



## **Project 3 : Web APIs and NLP**

Ramesh Babu




# What is this ....?

I thought ,

I need to know **Something** to learn **Anything** ..

But this course taught me ..

You need not know **Anything** to learn **Something** ..



I should have willingness to  
accept my insufficiency...

this is the first step towards learning...





# Who exposed this thought....?

I needed a Phd person to expose this to me....



And his name is **Jacob Frias Koehler, PhD**

Teacher of Mathematics and Computing

New York, United States



So my topic naturally fell into.....

Ask - Philosophy	Change - My - View
------------------	--------------------



## Ask - Philosophy

Is there such thing as objective morality?

What is more harmful, complete trust in everything or complete distrust in everything.

Reading list for A/H Philosophy

can absurdist dark memes be morally justified?

How do we reconcile with the self-undermining nature of Wittgenstein's ladder? "He must, so to speak, throw away the ladder after he has climbed up it."

what determines duty?

Is it ever right to sacrifice some to save the many?

Question About Occam's Razor and the Multiverse Theory

What does it mean to say concrete entities can have causal relationships while abstract entities cannot?

What makes pseudoscientific treatment convincing to some people?

Does talk of any value presuppose life having meaning? (also the contrapositive: If life had no meaning, would all values not exist?)

Cambridge university making textbooks free online, what,Âs good to snag while available?

What is Ecofascism?

Is it possible that an innocent could murder someone?

Mods in this subreddit

Is there a philosopher who considers the idea that all human beings end up in an afterlife consisting of eternal torture, regardless of their beliefs and actions?

The Phaedo

Why is stealing immoral??

What ethical considerations or approaches bear on encouraging or discouraging intelligence in non-human animals?

Being and Nothingness

Does Hobbes uphold the idea that social cooperation can function within the state without the need or the development of an absolute monarchy when he notes cases in which small groups banded together like families?



## Change my view

The Nintendo Wii is the best video game console ever.

Lawyers rule the world

There is not a rationally consistent abortion rights argument

In the Ballad of Buster Scruggs, the quadruple amputee does not \*spoiler\*

the majority of posters here have little to no interest in actually having their views challenged and instead simply wish to show off their "big brain" intelligence and breath-taking rhetorical skills

High school track is the hardest high school sports

Garth Brooks deserves Entertainer of the Year of Eric Church at the CMAs

High school track is the hardest high school sports

It's called the "China virus" because....IT CAME FROM CHINA

Suicide is selfish

Led Zeppelin is a superior to the Beatles.

xxxxxxxxxxxxx Should be Legalized Across America

I am tired of seeing celebrities with extreme amounts of money getting praised for using there money in a charitable way.

Even though it is true that COVID-19 came from China, calling it the "Chinese virus" will only spark further racism against Asian Americans/Asians living as a minority in Western nations.

The "Wuhan flu" would be fine, in my opinion, even.

We need to start using words like ..., etc.

The quality of new video games has diminished because of lack of imagination on the part of developers.

We are not gonna change anything after this pandemic.

It's unreasonable that the U.S. is still getting so much flak over its invasion of Native American land 400 years ago, especially as Native American tribes did the same to predecessor tribes



What is the  
source and  
how I got  
these posts...

It took about 15 mins to pull 10,000 posts without rate limiting  
`base_url = 'https://api.pushshift.io/reddit/search/submission'`

pull a 100 posts in one shot

```
def get_data(subreddit, after):  
    params = {  
        'subreddit': subreddit,  
        'size': '500',  
        'after': after  
    }  
  
    base_url = 'https://api.pushshift.io/reddit/search/submission'  
    res = requests.get(url, params)  
    raw_data = res.json()  
    posts = raw_data['data']  
    df = pd.DataFrame(posts)  
    return df
```

simple call to initiate

```
subreddit = 'askphilosophy'  
df_final_list = start_process()  
  
subreddit = 'changemyview'  
df_final_list = start_process(subreddit)
```

Iterate over 100 times to get 10\_000 posts

```
def start_process(subreddit):  
    start_time = time.time()  
    df_final = pd.DataFrame(columns=['subreddit', 'selftext', 'title', 'created_utc'])  
    after = '1584583125'  
    subreddit = subreddit  
    for i in range(100):  
        print(f" started .. {i}")  
        df_raw = get_data(subreddit, after)  
        df_temp = df_raw[['subreddit', 'selftext', 'title', 'created_utc']]  
        df_final = df_final.append(df_temp)  
        after = df_temp['created_utc'].max(axis = 0)  
    print(f"--- %s seconds ---" % (time.time() - start_time))  
    return df_final
```



