Reddit - Classification of posts

Evan, Su Ying, Shamini, Wilson

WarFrame Apex Legends

MMORPG/shooter (3rd)

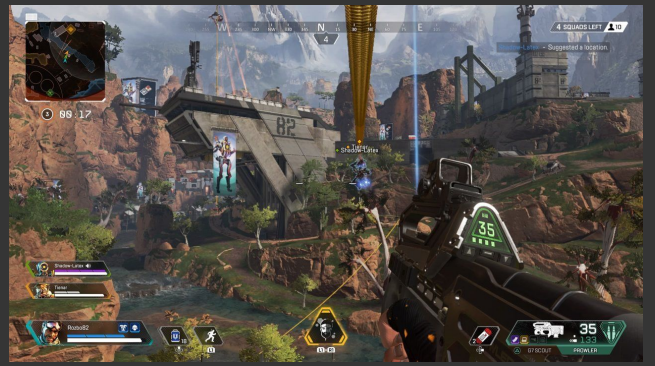
Free to play

Launched 2013

First Person Shooter (FPS)

Free to play

Launched 2019



MANCHESTER
UNITED

YOU'LL NEVER WALK ALONE
LIVERPOOL
FOOTBALL CLUB
EST·1892

Manchester Utd vs Liverpool FC

English Premier League

Champions - 20 times

RED

English Premier League

Champions - 18 times

RED

Fierce rivals who **REALLY** hate each other

No Sleep

VS

AsianBeauty

Caution

You have been warned, read them at your own risk

Ethereum
Litecoin
+1.2
Bitcoin
+12.8
Ripple
Ethereum

APPLE (AAPL) STOCK NAS
167.47 USD 0.04 (0.02%) 02:11:04 PM EST BTT
ADD
SHARE
Prev. Close 167.43
Open 167.43
Market Cap (USD) 845.36 B
Volume (Qty.) 1,289,345
Day Low 166.69
Day High 168.62
52 Week Low 125.40
52 Week High 180.10
INTRADAY 1W 1M 3M 6M YTD 1Y 3Y 5Y 10Y MAX
CHART OPTIONS
EXCHANGE: BTT
169.00
168.50
168.00
167.50
167.00
166.50
+0.75%
+0.5%
+0.25%
0%
-0.25%
-0.5%
10:00 AM
11:00 AM
12:00 PM
1:00 PM
2:00 PM
3:00 PM
4:00 PM

IT'S DOWN 10%: SELL SELL SELL!!!
IT'S DOWN 10%: THINK I'LL BUY MORE
Btcsections

STOCKS AND BITCOIN

- Stocks is the ownership of a company.
- Bitcoins are crytocurrencies. Send and receive money without being linked to a real identity.
- People are using reddit to gather information when they want to invest.
- Many newcomers believe investing in bitcoins are the same as buying stocks.

Why Reddit

Straight forward API

Entertaining

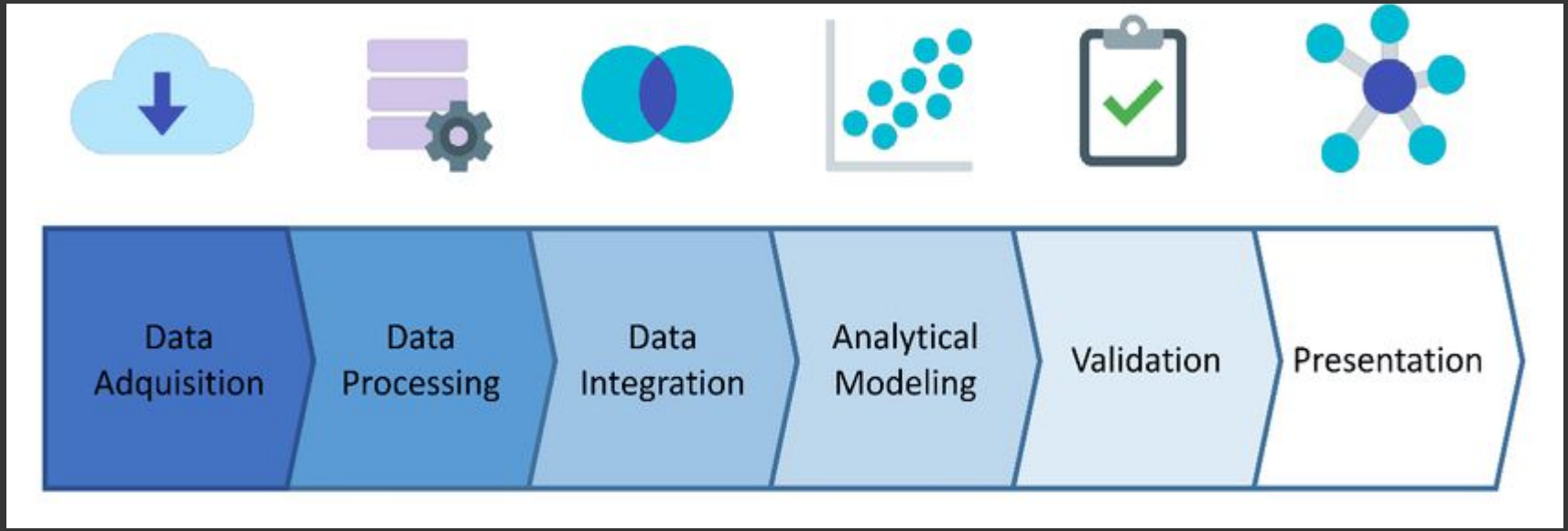
Forum-like

Active community

First time be like ???

