localhost:8501

Navigation

Go to

- ● Introduction
- ○ Dataset
- ○ Features
- ○ Model
- ○ Predict

Are you being sarcastic ???

Bites Makes Right

```
Team members :                                                    −

Josmi Agnes Jose, 20BDA27| Megha Roy, 20BDA41| Ajay Bhargav, 20BDA64| Abraham G.K, 20BDA20
```

Sarcasm refers to the use of words that mean the opposite of what you really want to say, especially in order to insult someone, or to show irritation, or just to be funny.

Aim : We have the Reddit Sarcasm Detection Dataset and our goal is to predict whether a given comment is sarcastic or not.

It's called sarcasm. Look it up.

Navigation

Go to

- ○ Introduction
- ● Dataset
- ○ Features
- ○ Model
- ○ Predict

Select Graph

Dataset ▾

Reddit Sarcasm Dataset.

Let us understand the dataset.

Shape of the dataset `(1010826, 10)`

Head of the dataset

	label	comment	author	subreddit	score	ups	downs	date	created
0	0	NC and NH.	Trumpbart	politics	2	-1	-1	2016-10	2016-10-16 23:
1	0	You do know west teams…	Shbshb906	nba	-4	-1	-1	2016-11	2016-11-01 00:
2	0	They were underdogs ea…	Creepeth	nfl	3	3	0	2016-09	2016-09-22 21:
3	0	This meme isn't funny …	icebrotha	BlackPeopleTwitter	-8	-1	-1	2016-10	2016-10-18 21:
4	0	I could use one of tho…	cush2push	MaddenUltimateTeam	6	-1	-1	2016-12	2016-12-30 17:
5	0	I don't pay attention …	only7inches	AskReddit	0	0	0	2016-09	2016-09-02 10:
6	0	Trick or treating in g…	only7inches	AskReddit	1	-1	-1	2016-10	2016-10-23 21:
7	0	Blade Mastery+Masamune…	P0k3rm4s7	FFBraveExvius	2	-1	-1	2016-10	2016-10-13 21:
8	0	You don't have to, you…	SoupToPots	pcmasterrace	1	-1	-1	2016-10	2016-10-27 19:
9	0	I would love to see hi…	chihawks	Lollapalooza	2	-1	-1	2016-11	2016-11-21 23:
10	0	I think a significant …	ThisIsNotKimJongUn	politics	92	92	0	2016-09	2016-09-20 17:

