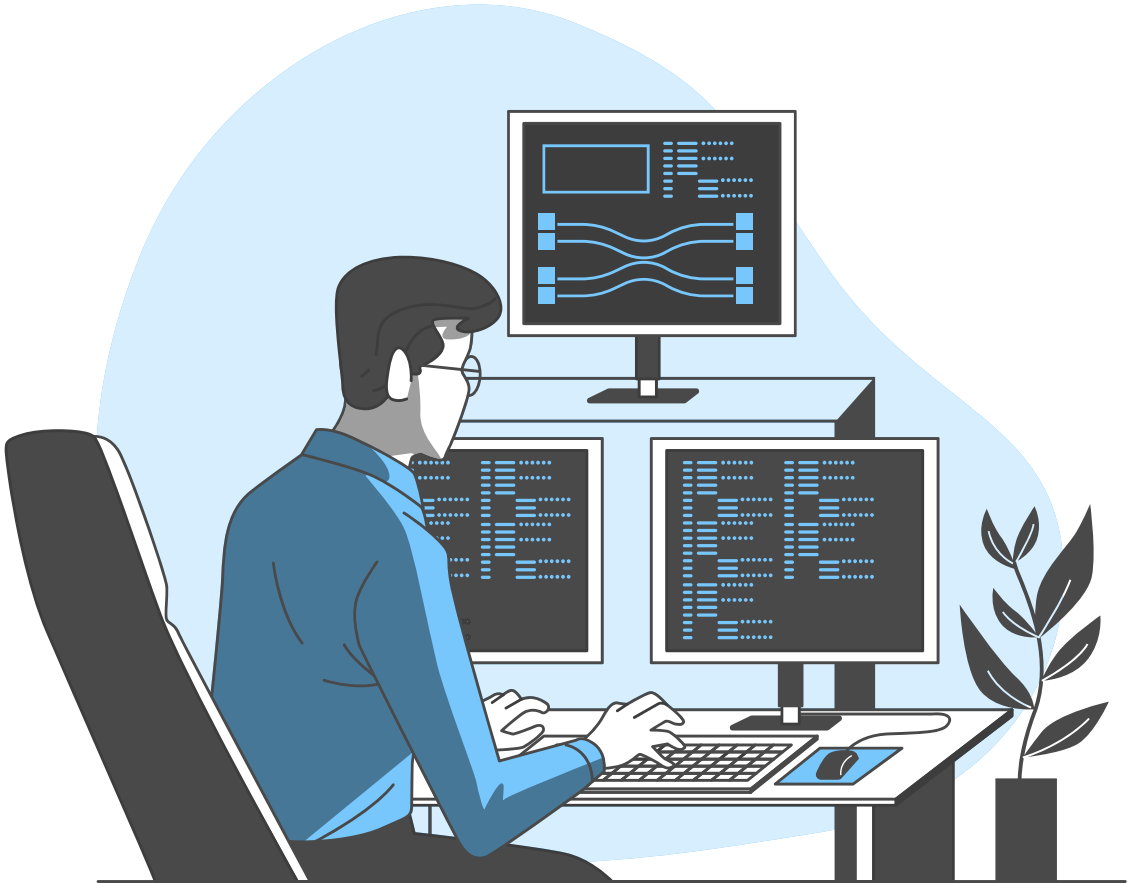


Restaurant Reviews NLP

Hernan Trujillo

01

Introduction

Intro Tips

01 Reviews are about 1 restaurant chain from different locations

02 “Liked” column is an insight provided by the customer

03 Initial findings on the “Liked” columns are exactly the same 50%/50%

02

Data

Data Gathering

01

Source

Kaggle.com
https://www.kaggle.com/vigneshwarsofficial/reviews

02

Type of Data

Restaurant Reviews (text) &
Likes (numbers)

