SDAIA
الهيئة السعودية للبيانات والذكاء الاصطناعي
Saudi Data & AI Authority

McDonald’s Review

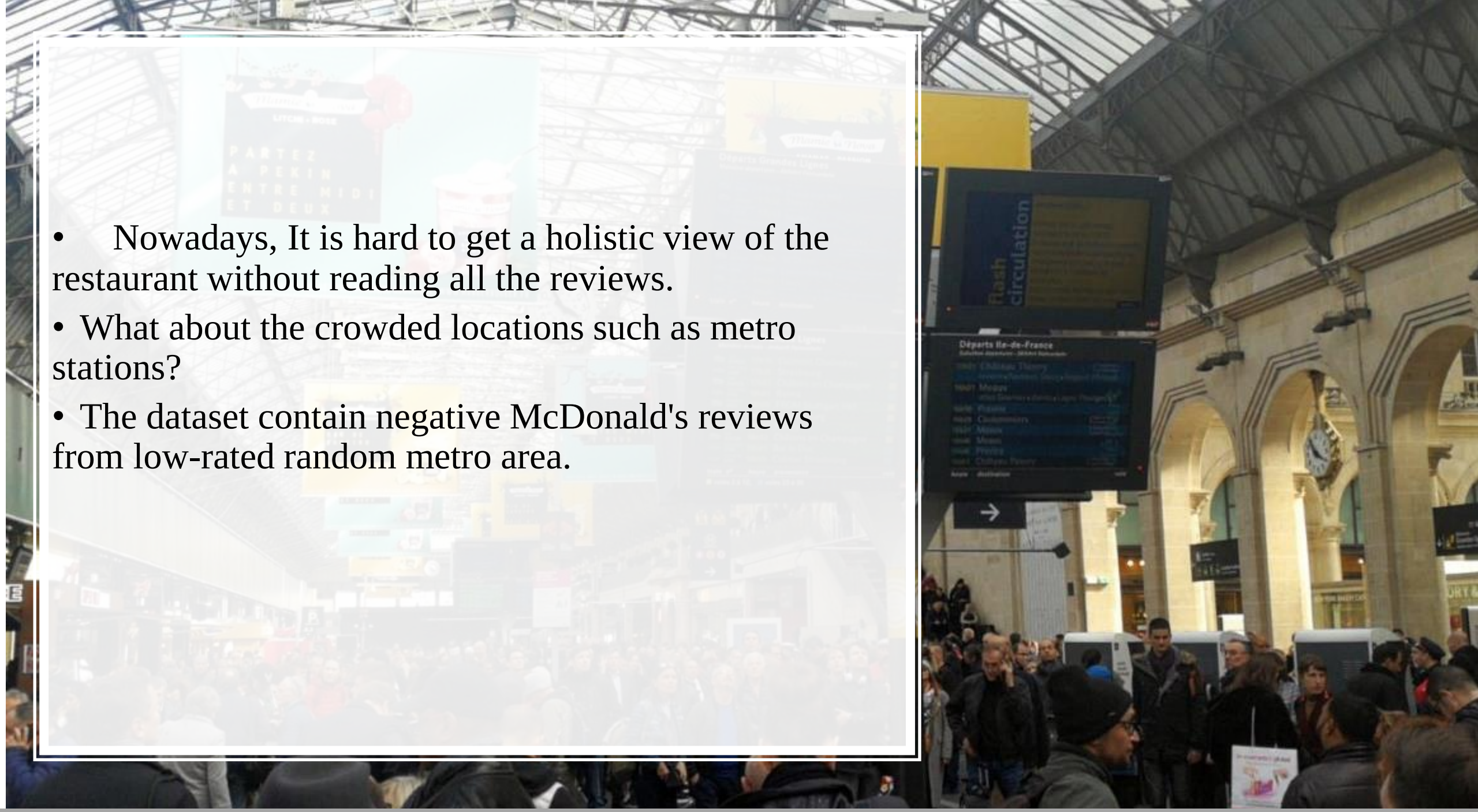
NLP project

Njood Alqhtani & Reem Alsuqaih

TABLE OF CONTENTS

- Introduction
- Problem Statement
- Methodology
- Data preprocessing
- Experiments
- Conclusion

