

**Which side
are you
on?**

Classifying a debate
discussion text
using NLP



How do online conversations go?

#debates 🗳️



??????



This is the
ans.....aw...



How dare
you!!!!



##\$@!!....



Problem Statement

To allow users/moderators to classify posts on various channels based on their preferences

Data from reddit using the pushshift.io api

2000 posts and 2000 comments from r/evolution

2000 posts and 2000 comments from r/Creations

2000 posts and 2000 comments from r/DebateEvolution

How to we achieve this?

Processing text from r/evolution and r/Creation:

- Tokenizing - regular expressions to extract only words
- Stopwords - english + ["https" ,"com","www",....]
- Lemmatizing
- Stemming

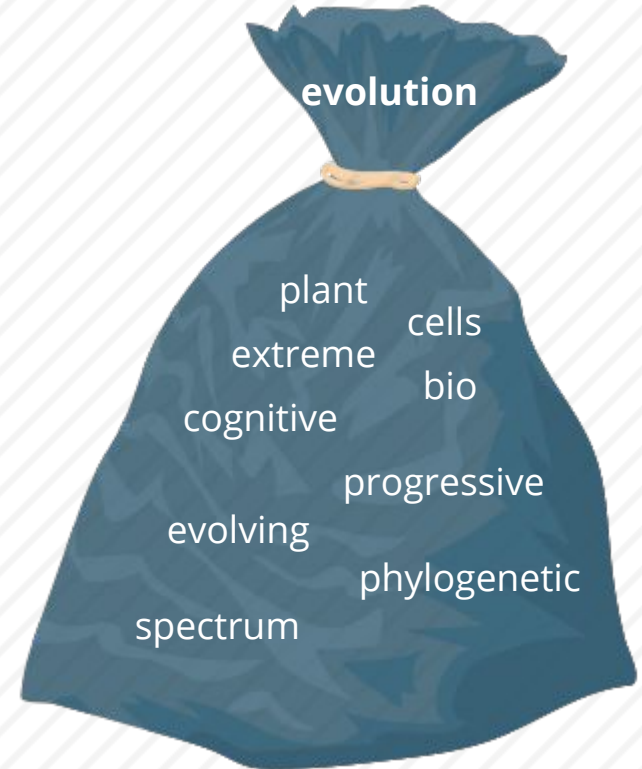
Converting it into bag of words using CountVectorizer or TfidfVectorizer

Running data through several classification models.

Evaluating the models

Choosing the best model and using that to classify data from
r/DebateEvolution

Bag of words using coefficient > 0.5



Best model out of 8640 models

Vectorizer

TfidfVectorizer() :
5000 - max features
(1,3) -n_gram range
English - stopwords

Classifier/Estimator Scores

LogisticRegressionCV():
0.1 - C
500 - max iteration
12 - penalty

Baseline score: .50

CV Score : .81

Train Score: .91

Test Score:.82

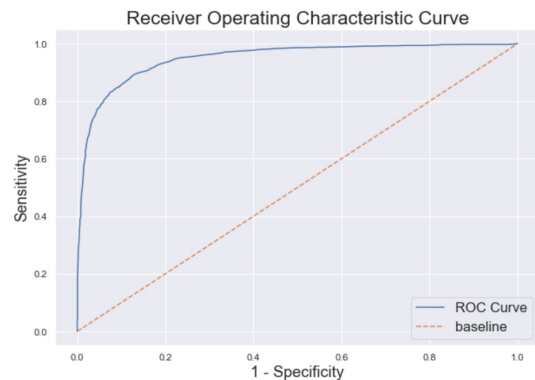
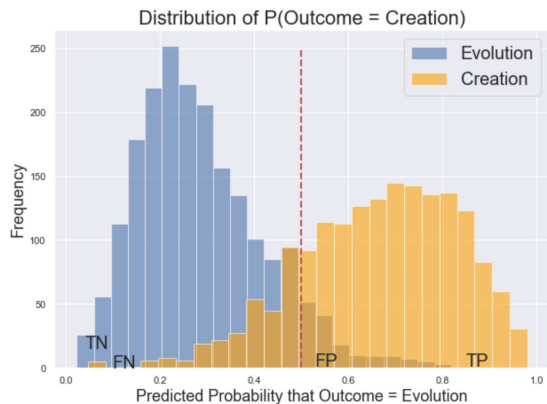
loss function (log loss): 6.54

Columns used:

Title

Selftext

Comment body



Classifying r/DebateEvolution

1288 posts got classified as Creation posts

712 posts got classified as Evolution posts

E.g. Evolution posts : Programmed Cell Death Is Vital to Life, but Where'd It Come From?

E.g Creation posts: Giving Up Darwin

Sentiment Intensity Analysis on all the subreddits :

More than 70% had a neutral tone.

Less than 20% posts had a negative tone.

Next Steps

Identify words/phrases commonly used by popular discussions from both the sides

To find out if there is a way to flag a post if the words are leaning towards an abusive language

Identify key arguments used to refute evolution