



R&D ML Workshop Tweet Classification

ROCS | Kayla Dubnick | dubnickk@vmware.com

2018-07-18 - 2018-08-31

The training set is 3,925 tweets that either acknowledge or deny the existence of climate change.

In **tweets-train.csv**:

Column 1 is the tweet itself.

Column 2 is the label. It's either "Yes" or "No" for acknowledgment and denial, respectively.

Your task is to build a classifier the correctly classifies the tweets to match the labels.

In **tweets-eval.csv**:

Column 1 is the ID of the evaluation tweet.

Column 2 is the text of the evaluation tweet.

Output evaluation file format:

Column 1 is the ID from tweets-eval.csv.

Column 2 is your predicted class, either "Yes", or "No".

Output Example:

```
ID, EXISTENCE
1, Yes
2, Yes
3, No
4, Yes
...
```

Leaderboard

#	Name	Kernel	Score	Entries
1	Bin Zan	en+votingClassifier	0.73333	3
2	Alex RN	TFIDF	0.68667	1
3	Rohan Thakkar	TF_IDF	0.58667	2
4	Rick Battle	Coin Flip	0.48333	1

[SUBMIT ANSWERS](#)[SEE ALL ENTRIES](#)

Good luck!

Files

[classify.py](#) (3.5 KB)

[eval.py](#) (1.2 KB)

[tweets-eval.csv](#) (34.3 KB)

[tweets-train.csv](#) (454.5 KB)

Comments

What are your thoughts?

ADD COMMENT