Fake or Real News

By :
Shahad Alomayrah
Rana Alzahrani

Outlines

1.Introduction

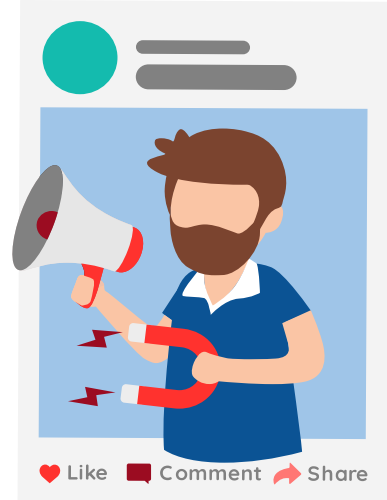2.Data set

3.Data Preprocessing

4.Topic Modeling

6. Classification Models

7.Tools

Overview

In this project, we are working on a dataset that consists of text about the News . Our observation is a Fake or Real News .

Goal

Building NLP model which is unsupervised learning that focuses on finding meaningful topics on News.
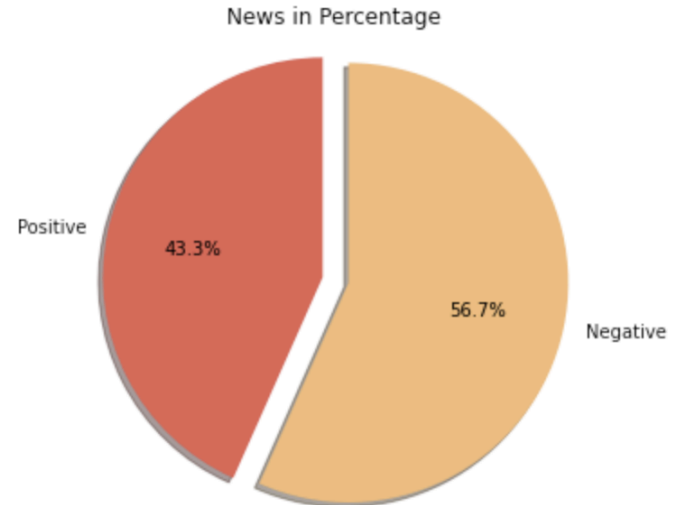
Dataset

20,761 Rows

5 Columns

We Collected the data from Kaggle website

10,361 Positive

7,924 Negative

News in Percentage

Positive 43.3%

Negative 56.7%

Data Preprocessing

Data Cleaning
- Remove spaces and punctuations
- Remove number
- Remove Stop words