Number of classes `2`

Navigation

Go to

○ Introduction
● Dataset
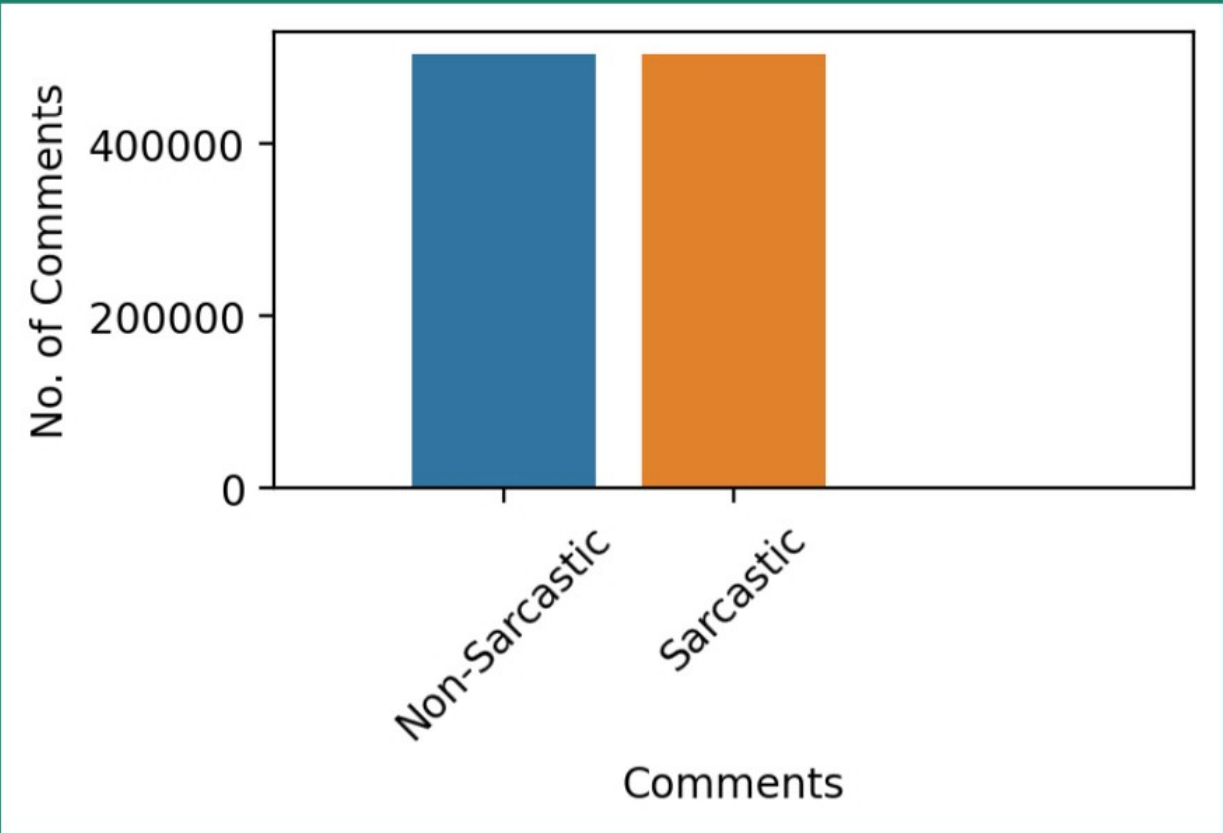○ Features
○ Model
○ Predict

Select Graph

Comments ▾

Dataset

Comments

Subreddit

Wordcloud

Reddit Sarcasm Dataset.
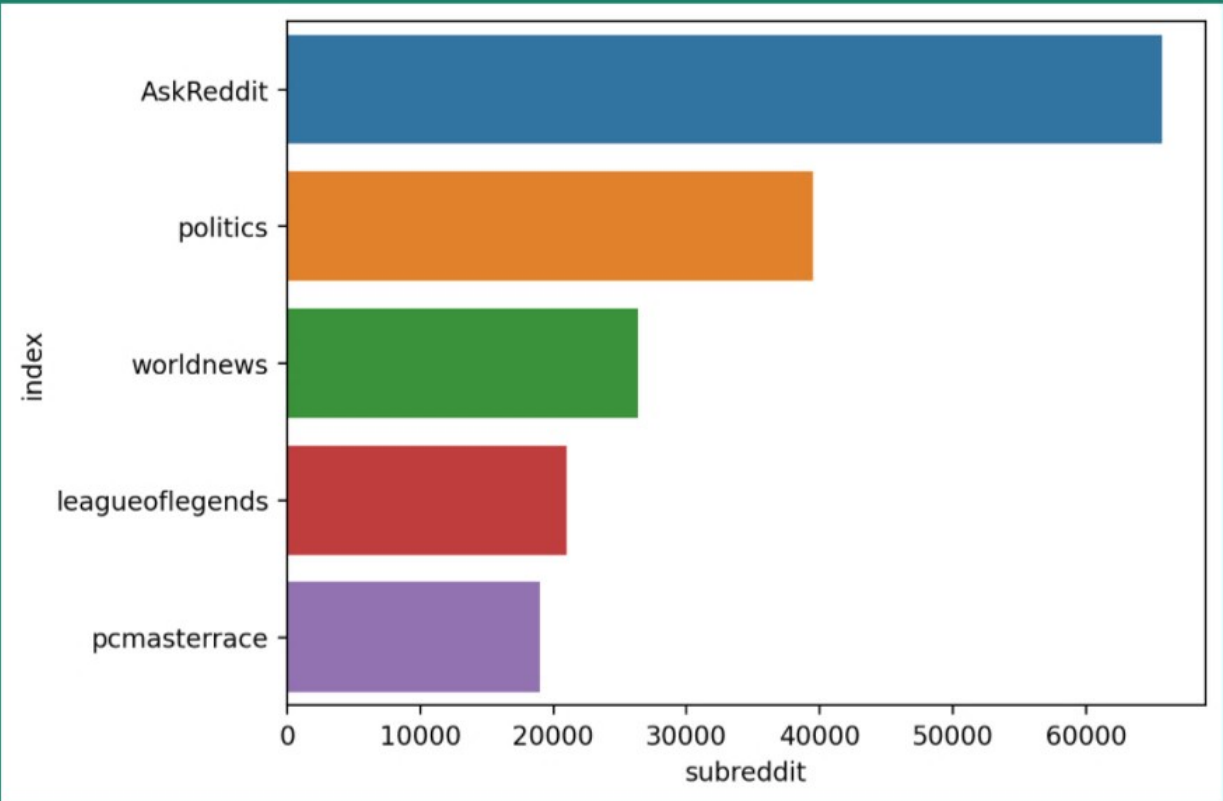
Count of Sarcastic and Nonsarcastic Comments

