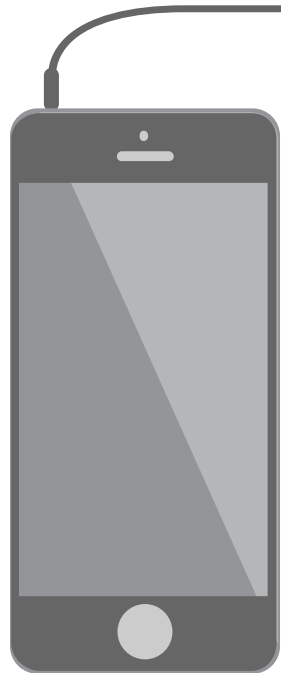


# Reviews For iPhone SE NLP



Teams:  
Lama Alziyad & Kholod alshehri

# Outline



01

**Introduction**

02

**EDA**

03

**Topic Modeling**

04

**Models**

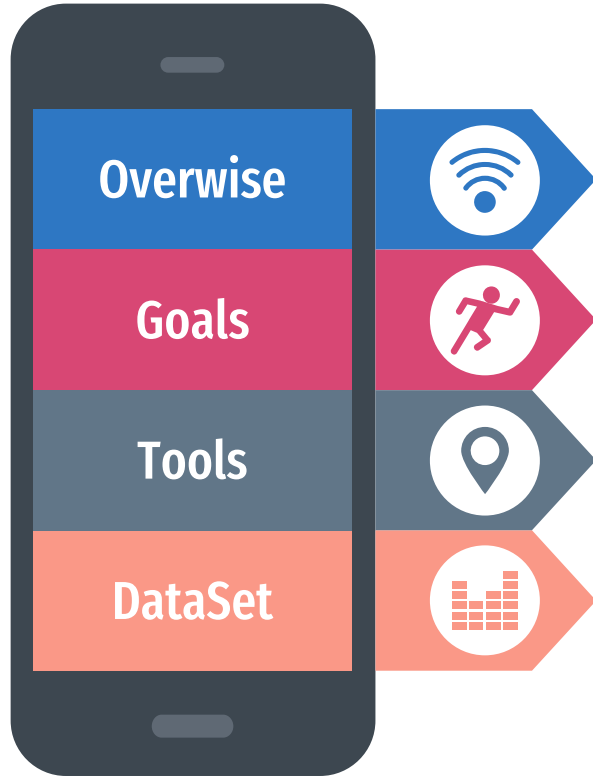
05

**Clustering**

06

**Conclusion**

# Introduction



The are sentiment analysis about I phone SE

Our goal is to build NLP to reviews about  
i phone SE is positive or negative

We use tools is : pandas, NumPy, SpaCy, NLTK,  
matplotlib, Jupyter, HTML, Seaborn, Python

The Data size is 9713 x 3 and 3 Columns it is ( Rating, Comments, Reviews )