Introduction



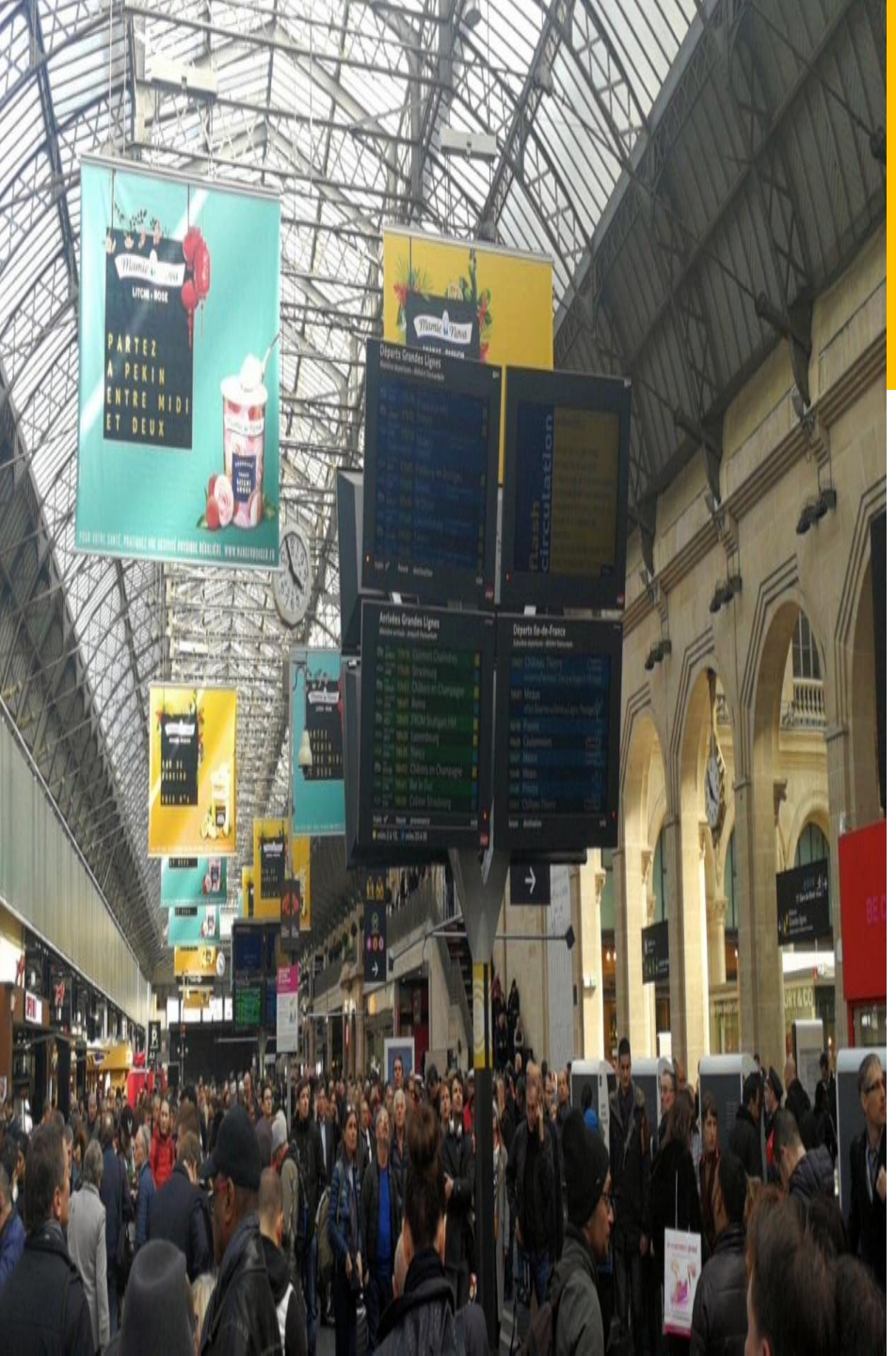
- Nowadays, It is hard to get a holistic view of the restaurant without reading all the reviews.
- What about the crowded locations such as metro stations?
- The dataset contain negative McDonald's reviews from low-rated random metro area.