localhost:8501

Apps · √ Malayalam Typin... · Dashboard | Hacker... · Slack | community-... · MSc. Big Data Anal... · GitHub - IvyZX/mus... · GitHub - vskadanda... · Fourier Transforms... · ipython_audio · Reading list

Navigation

Go to

○ Introduction
● Dataset
○ Features
○ Model
○ Predict
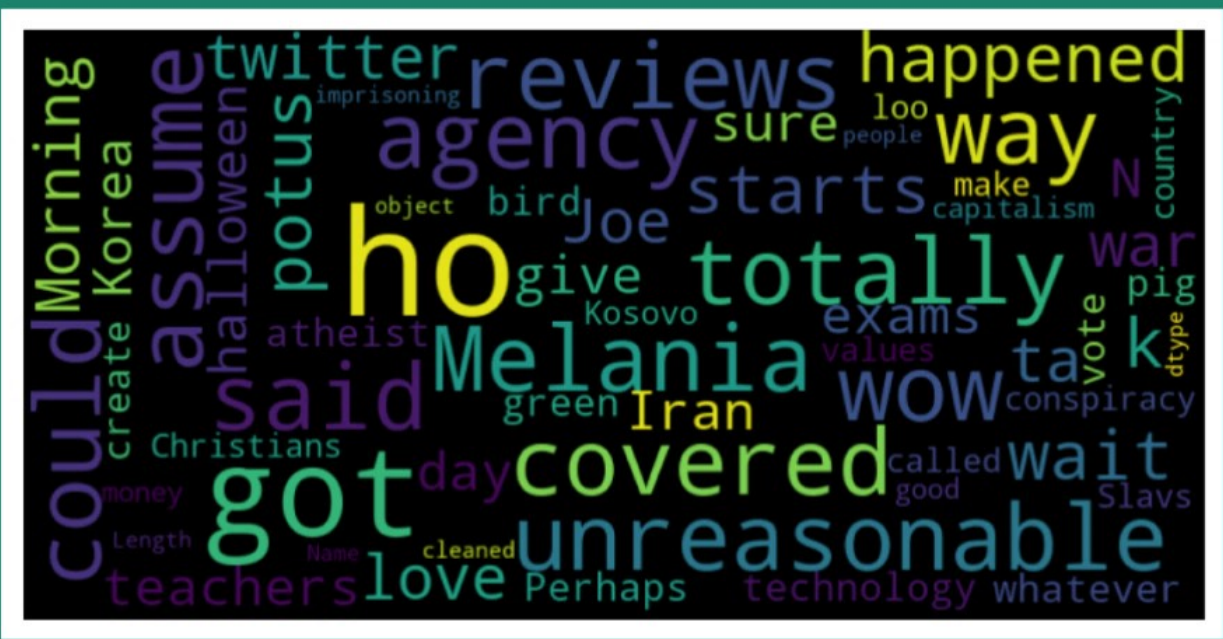
Select Graph

Wordcloud ▾

Reddit Sarcasm Dataset.

