



Twitter **Sentiment Analysis**



Objective

In this project, I scraped 40000 tweets from Twitter about the "UCL Final" between Liverpool and Real Madrid, the tweets are categorized into Neutral, Positive, and Negative. Because limitation of time and resource, only 10000 tweet proceed.

Objective

The roBERTa-base model is used to do the Twitter sentiment analysis.

The screenshot shows the Hugging Face interface for the **roberta-base** model. At the top, the Hugging Face logo and a search bar are visible. Below the model name, there are tags for various frameworks (Fill-Mask, PyTorch, TensorFlow, JAX, Rust, Safetensors, Transformers) and datasets (bookcorpus, wikipedia, English, arxiv:1907.11692). The model card is titled "Model card" and includes a description of the RoBERTa base model, its objective, and a disclaimer. On the right, there is a section for "Hosted inference API" with a text input field for the mask token and a dropdown menu for examples.

roberta-base like 59

Fill-Mask PyTorch TensorFlow JAX Rust Safetensors Transformers bookcorpus wikipedia English arxiv:1907.11692

arxiv:1806.02847 roberta exbert AutoTrain Compatible License: mit

Model card Files and versions Community 3

RoBERTa base model

Pretrained model on English language using a masked language modeling (MLM) objective. It was introduced in [this paper](#) and first released in [this repository](#). This model is case-sensitive: it makes a difference between english and English.

Disclaimer: The team releasing RoBERTa did not write a model card for this model so this model card has been written by the Hugging Face team.

Downloads last month
10,306,773

Hosted inference API

Fill-Mask Examples

Mask token: <mask>

The goal of life is <mask>.

Dataset

40000 tweets contain "UCL Final" word

✓ f_x No	
C	D
Username	Tweet
:59:4; fednanen	@arkulari Makes me sick that some CFC fans have forgiven Courtois for winning ucl and stopping Liverpool- this cha
:59:3; thvstwt	@Louis_Tomlinson did you watch the UCL final?
:59:2; lunatik1887	@MadeiraCakes @MoBob8208 @ethanpvt @Ultra_Suristic Stop talking about the UCL, he needed to join the team th
:59:1; Hxncholini	@mackutdd Says all of this but has only seen his UCL games in which he is rarely impressive bar his lucky goals
:59:1; mxifahad3	Vini, Havertz, Coman, Origi, Salah, Bale, Benzema, Mane, Asensio, Ronaldo, Mandzukic, Casemiro.
:59:0; izzmeMaya	Apparently I read comics now
:58:4; silva_enjoyer	Scenes when Haaland bags 50+G/A next season under Pep and wins us the UCL... https://t.co/CXOEXAah66
:58:3; mo374m3d	leicester will go back to the ucl https://t.co/TC30AOD4kg
:58:3; RooneyLaBest	@ILikeBeers22 @PeterMcCormack Far from being the best) and Real Madrid, one of the best teams in the world. Ita
:58:2; 247sportscard	2018 19 Topps Chrome UCL Refractor Au #ad #thehobby https://t.co/n71pWGIOob
:58:2; yuan_basuki	@DaveOCKOP @DivockOrigi thanks for everything, Div @DivockOrigi .. Wouldn't have won UCL without you ðŹ
:58:1; Asfarniazi07	Ok so UCL final has done but I don't understand why the supporters of other clubs are making memes for liverpo
:58:0; bigbluefc	wait those canola tweets about the fans at the ucl final were real?? i just assumed someone photoshopped them ðŹ
:57:1; FkcZBelgrade	@dirtyharry230 @lildxnte @Ayoola50247910 @FabrizioRomano @hadrien_grenier On the side of who controls all th
:56:4; HalaMadrid1D	@Louis_Tomlinson did you watch the UCL final?
:56:4; Bossman_Tips	@FabrizioRomano Crazy what UCL pull does at a club like Madrid. UCL pull at Chelsea is a bit shit.
:56:1; Kingoronoa1	@presidentjoe3 @JikaMuh @JesusRole What a ucl / pl this season, enjoy your shit season https://t.co/ZODF7d0IWx
:56:0; 15ucl	@Lfcrazyyy_ @WedressiPxcked @JaimyRm own some b*tches
:56:0; ShownSolid	@ConnCFC I mean apart from the blunder of a UCL they had, i dont see why he would be unhappy.
:55:5; benktib_sofien	fun fact : Barcelona & Atletico Madrid fans cried more than Liverpool fans after the UCL final lmao.
:55:3; ichigoluvrs	@jounocult shit is so heartbreaking and like you dont see it coming at all pls ginran hurts me so bad
:55:3; Adamdz_7	@x1x_ton @NerdolaKrl @paulohe3 @richarlison97 Hello Im a Madridista, who's talking about the UCL?!!

Clean Tweet From Unwanted String

```
df.Tweet = df.Tweet.replace(r'\n', ' ', regex=True)
```

```
def preprocess(text):  
    new_text = []  
    for t in text.strip().split(" "):  
        t = '@user' if t.startswith('@') and len(t) > 1 else t  
        t = 'http' if t.startswith('http') else t  
        new_text.append(t)  
    return " ".join(new_text)
```

```
df['Tweet'] = df['Tweet'].map(lambda x: preprocess(x))
```

```
df.Tweet.head()
```

```
0    @user Makes me sick that some CFC fans have fo...  
1                @user did you watch the UCL final?  
2    @user @user @user @user Stop talking about the...  
3    @user Says all of this but has only seen his U...  
4    Vini, Havertz, Coman, Origi, Salah, Bale, Benz...  
Name: Tweet, dtype: object
```

Load Model and Tokenizer

```
roberta = f"cardiffnlp/twitter-roberta-base-sentiment"  
model = AutoModelForSequenceClassification.from_pretrained(roberta)  
tokenizer = AutoTokenizer.from_pretrained(roberta)  
model.save_pretrained(roberta)  
tokenizer.save_pretrained(roberta)  
labels = ['Negative', 'Neutral', 'Positive']
```

Sentiment Analysis

```
def sentiment(tweet):  
    encoded_tweet = tokenizer(preprocess(tweet), return_tensors='pt')  
    output = model(encoded_tweet['input_ids'], encoded_tweet['attention_mask'])  
    scores = output[0][0].detach().numpy()  
    scores = softmax(scores)  
    max_value = max(scores)  
    list_scores = list(scores)  
    index = list_scores.index(max_value)  
    # print(labels[index], max_value)  
    return pd.Series([max_value, labels[index]])
```

```
df[['Score', 'Label']] = df['Tweet'].apply(lambda x: sentiment(x))
```

Result of model

	No	Date	Username	Tweet	Score	Label
0	1	2022-05-30 23:59:42+00:00	fednanen	@user Makes me sick that some CFC fans have fo...	0.859798	Negative
1	2	2022-05-30 23:59:39+00:00	thvstwt	@user did you watch the UCL final?	0.880020	Neutral
2	3	2022-05-30 23:59:22+00:00	lunatik1887	@user @user @user @user Stop talking about the...	0.625723	Neutral
3	4	2022-05-30 23:59:18+00:00	Hxncholini	@user Says all of this but has only seen his U...	0.506189	Neutral
4	5	2022-05-30 23:59:16+00:00	mxifahad3	Vini, Havertz, Coman, Origi, Salah, Bale, Benz...	0.716288	Neutral

Result of model