Problem Statement

A sentiment analysis of negative McDonald’s reviews. Contributors were given reviews culled from low-rated McDonald’s from random metro areas and asked to classify why the locations received low reviews. Options given were: Rude Service ,Slow Service ,Problem with Order, Bad Food Bad Neighborhood , Dirty Location ,Cost ,Missing Item .

Methodology

- Data Understanding
- Data preprocessing
- Topic Modeling
- Classification
- Clustering
- Conclusion



Data Understanding

unit_id	id of record
golden	value FALSE
unit_state	value finalized
trusted_judgments	value 3
last_judgment_at	Time. Example 2/21/15 0:36
policies_violated	The type of policies, violated. Example: *RudeService
policies_violated.confidence	The confidence of policies, violated. Example: 1.00.66670.6667
city	City name
policies_violated_gold	value NA
review	review detail

Data preprocessing

- To get started, we need to import some useful libraries that will help us import the dataset into our python environment, manipulate and analyze the same and later help us to visualize it.
- Tokenization, normalization, stop word filtering, and lemmatization the dataset.
- Remove unnecessary columns.

Experiments

Topic Modeling

Topic Modeling (NMF)

Topic 0
order, time, wrong, minutes, right, breakfast, wait, window, ordered, waiting

Topic 1
mcdonald, review, people, ve, know, way, line, ξthe, say, good

Topic 2
food, fast, time, waiting, eat, restaurant, customers, line, fresh, quick

Topic 3
mcdonalds, ve, breakfast, good, want, really, people, ξthe, employee, pretty

Topic 4
drive, window, car, way, inside, went, cars, location, sure, said

Topic 5
fries, chicken, ordered, got, said, ξthe, went, cold, asked, meal

Topic 6
coffee, burned, iced, morning, breakfast, cream, hot, time, starbucks, line

Topic 7
place, time, people, location, staff, ve, really, kids, inside, work

Topic 8
like, mcd, want, people, really, sweet, eating, old, shake, looks

Topic 9
service, customer, location, ve, manager, worst, bad, slow, times, rude

Topic Modeling (LSA)

Topic 1
mcdonald, like, review, mcwrap, fries, people, old, north, ξand, way

Topic 2
food, mcdonalds, place, fast, service, eat, like, good, people, want

Topic 3
food, order, fast, mcdonald, service, window, waiting, customers, time, longer

Topic 4
order, time, fries, manager, coffee, minutes, waiting, counter, ordered, chicken

Topic 5
fries, food, chicken, ordered, window, drive, asked, got, said, went

Topic 6
coffee, food, window, line, burned, cream, got, waiting, morning, wanted

Topic 7
mcdonalds, service, mcdonald, customer, fries, breakfast, employee, manager, ξthe, worst

Topic 8
time, mcdonalds, drive, line, minutes, fries, wait, 10, north, ve

Topic 9
people, line, window, like, waiting, minutes, mcdonalds, place, lot, north