# Got fairly distributed data

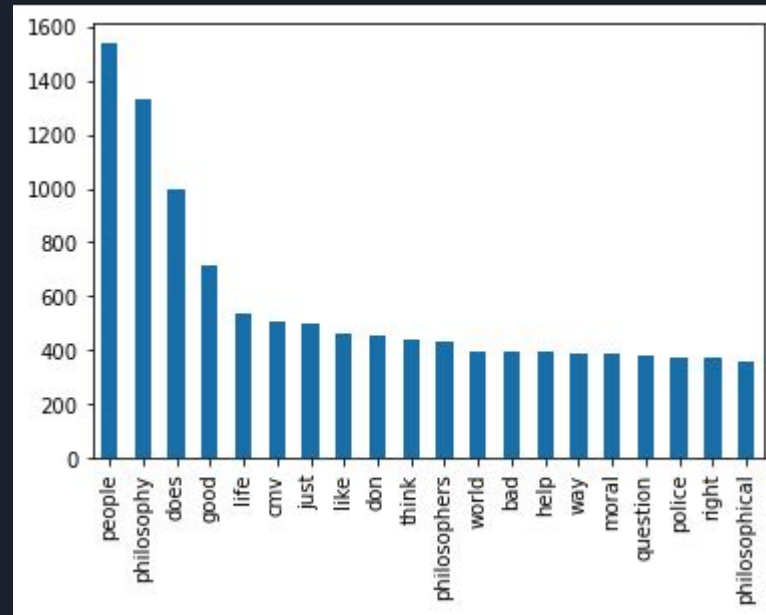
Ask - Philosophy	50.25 %
Change - My - View	49.74 %



# EDA: Exploratory Data Analysis

1. For some reason the people in Change in View wanted to add “CMV: “ in all their topic .. because they want to tell the world they will `Change Their View` pretty soon
2. There were people want to add 26 times their view again and again .. to prove that they will not Change their View
3. I need to deal with profane and cursing words

people	1538
philosophy	1329
does	995
good	711
life	536
...	
feel	130
important	129
term	129
different	129
china	128

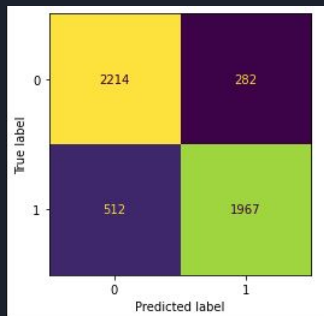


# Model Comparison

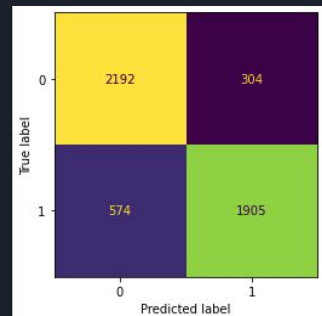
	LogisticRegression	Decision Tree	Random Forest	Multinomial NB
Train Score	86.15	87.68	75.57	83.71
Test Score	84.04	82.35	74.49	82.77
Precision	87.46	86.23	92.96	64.72
Recall	79.34	76.84	52.80	94.71
Specificity	79.34	76.84	52.80	94.71
Balanced Accuracy	84.02	82.33	74.41	71.71
Accuracy Score	84.04	82.35	74.49	71.63
False Positive Rate	11.30	12.17	3.96	51.28