Wordcloud for comments.

Sarcastic or Non-Sarcastic

Sarcastic ▾

Navigation

Go to

- ○ Introduction
- ◉ Dataset
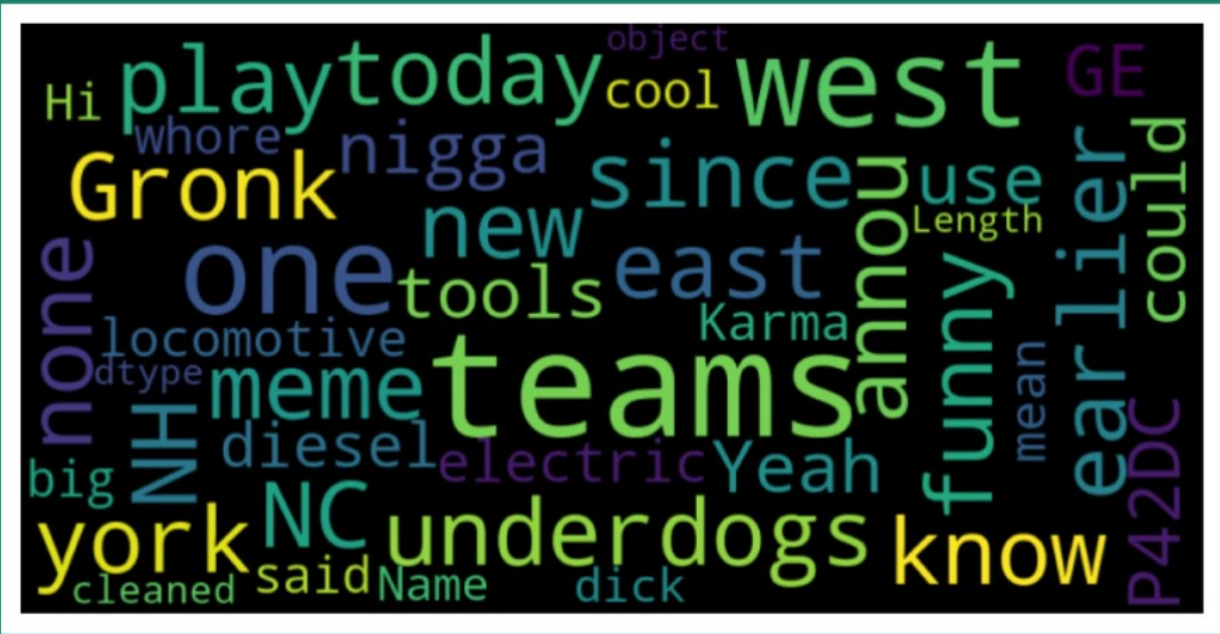- ○ Features
- ○ Model
- ○ Predict

Select Graph

Wordcloud ▾

Reddit Sarcasm Dataset.

Wordcloud for comments.

Sarcastic or Non-Sarcastic

Non Sarcastic ▾

Extracted Features.

Head of the dataset

	Comment	Capital words	Total Characters	Unique Characters	Characters repeated	Capital words	Sente
0	NC and NH.	2	10	8	0	2	
1	You do know west teams…	0	74	21	0	0	
2	They were underdogs ea…	0	121	30	0	0	
3	This meme isn't funny …	0	60	21	0	0	
4	I could use one of tho…	1	31	14	0	1	
5	I don't pay attention …	2	112	26	0	2	
6	Trick or treating in g…	0	45	19	0	0	
7	Blade Mastery+Masamune…	1	31	21	0	1	
8	You don't have to, you…	0	62	21	0	0	
9	I would love to see hi…	1	33	16	0	1	

Number of Features extracted 16

Extracted Features.