Lemmatization

Delete Meaningless Words

Vectorization
- Count Vectorizer

Topic Modeling Algorithms

1
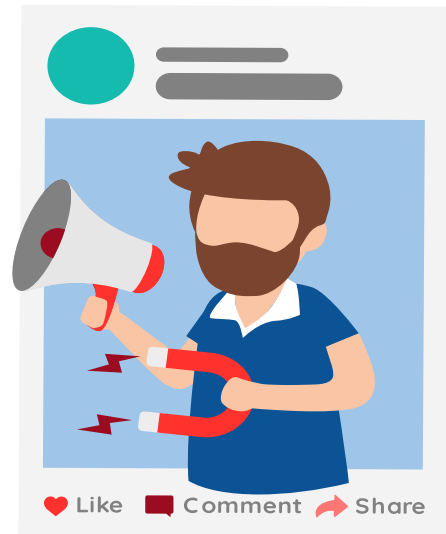
2

3

4

LSA

NMF

CorEX

LDA

The best algorithms is **NMP**
with 10 topic

Like   Comment   Share

Positive Topic

pay, job, tax, money, hard, bad, economy **Topic 1 'Taxs'**

money, bank, debt, reserve, loan, trillion   **Topic 2 'Financial'**

military, force, nuclear, attack, Iran, weapons **Topic 3 'Wars'**
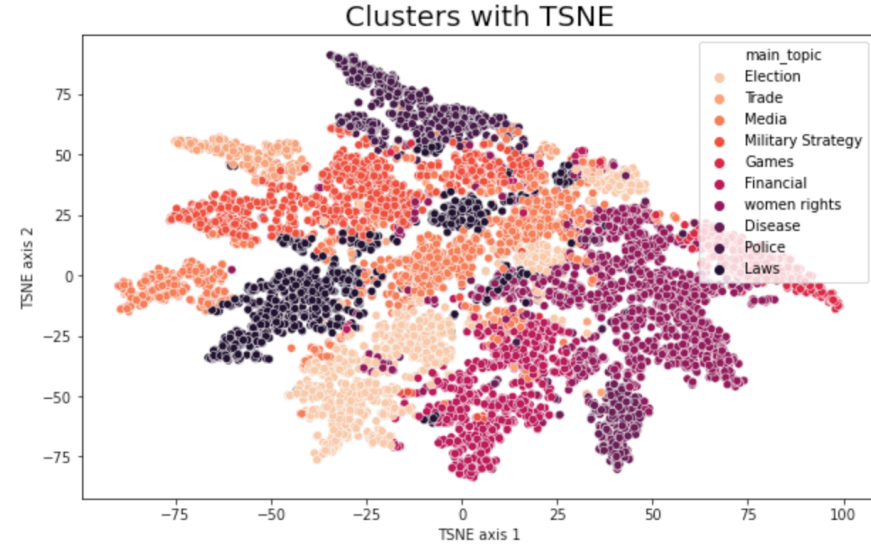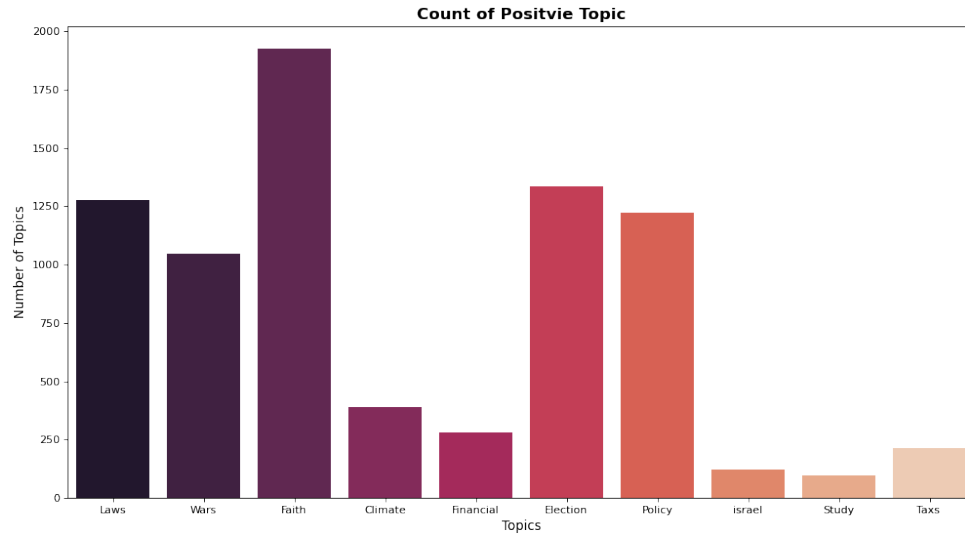
Arabs , jewish , expel, attack, Palestinians , Zionists **Topic 4'Israel'**

fbi, investigation, case, evidence, source **Topic 5'Laws'**

vote, win, media, candidate, presidential, democratic **Topic 6' Election'**

Positive Topic

Negative Topic

women, life, feel, right, mother, include **Topic1'Women Rights'**

vote, political, win, presidential, conservative, candidates **Topic2'Election'**
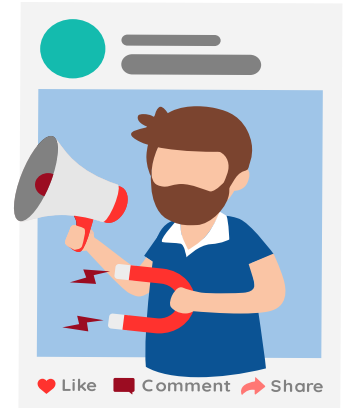
kill, security, fight, bomb, strike, troop, minister **Topic3'Militray Strategy'**

officer, gun, arrest, drug, suspect, gang, protest **Topic 4'Police'**

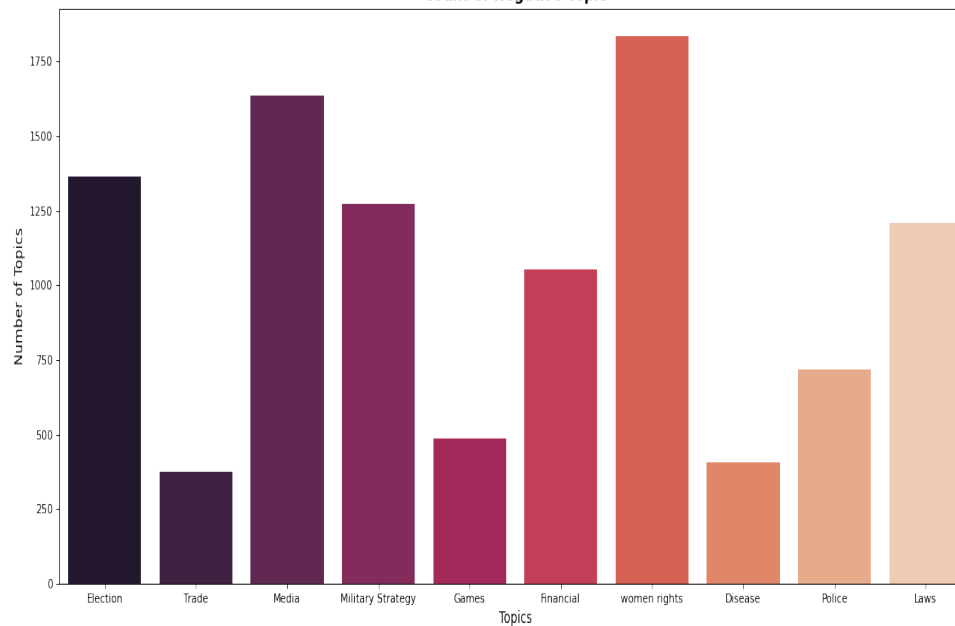tax, pay , money, bank, billion, increase, deal, firm **Topic 5 'Financial'**