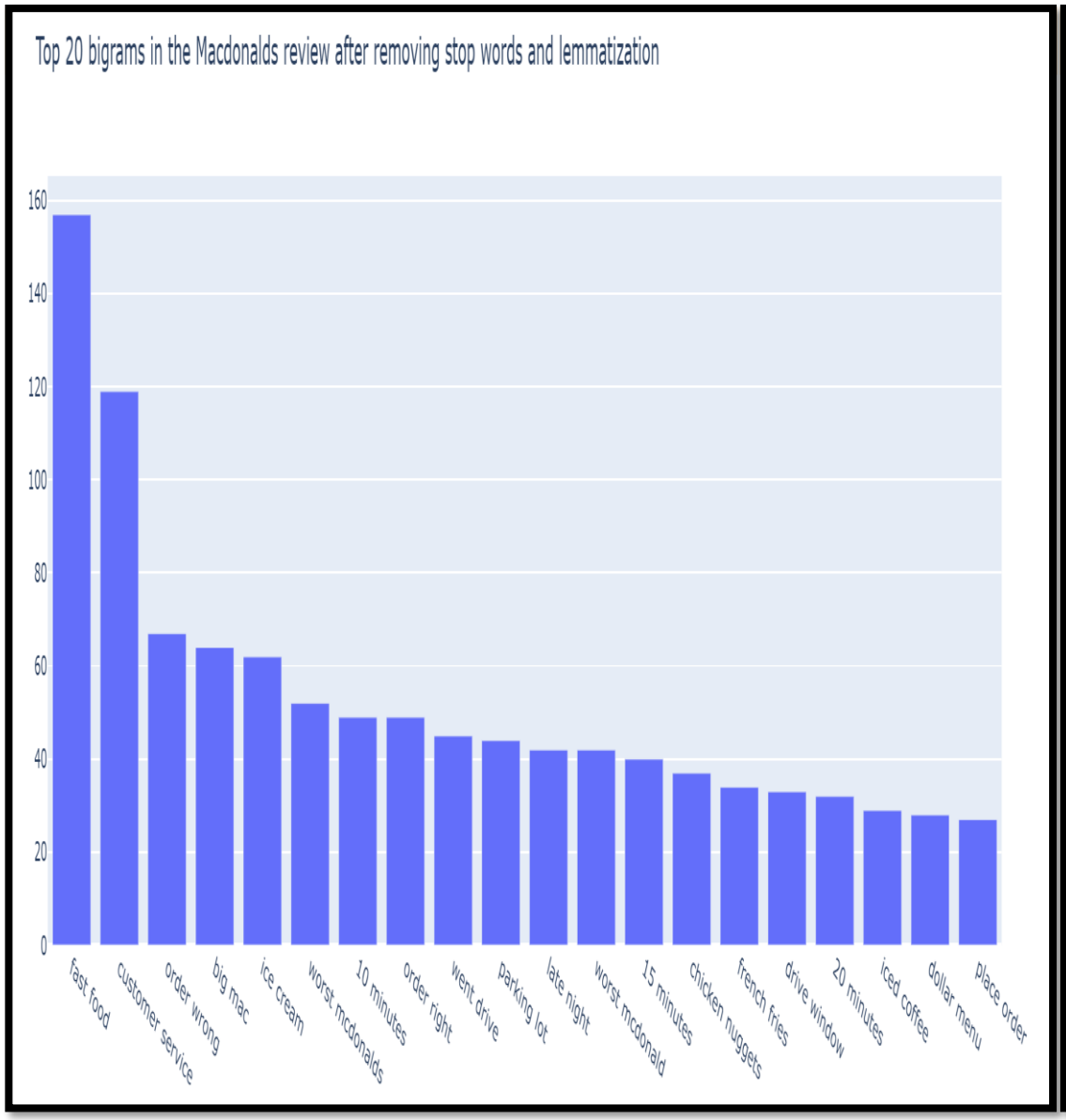
Topic Modeling (LDA)

```
Topic  1
food, mcdonald, order, drive, like, fries, fast, time, mcdonalds, place

Topic  2
mcdonald, mcdonalds, food, like, time, location, place, good, fries, ve

Topic  3
service, order, drive, food, asked, manager, mcdonald, place, say, like

Topic  4
mcdonald, time, food, mcdonalds, order, breakfast, worst, orders, like, drive

Topic  5
order, drive, service, food, mcdonald, mcdonalds, time, window, place, minutes

Topic  6
mcdonald, food, like, service, time, people, breakfast, biscuit, place, egg
```