Navigation

Go to

- ○ Introduction
- ○ Dataset
- ● Features
- ○ Model
- ○ Predict

Select Graph

Swear Words ▾

Swear Words.

Extracted Features.

Capital words.
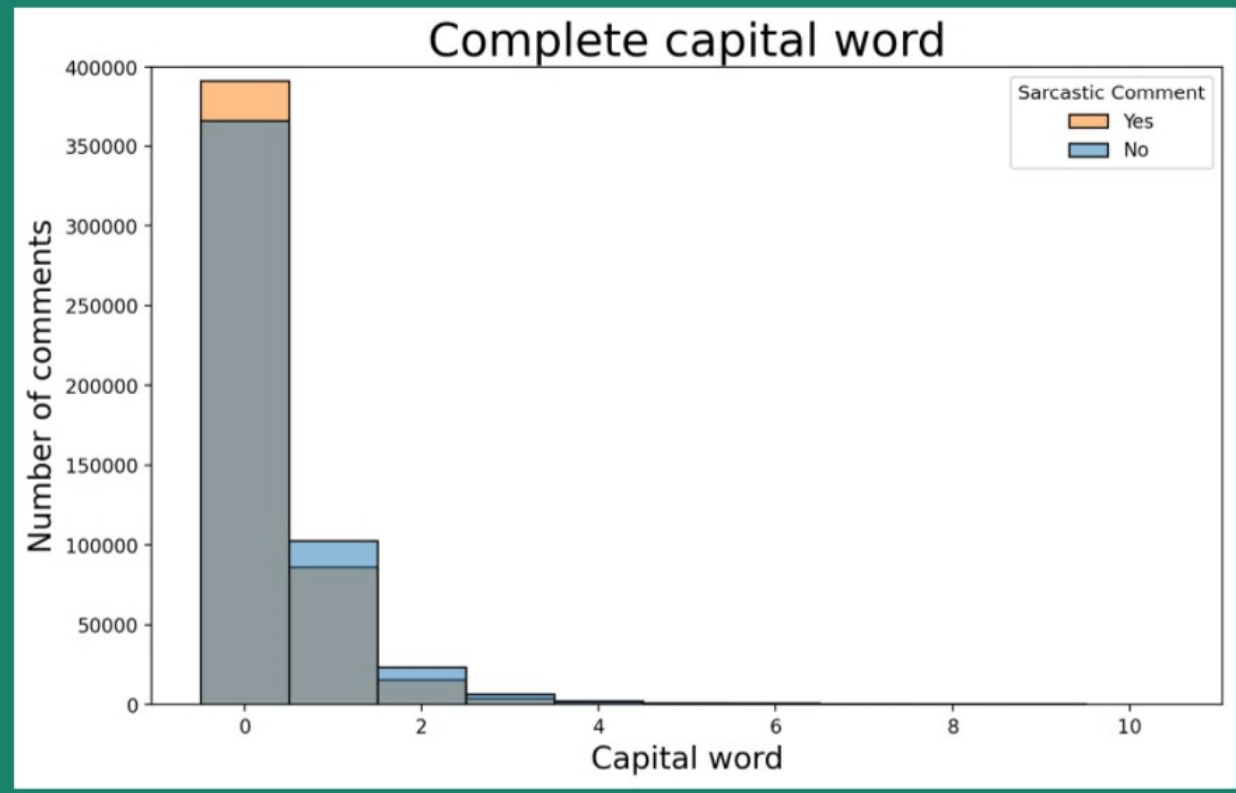
rough · Streamlit

localhost:8501

Apps ✓ Malayalam Typin... Dashboard | Hacker... Slack | community-... MSc. Big Data Anal... GitHub - IvyZX/mus... GitHub - vskadanda... Fourier Transforms... ipython_audio Reading list