03

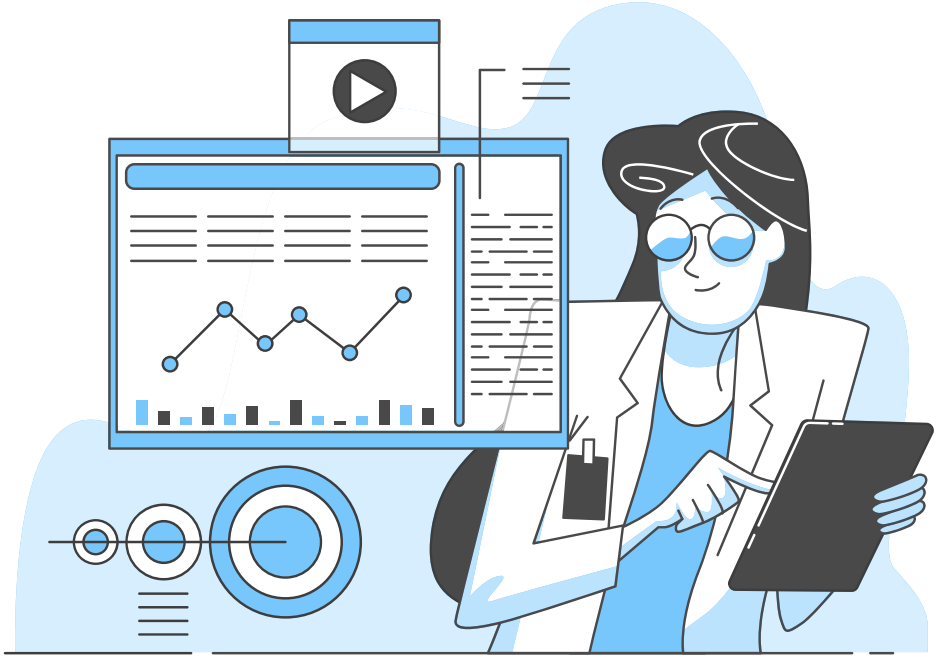
Amount of Data

1000 Reviews

04

Our Consultants

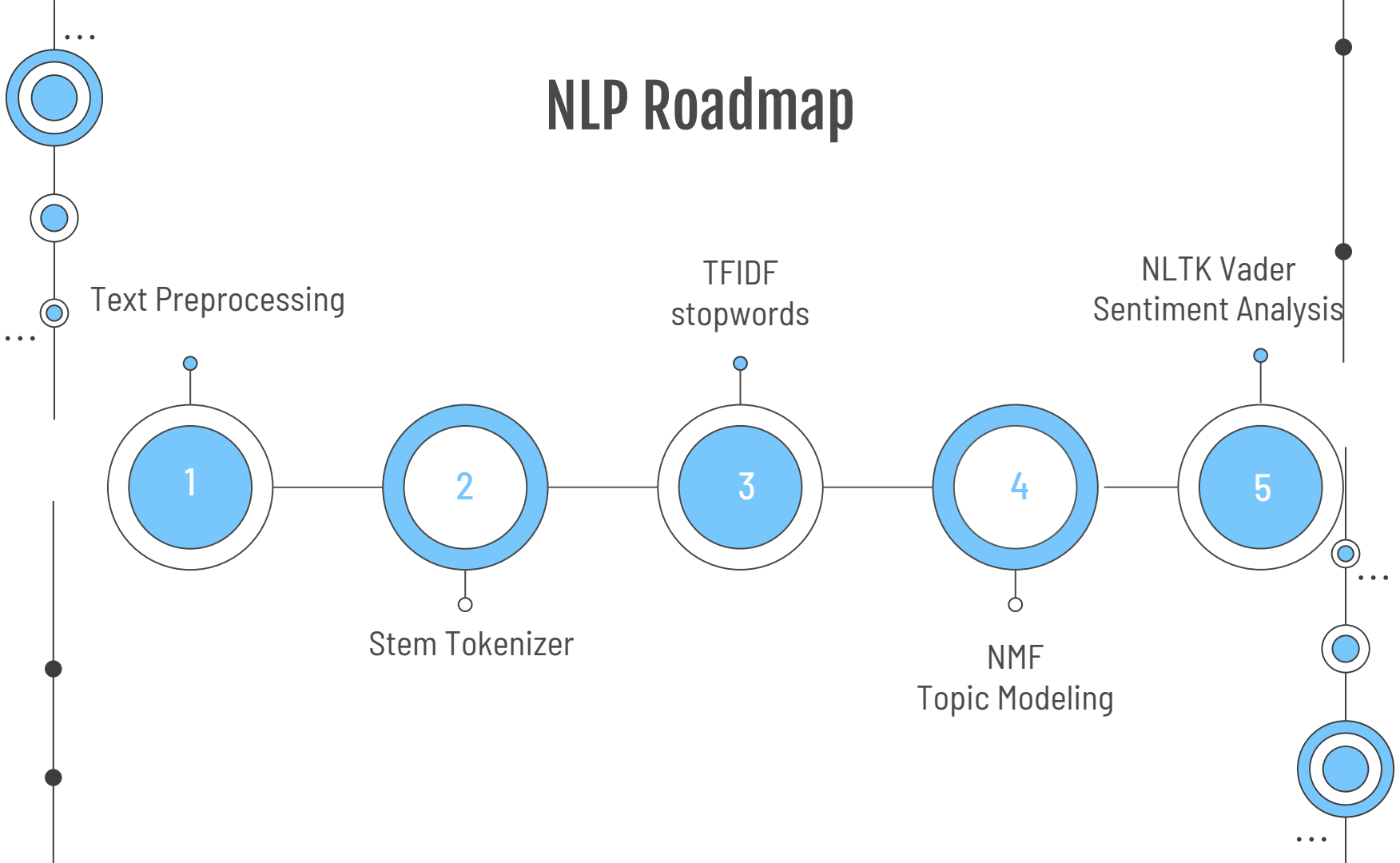
Review Column: customer reviews
Like column: Positive and Negative



