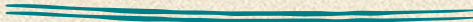


# Define Quotes Tags Using NLP

Shaikha Bin Ateeq  
Raghad Althunayan



# Introduction:

Quotes have the power to motivate, inspire and encourage At such time, read a few quotes, and you will see how your mood and state of mind begin improve

## Goal:

predict the quote tags whether 'Sadnesses','Knowledge','Universe','Feelings'







Tell me and I forget. Teach me and  
I remember. Involve me and I learn



**- Benjamin Franklin**



# Dataset



## Data Source

- From Kaggle
- Web scraping (Goodreads)



## Rows & Columns

- Rows & Columns (499709,3)
- Rows & Columns (83618,3)



## Features

- Quotes
- Author
- Category





# Tools



## Data Cleaning & Manipulation



## Model Building

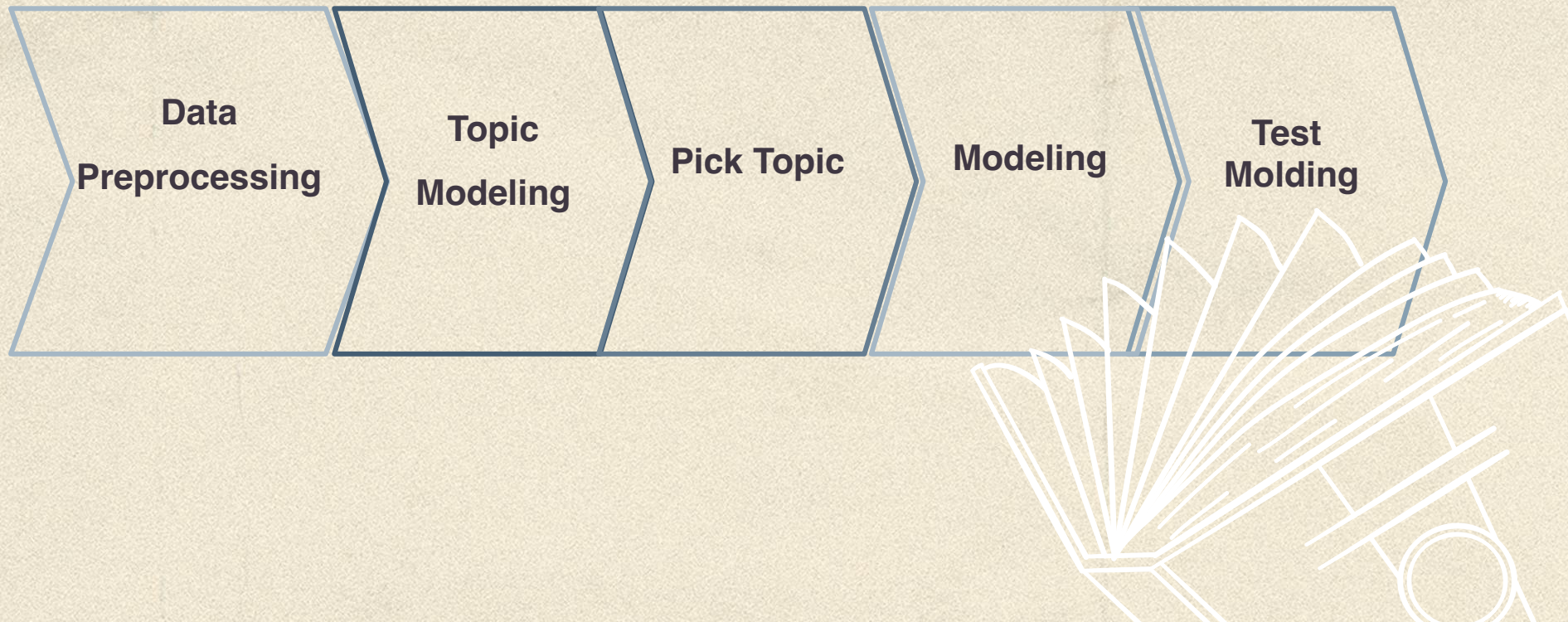


## Visualization



## NLP

# Workflow





# Data Preprocessing

## SubSet Of Data

- Quotes Where The length of word more than 150

## NLP preprocessing Pipeline

- Contraction
- Tokenization
- Remove Punctuations & Number
- Remove Non English word
- Remove rare and frequent words
- Remove stop word
- Lowercase
- Remove Repeating Char
- Spell Corrections
- Stemming
- Lemmatization
- Remove Word Less Than 3
- Remove Verbs

## Another preprocessing

- Check Null values
- Check duplicate values
- Remove Un-wanted Colman

**Data After Cleaning (11286, 10)**



# Topic Modeling



## Modeling

- Latent Semantic Analysis (LSA)
- Non-Negative Matrix Factorization (NMF)
- Latent Dirichlet Allocation (LDA)
- CorEx
- CorEx With Anchors



## Embedding

- CountVectorizer
- Tf-IDF



# CorEx Topic

## **Sadnesses**

death,light,earth,dead,blood,beneath,great,fire,soul.