After: 9049 x 2

# EDA and Pre-processing

## Data Cleaning

- Drop comments
- Drop duplicated

## Lemmazation



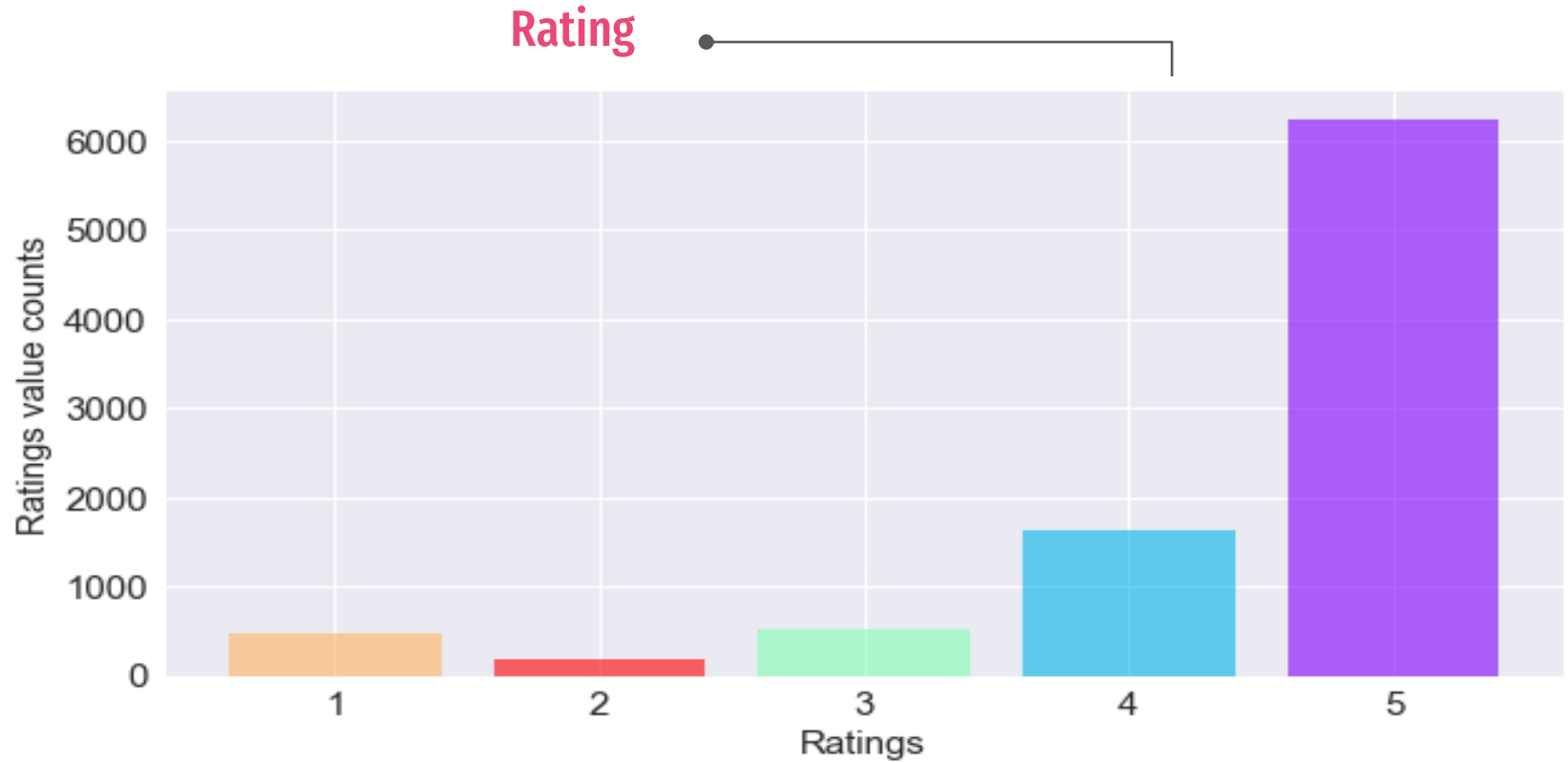
## Vectorization

CountVectorizer

## Tokenization

NLP Pipe  
ADV  
None  
Verb

# Data Analysis



# Fauters Engineering

## Positive

Add new columns it is  
sentiments more than and  
equal 3

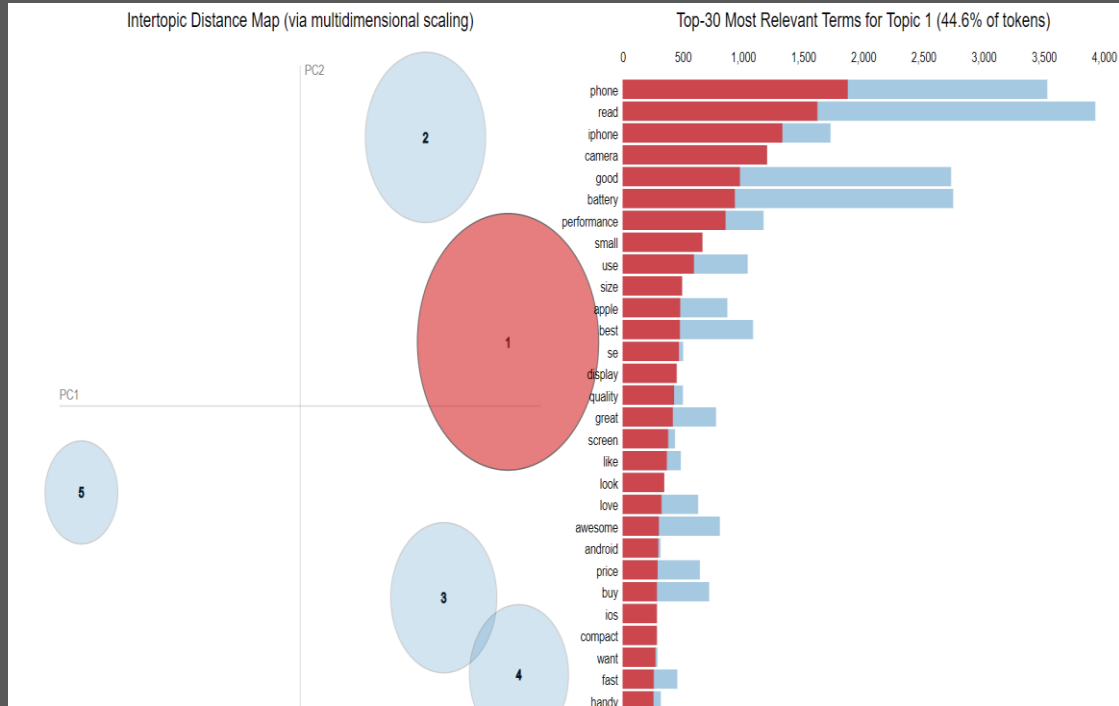


## Negative

Add new columns it is  
sentiments less than 3