trade, missile, deal, economic, global, weapons **Topic 6'Trade'**

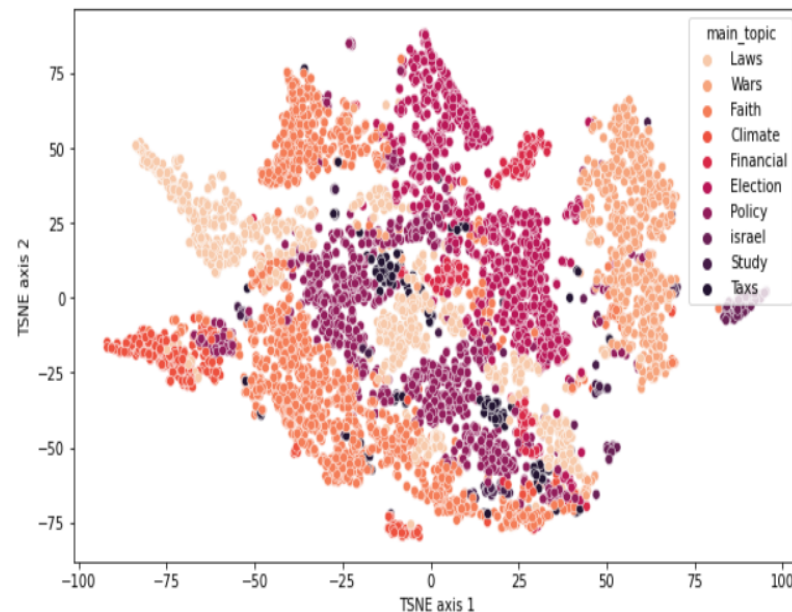season, players, win, sport, fan, league, apocalypse **Topic 7'Games'**
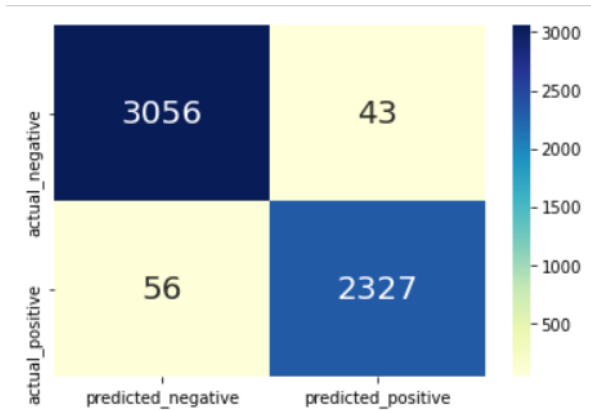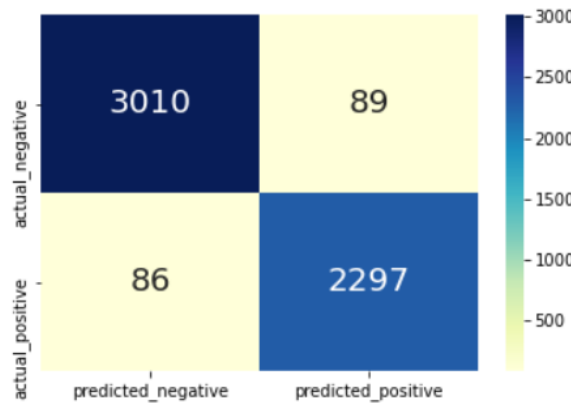
Negative Topic

Confusion Matrix

Classification

Model	BernoulliNB	LogisticRegression	MultinomialNB
Accuracy	0.90	0.98	0.96
Precision	0.99	0.98	0.96
Recall	0.79	0.97	0.96

↑

The Best Model

NEWS

search

Tools

- **Sklearn**
- **Numpy**
- **Pandas**
- **Seaborn**
- **WordCloud**
- **Anaconda**
- **Github**
- **Jupyter Notebook**

THANKS!

Do you have any questions?