## **Knowledge**

work,book,public,inform,write,read,abus,medium,group

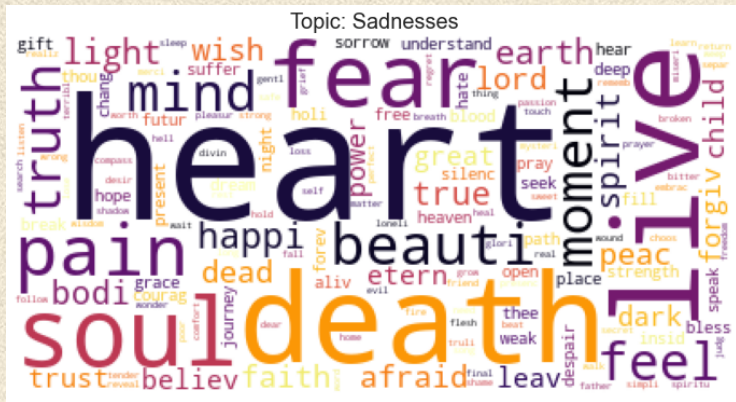
## **Universe**

human,natur,social,societi,individu,moral,modern,polit,scienc

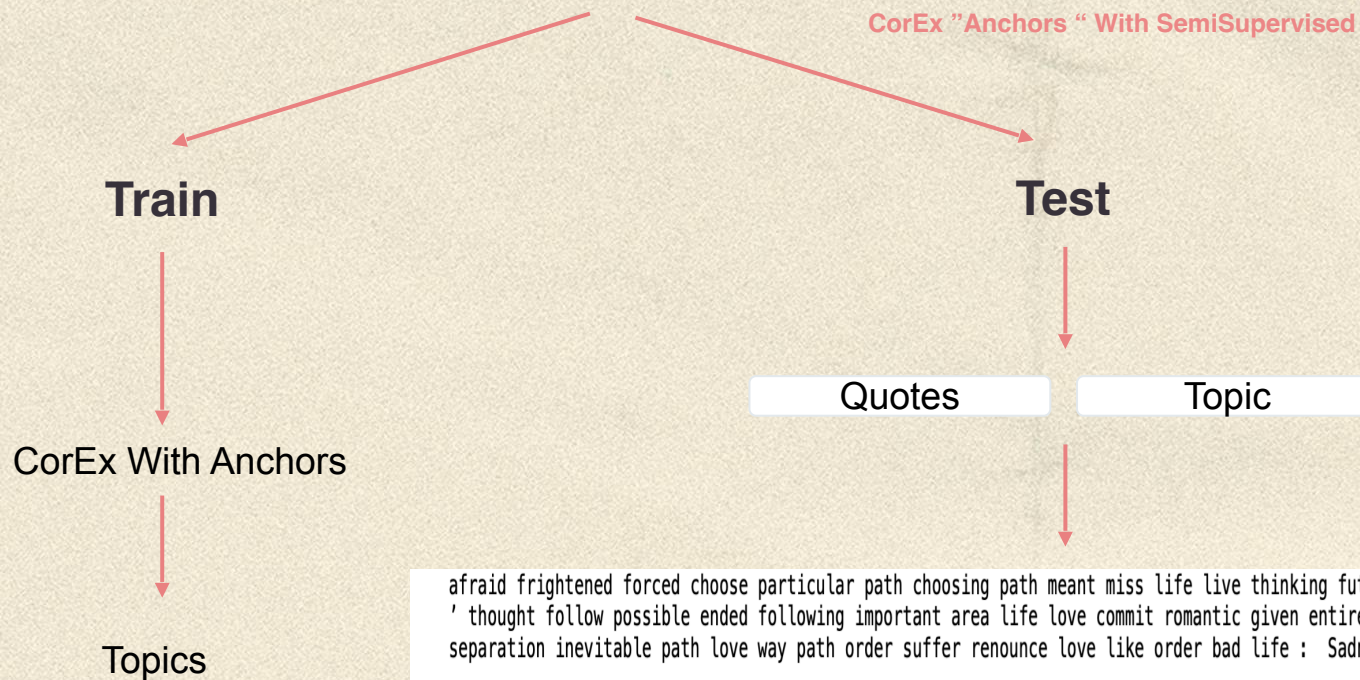
## **Feelings**

head,night,hand,hair,face,breath,mouth,black,look

# Word clouds







# Supervised learning

- Logistic Regression
- Random forest

## Classification Accuracy

	Logistic Regression	Random forest
Train	0.81	0.82
Val	0.67	0.44
F1-score	0.72	0.46



# Challenges And Limitations

## Deep Meaning

You were never created to live depressed, defeated, guilty, condemned, ashamed or unworthy. You were created to be victorious



To Be Or Not To Be

Stop Words





**Feed classifier more features**

**Model**

**Future Work**

**Recomndtion System**

**Deployment**



“

**Thanks**

”