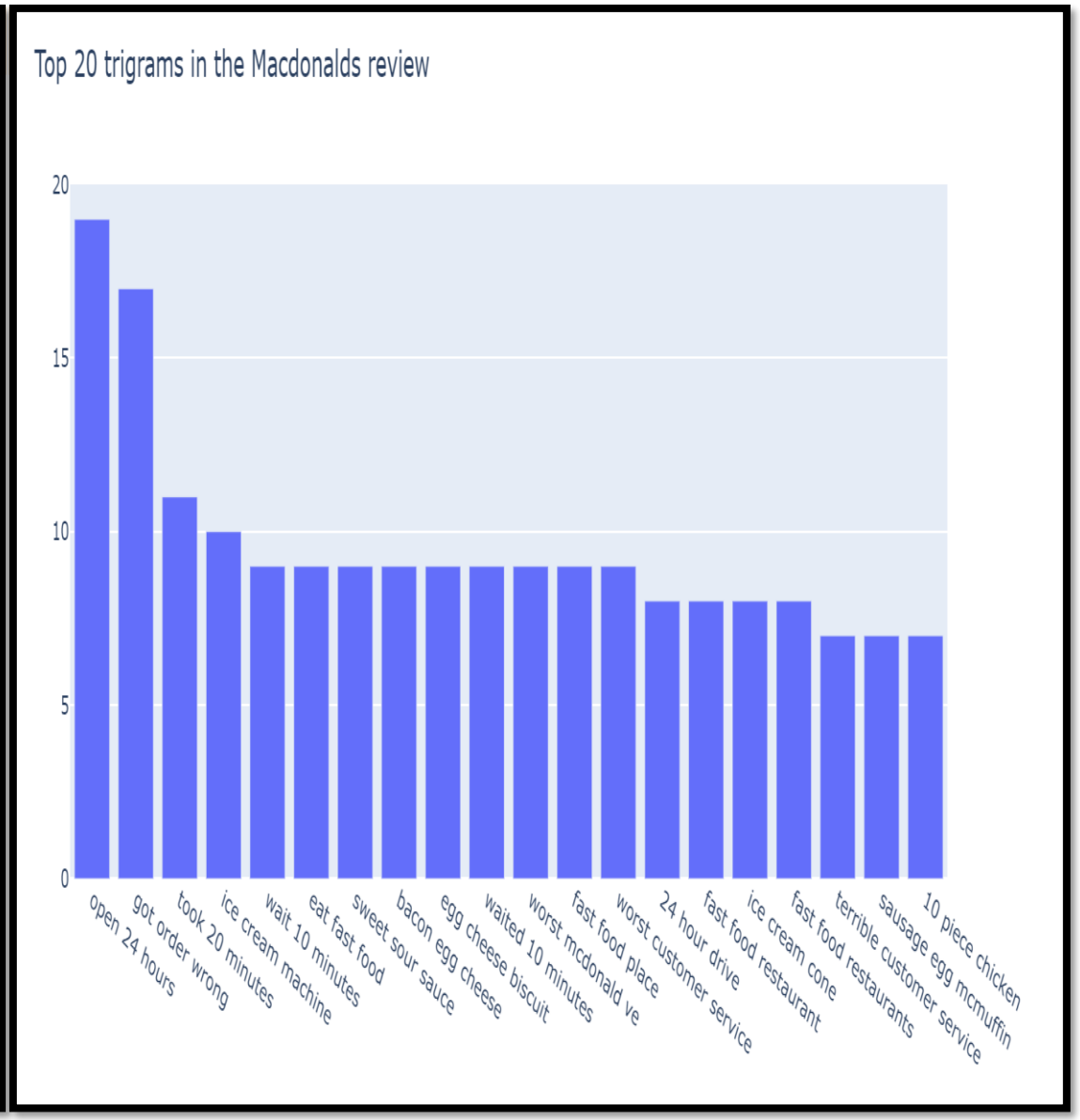
The best Model of Topic Modeling is (LDA) so we take the word cloud of the best model. It shows the most common word.

