

# Sentiment Analysis with One-Hot Encoding

# Sentiment Analysis with OHE

## Dataset

- 27481 tweets
  - Text and corresponding sentiment
  - file Data/Tweets.csv
- 
- Source:  
<https://www.kaggle.com/code/irasalsabila/twitter-sentiment>

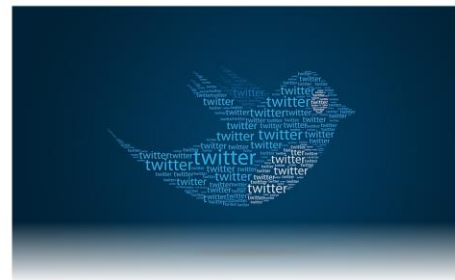
### Twitter Tweets Sentiment Dataset

Twitter Tweets Sentiment Analysis for Natural Language Processing



Data Card Code (8) Discussion (0)

#### About Dataset



Usability ⓘ  
10/00

License  
CC0: Public Domain

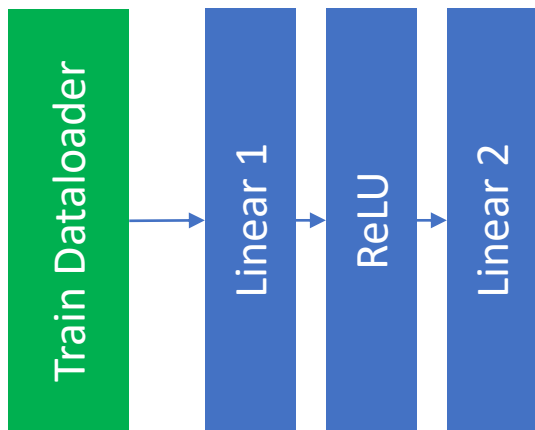
Expected update frequency  
Annually

| textID     | text  | selected_text                       | sentiment |
|------------|---|-------------------------------------|-----------|
| cb774db0d1 | I'd have responded, if I were going           | I'd have responded, if I were going | neutral   |
| 549e992a42 | Sooo SAD I will miss you here in San Diego!!! | Sooo SAD                            | negative  |
| 088c60f138 | my boss is bullying me...                     | bullying me                         | negative  |

# Sentiment Analysis with OHE

## Modeling

- Model trained for sentiment
- Data is One-Hot encoded



Naive Classifier: 41.0 %

```
cat_id = {'neutral': 1,  
          'negative': 0,  
          'positive': 2}
```