# Topic LDA

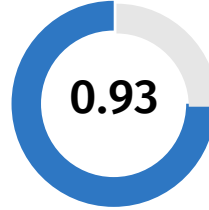
## pyLADvis



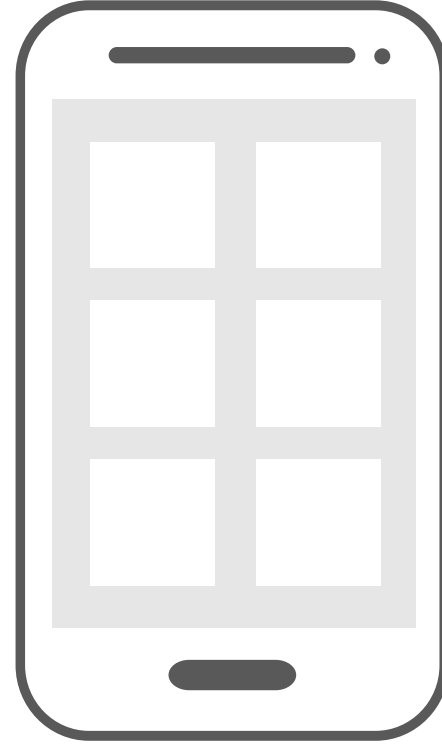
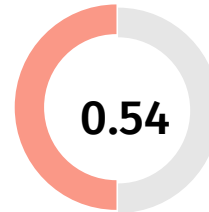
# Classification Modeling

Accuracy Score

Random Forest



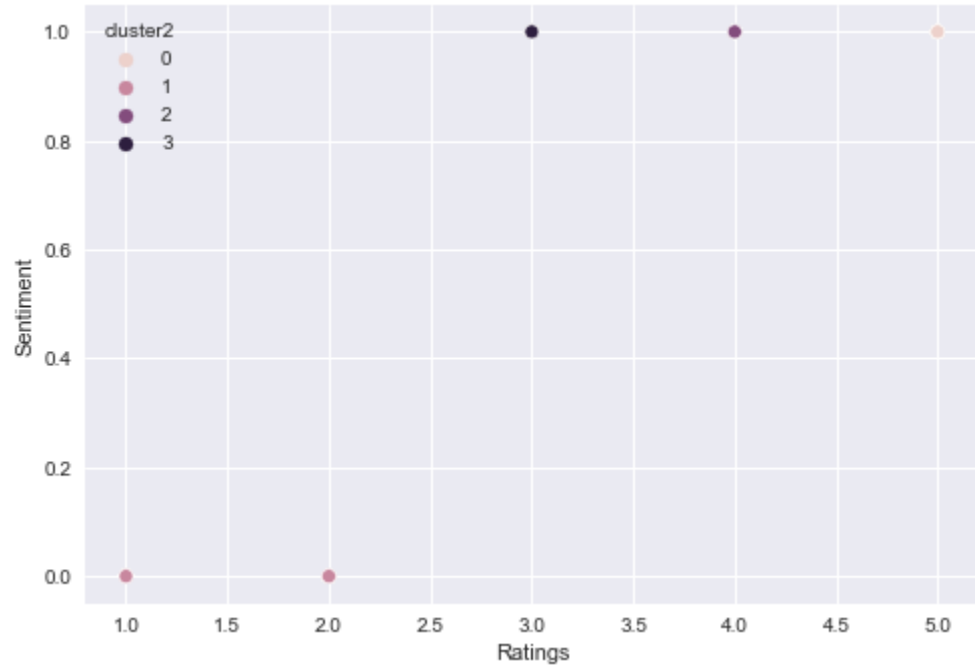
Naive Bayes





# Clustering

**Kmeans:**



# Tools



NLTK



Spacy

Word Colud



pyLADvis



# Conclusion



Random Forest

Accuracy: **93%**



**Any Questions?**

Please Say No 😊