one
ξ
This
lover
Dirt
race
know
sauc
customer
points
review
long
lost
clean_text_stopward
time
Length
McDonald
First
really
stood
wanted
repeatedly
see
service
mcds
Terrible
seems
part
giving
actually
fa
gave
huge
Two
grab
Well
work
order
dtype
right
came
breakfast
asked
30pm
morning
star
enjoyed
gate
Name
craving
someone
better
Worst
food
object
Thuggery
closes
everywher

Top 20 Bigrams

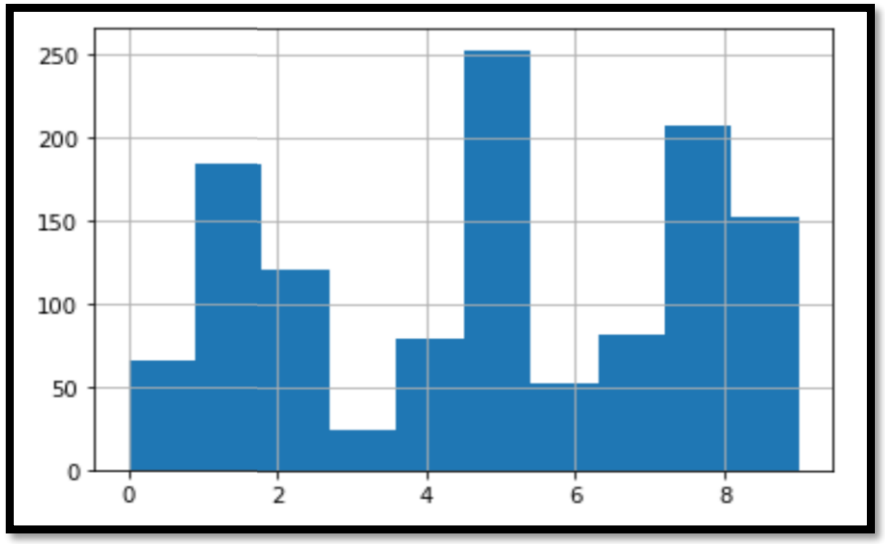


Top 20 Trigrams

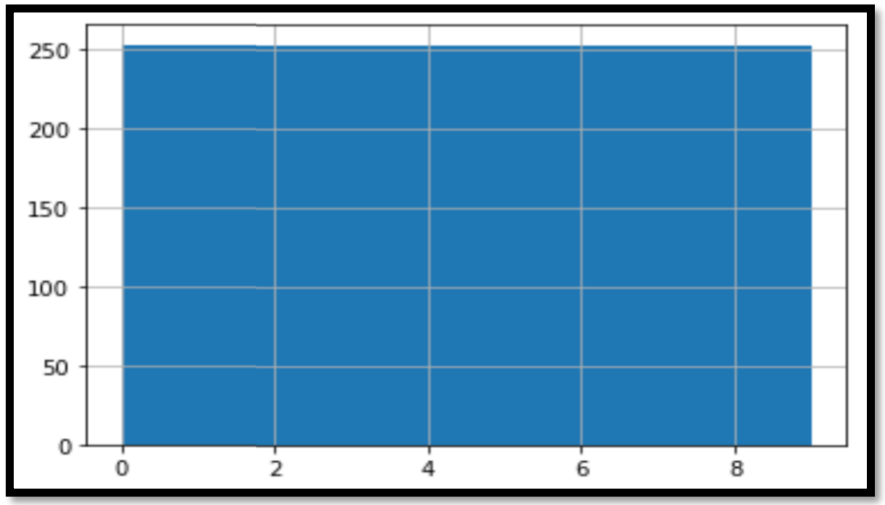


Classification

SMOTE for Data Balance



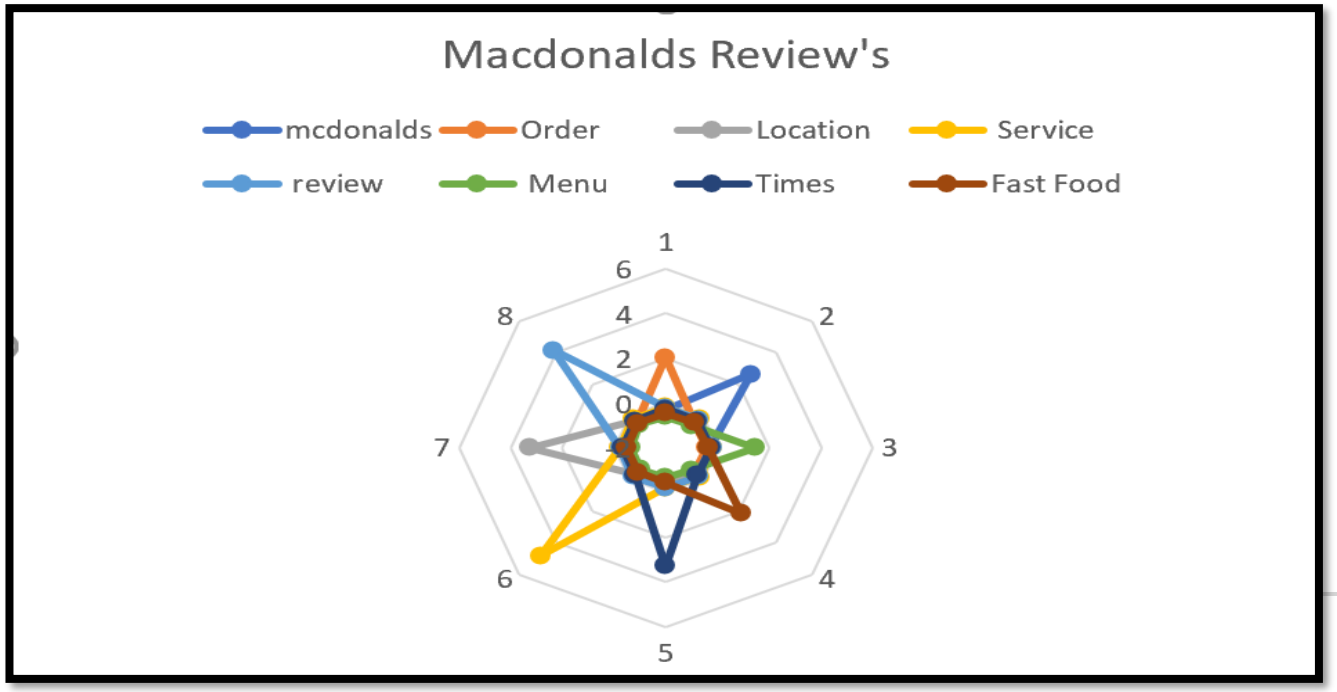
Imbalance

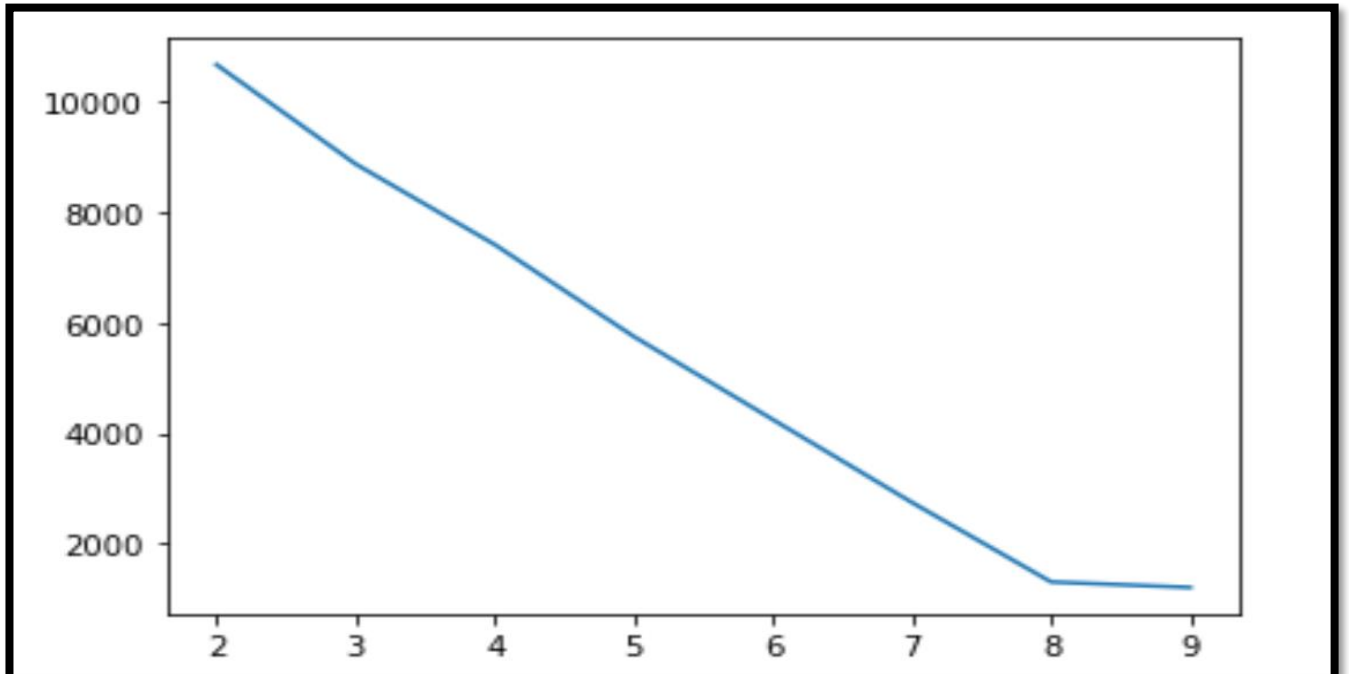