Nobody: 'Cause it is part of the project requirements. Duhhh.'

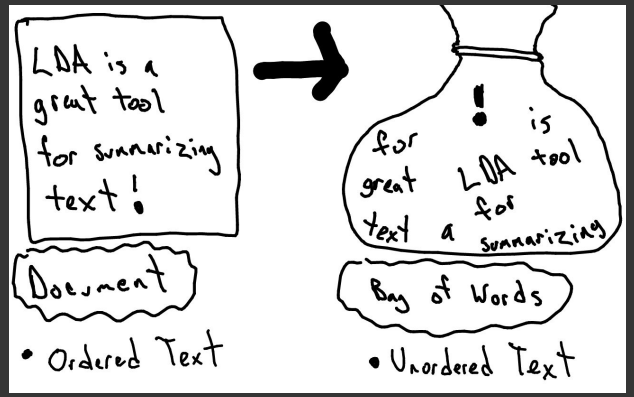
Pipe


Data Adquisition
Data Processing
Data Integration
Analytical Modeling
Validation
Presentation

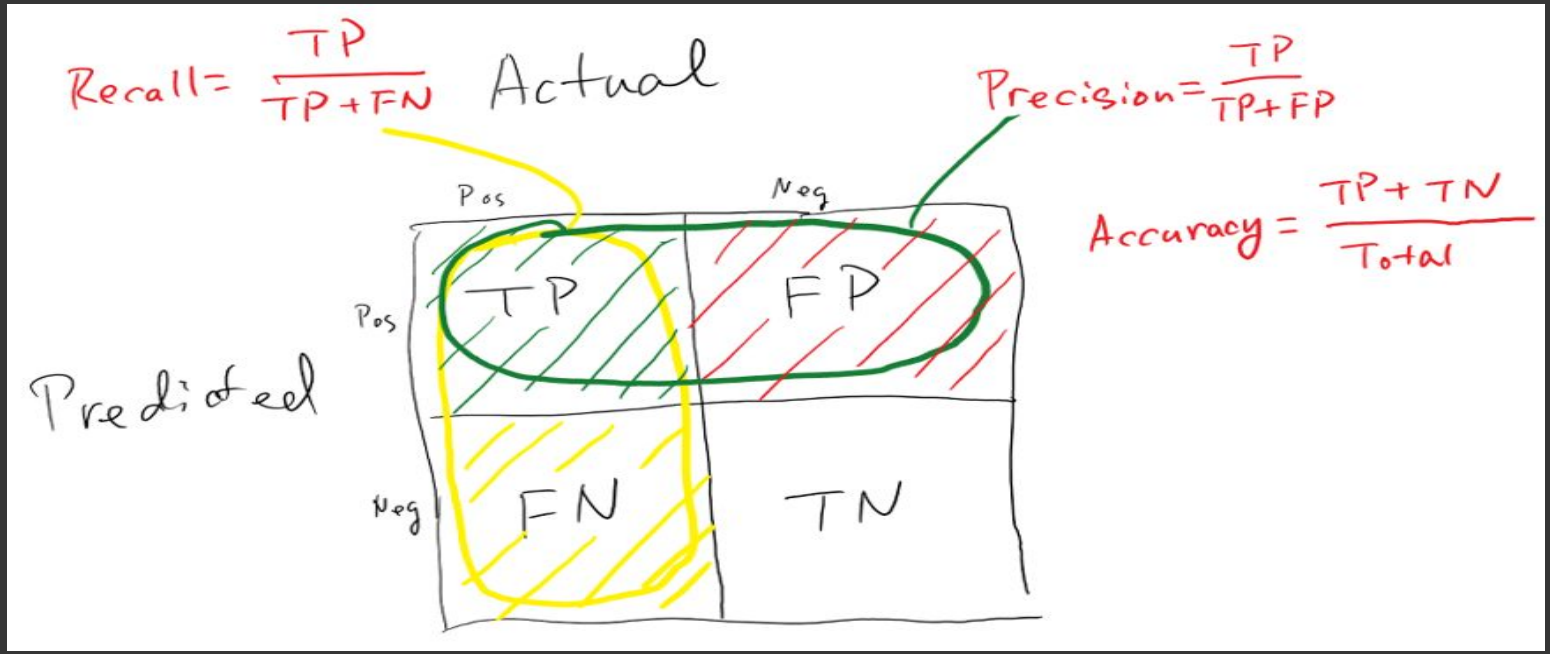
Preprocessing

- Cleaning up any null values
- Token pattern
- Lemmatization
- TF-IDF vs CountVectorizer
  - Stop_words
  - *ngram range*

"Now i know my abc."



Model Evaluation



Conclusion

- Machine Classification Predictor
  - Model will be able to accurately suggest the most suitable subreddit a new post should belong to.
  - Flag out irrelevant posts before users report