Navigation

Go to

○ Introduction
○ Dataset
● Features
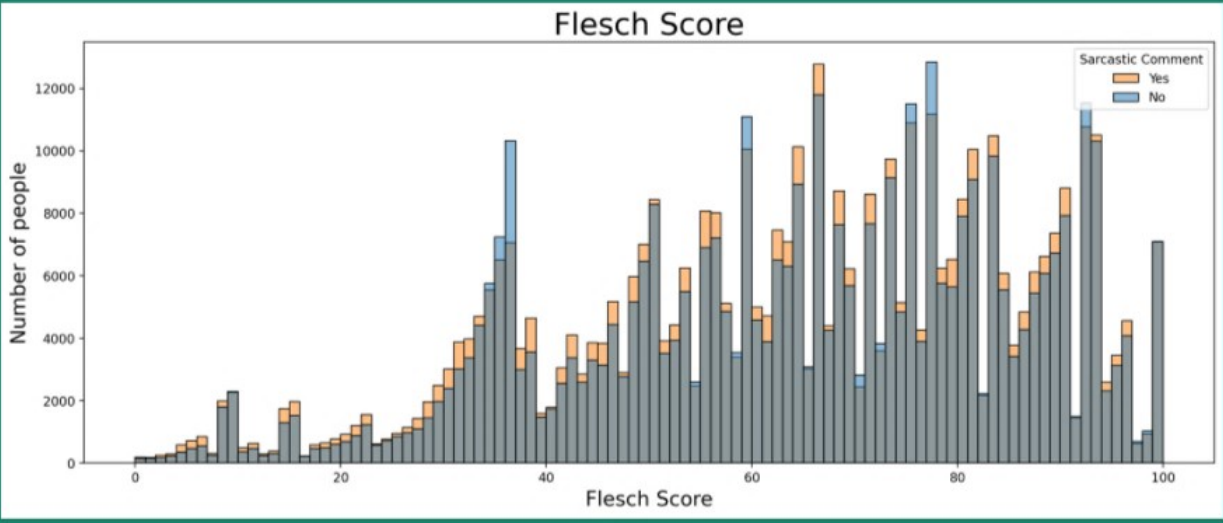○ Model
○ Predict

Select Graph

Complexity ▾

Extracted Features.

Complexity.

Measure of Complexity

Flesch Score ▾

Navigation

Go to

- ⦿ Introduction
- ⦿ Dataset
- 🟢 Features
- ⦿ Model
- ⦿ Predict

Select Graph

Complexity ▾

Extracted Features.

Complexity.

Measure of Complexity

Flesch Score ▾

Flesch Score

Sentence Length

Comment Length

Syllables per word

rough · Streamlit

localhost:8501

Apps   ✓ Malayalam Typin...   Dashboard | Hacker...   Slack | community-...   MSc. Big Data Anal...   GitHub - IvyZX/mus...   GitHub - vskadanda...   Fourier Transforms...   ipython_audio »   Reading list

Navigation

Go to

◯ Introduction
◯ Dataset
� Features
◯ Model
◯ Predict

Select Graph

Sentiment|

Y limit

22500

25000   225000

X Range

-1.10   1.10

-1.10   1.10

Extracted Features.

Sentiment in sarcasm.