After SMOTE

	Logistic Regression CV		Logistic Regression TFIDF		Naive Bayes Multinomial CV		Naive Bayes Multinomial TFIDF		Naive Bayes Bernoulli CV		Naive Bayes Bernoulli TFIDF	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
Accuracy	99%	51%	80%	47%	83%	40%	47%	29%	48%	25%	48%	25%
Precision	99%	52%	83%	43%	87%	42%	54%	21%	79%	23%	79%	23%
Recall	99%	51%	80%	47%	83%	40%	47%	29%	48%	25%	48%	25%
F1 Score	99%	48%	78%	40%	82%	33%	38%	19%	45%	15%	45%	15%

Clustring


Macdonalds Review's
mcdonalds
Order
Location
Service
review
Menu
Times
Fast Food
1
2
3
4
5
6
7
8
0
2
4
6


10000
8000
6000
4000
2000
2
3
4
5
6
7
8
9

Conclusion

We have noticed from the above the low service in crowded places and the frequent inconvenience of customers so it is necessary to take into account the negative comments and develop them and consider them, from this point of view must take into account the project Of Riyadh Metro, which will open soon where it is considered one of the most important projects for public transport in Riyadh city, it is necessary to take into account the locations of restaurants and the number of employees in them and work to improve the service in them and take into account the opinions of customers.

Thanks, you all