

Assignment 3 – sentiment analysis of tweets

In this assignment you will do sentiment classification on tweets that you have downloaded with the Twitter API.

Deadline

The assignment deadline is on December 5, at 2:30 pm.

Instructions

Getting the tweets

For you to get Tweets with Python, you will need access to the Twitter API. For this purpose, I have set up an educational developer account. By joining it, you will get access to the API.

- Sign up for a Twitter account (unless you already have one)
 - <https://apps.twitter.com/>
- Send me your Twitter @handle (send to this email: hturesson@schulich.yorku.ca).
- Once you have joined the educational account (MBAN6500A), create a Twitter app.
 - See [Twitter's instructions](#)
 - For more detail, see [How to make a twitter bot in Python with Tweepy](#)
- Get the API keys
- Install tweepy (a Python wrapper for the Twitter API)

You can test that your API access works with this Python snippet:

```
import tweepy

# Authenticate to Twitter
auth = tweepy.OAuthHandler("<your API key>",
    "<your API secret key>")
auth.set_access_token("<your Access token>",
    "<your Access token secret>")

api = tweepy.API(auth)

try:
    api.verify_credentials()
```

```
print("Authentication OK")
except:
    print("Error during authentication")
```

Training the sentiment analysis classifier

Download the sentiment analysis code and model from GitHub.

URL: <https://github.com/openai/generating-reviews-discovering-sentiment>

Adapt the code to fit your needs and train the LogisticRegression classifier and save it using **joblib**.

Example Python snippet:

```
from joblib import dump, load

# clf is your trained LogisticRegression Classifier
# save it:
dump(clf, 'logregress_clf.joblib')

# load it:
clf = load('logregress_clf.joblib')
```

Combining the API with the sentiment classifier

In one PY file (named according to the scheme below) combine the Twitter API query with sentiment analysis using your trained classifier. No training should occur here.

Submission

You should submit a PY file named **SA_<YourFirstnameLastname>.py**. When run (e.g. from IPython/Jupyter with `%run SA_HjalmarTuresson.py`) it should use the twitter API, download the 10 latest tweets associated with a twitter handle or hashtag of your choosing. The search tag (handle/hashtag) should be printed, and below that, one on each line, the 10 tweets along with their respective binary sentiment score (0 for negative and 1 for positive). The sentiment score should be printed before the tweet.

Example using **@united** as the search tag:

Search tag: @united

0 @united Try to use points for a flight, if it doesn't work

out you charge \$50 per ticket (\$150 this time) JUST to r...
<https://t.co/Z78GWLtyzM>
1 @lilgayafg @united @AppleCard Tbh my Costco card has some of
the best cash back
0 @ariannahuff Ridiculous! @united stop being such
cheapskates!
0 RT @jbadomics: Looking at you, @AmericanAir @united

<https://t.co/kIFfo3PwzG> <https://t.co/VckdriC0nY>
<https://t.co/qHdjSDa85G>
0 So now that @united is changing its whole status program I'm
trying to find a better card & I thought maybe...
<https://t.co/S4PTG39ByW>
0 @DeborahDefer @MeganTarnow @CathyIconis @lynda_artesani
@ChristineGalli @HSatterley @LizScottQB0 @MarnieStretch...
<https://t.co/JQHxT86Pbj>
0 Drinking a Next Coast by @GooseIsland at @united –
<https://t.co/lPzf4Pzx00>
1 Thanks United for making sure bag made tight International
connection. And congratulations Phil and Mimi, newly mar...
<https://t.co/AmpdPeZ9a9>
0 @Airdrew5 @united @KevinKead0 That is gorgeous. This livery
was gone by the time I started to notice. I wish I cou...
<https://t.co/UCkCYz3wXg>
1 @SteveDahlShow @united God bless and safe flight.

Good luck!!