Polarity or Intensity

Polar

Extracted Features.
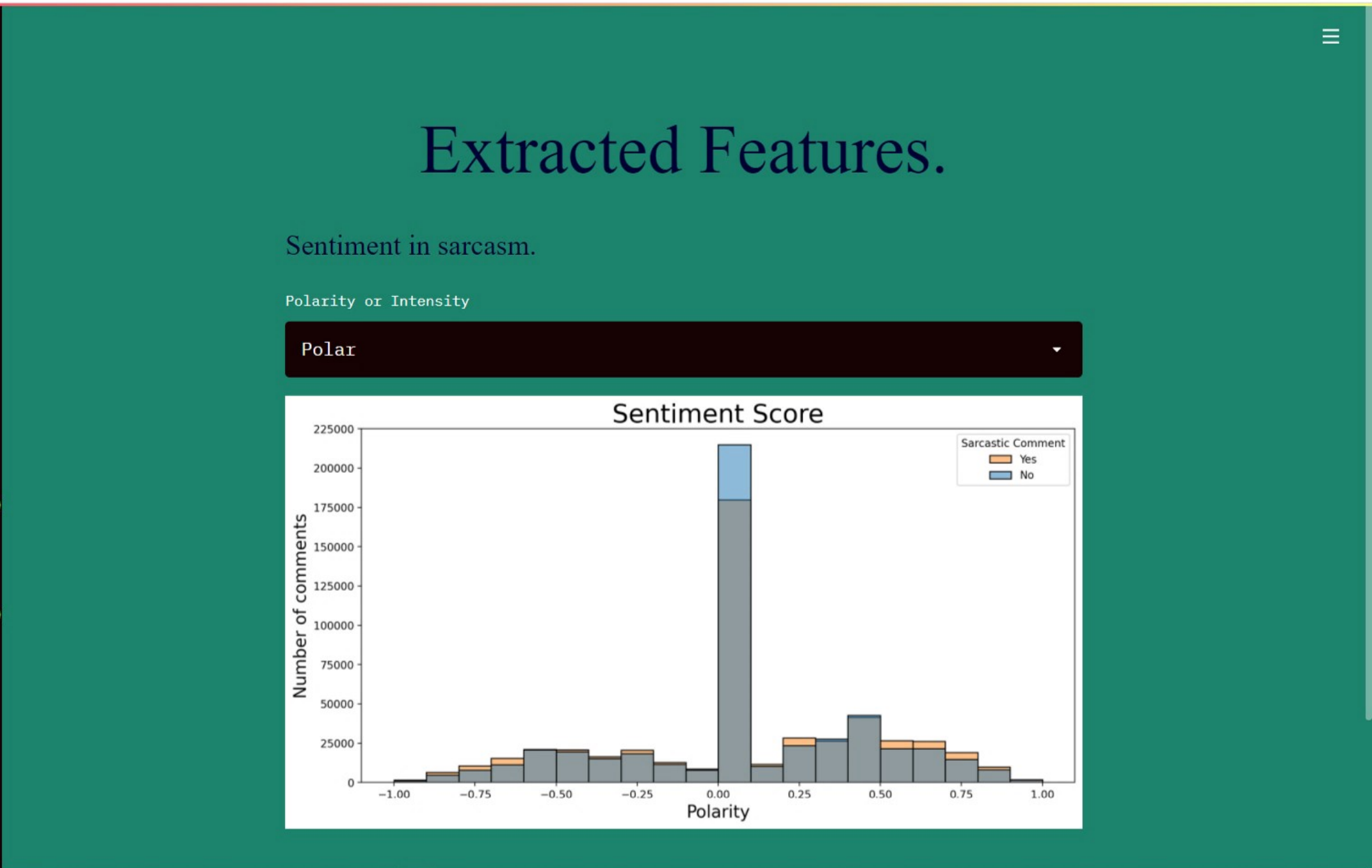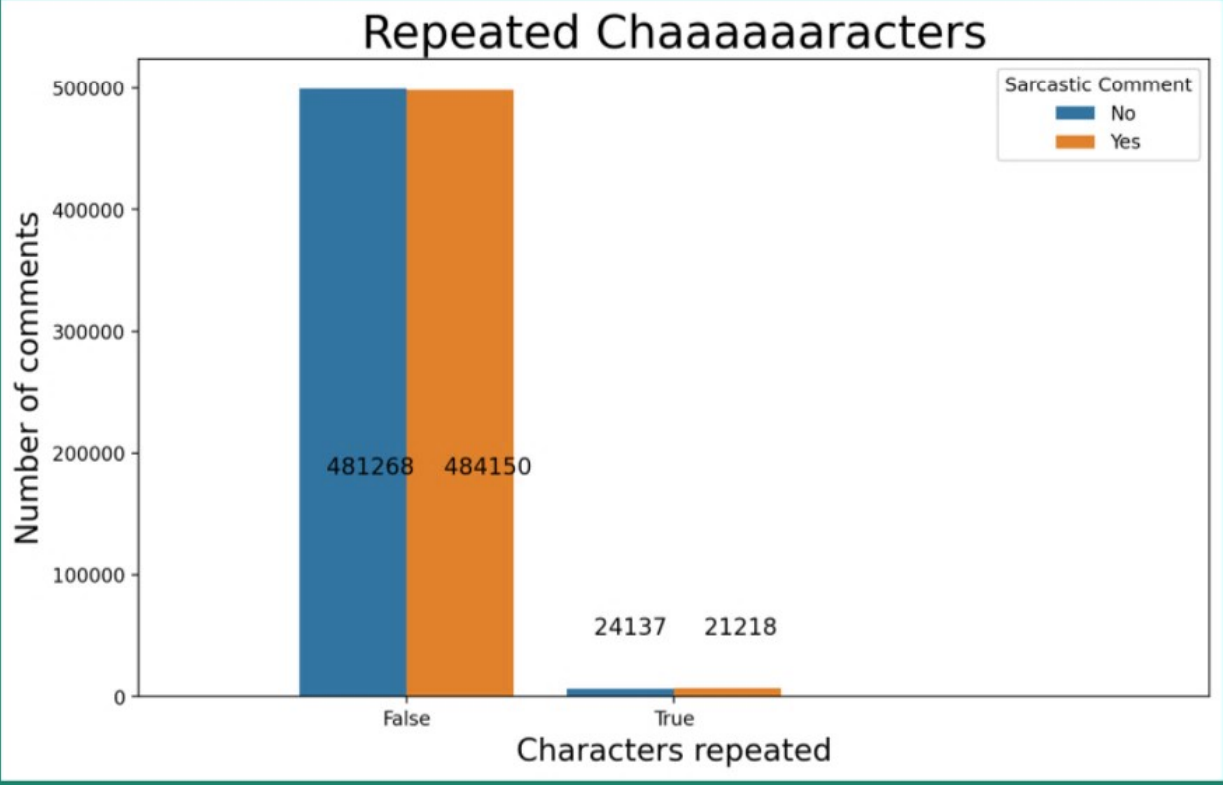
Navigation

