**Shaista Hussain Cohort 10**

**What do you notice about the difference in the results?**

The code run by the [distilbert-base-uncased-finetuned-sst-2-english](https://huggingface.co/distilbert-base-uncased-finetuned-sst-2-english) (default model) in the pipeline yielded a EOL while scanning string literal error when i ran the command line. The bert\_sentiment = bertweet\_pipeline, gave me the following results:

0 (0.00%) of the tweets classified are positive.

100 (100.00%) of the tweets classified are neutral.

0 (0.00%) of the tweets classified are negative.

**❓ Do the results for the bertweet-base model look better, or worse, than the results for the distilbert-base model? Why?**

The results for the bertweet-base model look better than the distilbert-base model. This is likely because Tweets are short, use irregular words (like abbreviations) and have typographical errors; BERTweet is good for text classification, pretrained for English language Tweets, and possibly also handles the large data better.

**❓ How did you do? Did you find any surprising results?**

Completed, yes, the surprise was when i found 100 (100.00%) of the tweets classified as neutral.

**❓ Are there any instances where the two models gave different predictions for the same tweet?**

When i ran

for tweet in example\_difficult\_tweets[0:1000]:

pprint(sentiment\_pipeline(tweet))

print(tweet + '\n')

, [{'label': 'POSITIVE', 'score': 0.5429091453552246}]

Kong vs Godzilla has record for most meth ever consumed in a writer's room

[{'label': 'POSITIVE', 'score': 0.9998587369918823}]

Avengers is the best movie

[{'label': 'POSITIVE', 'score': 0.9998866319656372}]

Animals have amazing energy

But when i ran :

for tweet in example\_difficult\_tweets[0:1000]:

pprint(bertweet\_pipeline(tweet))

print(tweet + '\n')

Output:

[{'label': 'NEG', 'score': 0.721301257610321}]

Kong vs Godzilla has record for most meth ever consumed in a writer's room

[{'label': 'POS', 'score': 0.9917110204696655}]

Avengers is the best movie

[{'label': 'POS', 'score': 0.9904809594154358}]

Animals have amazing energy

The latter command produced different sentiment scores for the respective examples.