03

Approach

NLP Roadmap


Text Preprocessing
1
Stem Tokenizer
2
TFIDF
stopwords
3
NMF
Topic Modeling
4
NLTK Vader
Sentiment Analysis
5

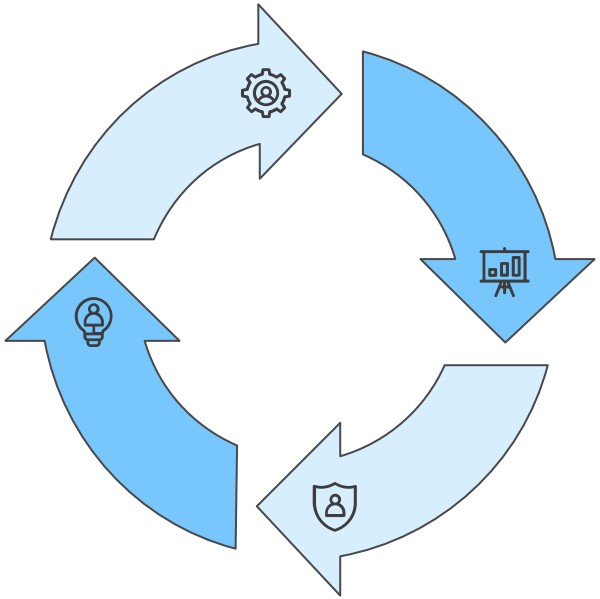
Modeling Workflow

10 Topics

NMF
(n_components=10)

30 Topics

NMF
(n_components=15)



7 Topics

NMF
(n_components=7)

5 Topics

NMF
(n_components=5)

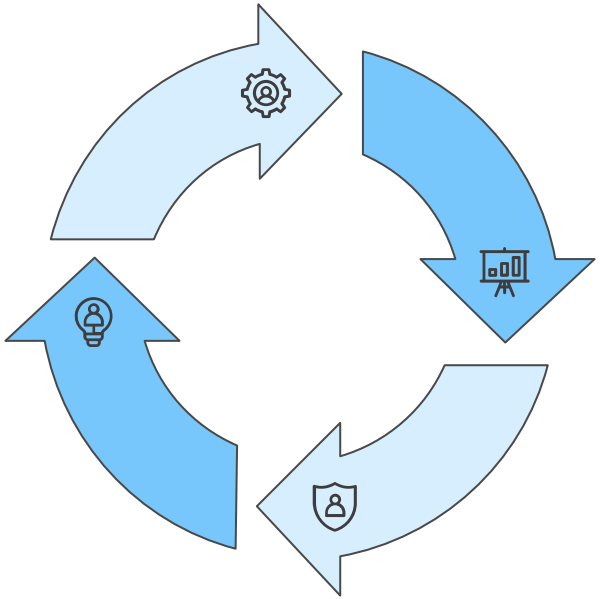
Modeling Workflow N=4

2 Topics

NMF
(n_components=2)