Go to

- ○ Introduction
- ○ Dataset
- ● Features
- ○ Model
- ○ Predict

Select Graph

Repeated Characters ▾

Dataset
Swear Words
Capital Words
Complexity
Sentiment
Punctuation
Repeated Characters

Repeated characters.

Machine Learning Model.

Logistic Regression

	precision	recall	f1-score	support
0	0.7328	0.7711	0.7515	100954
1	0.7591	0.7196	0.7388	101201
accuracy	0.7453	0.7453	0.7453	0.7453
macro avg	0.7460	0.7453	0.7451	202155
weighted avg	0.7460	0.7453	0.7451	202155

Train score : 0.7465453403213879

Confusion matrix:

	0	1
0	77845	23109
1	28381	72820

Machine Learning Model.

Random Forest Regression

	precision	recall	f1-score	support
0	0.7271	0.7534	0.7400	100954
1	0.7448	0.7180	0.7311	101201
accuracy	0.7356	0.7356	0.7356	0.7356
macro avg	0.7360	0.7357	0.7356	202155
weighted avg	0.7360	0.7356	0.7356	202155

Train score : 0.9976057916098825

Confusion matrix:

	0	1
0	76055	24899
1	28542	72659

localhost:8501

Navigation

Go to

- ○ Introduction
- ○ Dataset
- ○ Features
- ○ Model
- ● Predict

Let us do some Predictions.

Here you can check whether a comment is sarcastic or not.

Type the comment here :

```
This is a genuine compliment!!!
```

Select Subreddit.

AskReddit ▼

Comment Score:

-10000   —   +

Predict