## Confusion Matrix

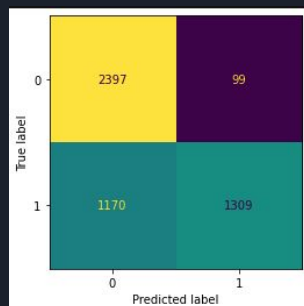
### Logistic Regression



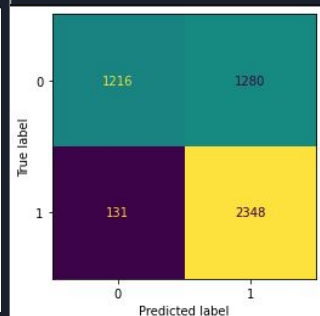
### Decision Tree



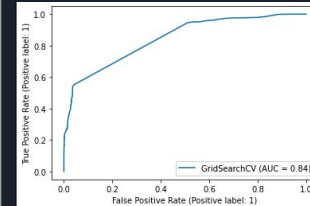
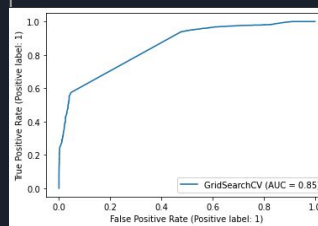
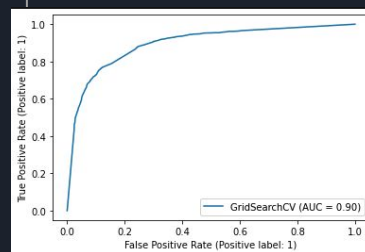
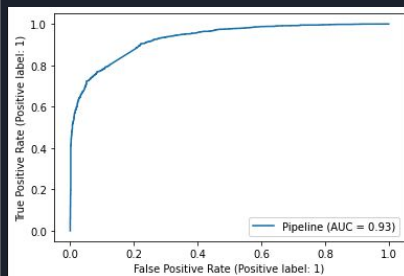
### Random Forest



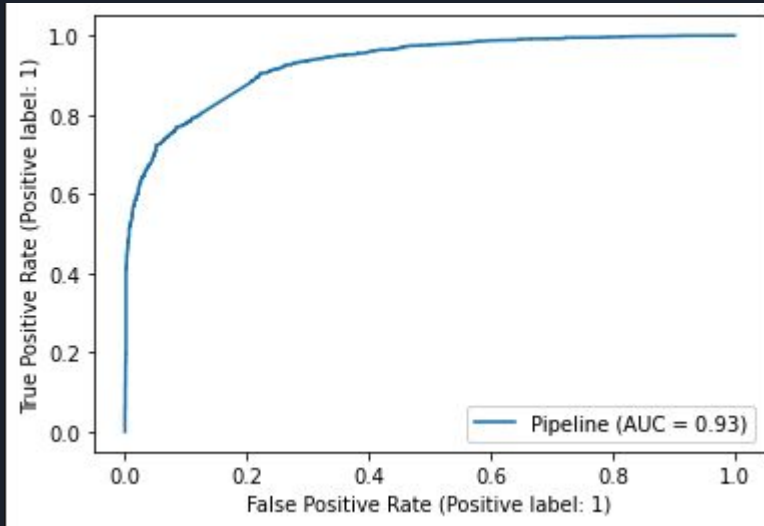
### Multinomial NB



## ROC AUC



## Interpreting ROC AUC



- If you have an ROC AUC of 0.5, your positive and negative populations perfectly overlap and your model is as bad as it can get.
- If you have an ROC AUC of 1, your positive and negative populations are perfectly separated and your model is as good as it can get.
- The closer your ROC AUC is to 1, the better. (1 is the maximum score.)
- If you have an ROC AUC of below 0.5, your positive and negative distributions have flipped sides. By flipping your predicted values (i.e. flipping predicted 1s and 0s), your ROC AUC will now be above 0.5.

## Ask - Philosophy

<b>philosophy</b>	1.636010
<b>philosophers</b>	1.172482
<b>kant</b>	0.839584
<b>philosophical</b>	0.766353
<b>philosopher</b>	0.680404
<b>question</b>	0.617820
<b>nietzsche</b>	0.606616
<b>plato</b>	0.583498
<b>realism</b>	0.569104
<b>socrates</b>	0.552234

## Change My View

<b>cmv</b>	-1.083951
<b>trump</b>	-0.516874
<b>usa</b>	-0.489465
<b>gun</b>	-0.478591
<b>biden</b>	-0.457380
<b>healthcare</b>	-0.452885
<b>looting</b>	-0.415743
<b>protests</b>	-0.412229
<b>protesting</b>	-0.398190
<b>police</b>	-0.367203





# Conclusion

Not knowing the meaning of a sentence, algorithms can accurately predict any topic thrown at it. But you cannot expect any changes in it .

On The Contrary:

We may find difficult some time to differentiate between Philosophy and Change-in-View.  
Once we understood, it will change our lives .



# Food for thought .....

Feb. 18, 1996

**Deep Blue** won **Chess Grandmaster Garry Kasparov** on the opening chess game,

Does **Deep Blue** takes **Deep Pink** for a date on that evening ....???