4 Topics (Best)

NMF
(n_components=4)



Topic Interpretation

Interpretation for meta data

Sentiment Analysis

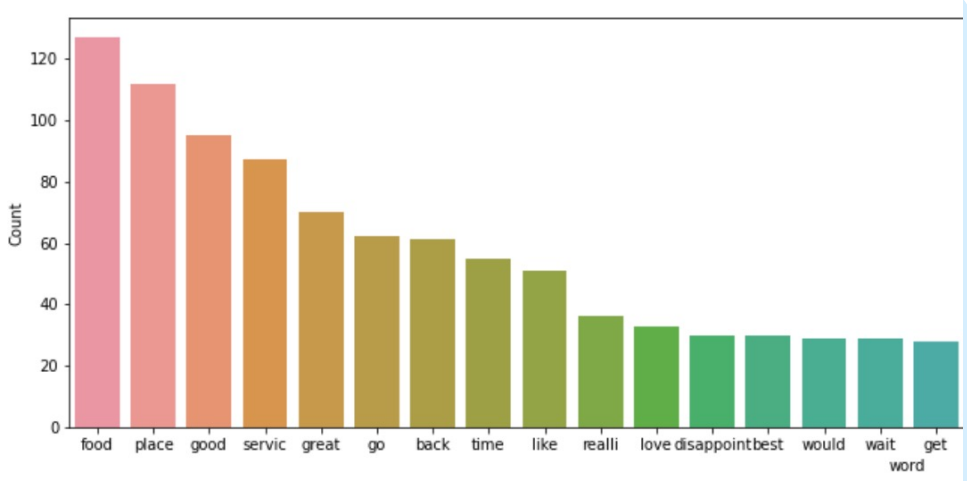
NLTK Vader Compound Score

04

Results

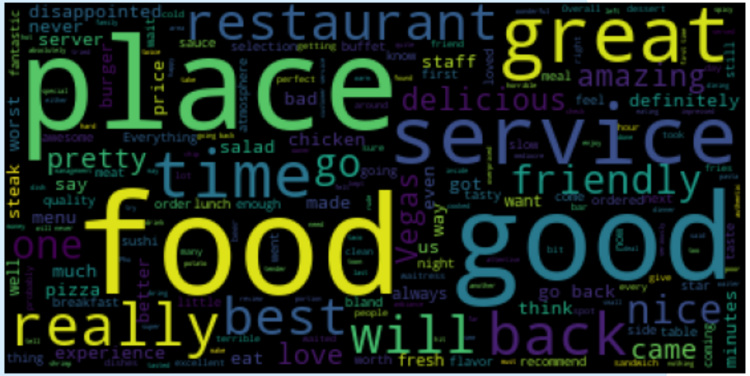
Top Words

	words	tf
79	food	127.0
1	place	112.0
3	good	95.0
41	servic	87.0
18	great	70.0
42	go	62.0
44	back	61.0
32	time	55.0
29	like	51.0
282	realli	36.0



Wordcloud


enough
took
draw
wound
check
wast
life
pour
time
salt
bring


disappointed
restaurant
great
place
service
delicious
amazing
definitely
friendly
time
pretty
food
good
really
best
will
back
nice
came
love
Vegas
server
never
burger
price
staff
bad
chicken
salad
go
menu
steak
quality
say
one
much
pizza
breakfast
experience
better
sushi
think
go back
fresh
recommend
minutes
worst
awesome
Everything
order
lunch
enough
made
got
tasty
want
ordered
next
us
night
always
bland
people
eat
flavor
excellent
atmosphere
sauce
selection
buffet
first
loved
meal
feel
slow
fries
taste

Topic Results

01 Food Quality

02 Place Approval

03 Fidelity

04 Service

Sentiment Analysis

69%

Positive

Positive reviews after processing the data

31%

Negative

Negative reviews after processing the data

05

Future Work

CorEx

Correlation **Ex**planation, topic modeling
Discriminative Model

Why:? Topic modeling via total correlation

Thanks!

Do you have any questions?

hernantru943@gmail.com

GitHub
https://github.com/hernantru943/Restaurant-Customer-Reviews