

## Konsolenoutput Aufgabe 1, Übung 6 von Jessica Hassibi, 6377520

\*\*\*\*\*  
\*\*\*\*\*

### Sheet 6 Exercise 1.1:

----->Creating Top 10 list of most frequent words in wrongly classified tweets...

\*\*\*\*\*  
\*\*\*\*\*

Most frequent words in wrongly classified tweets:

- No. 1 : mal with 97 occurrences.
- No. 2 : hab with 49 occurrences.
- No. 3 : einfach with 45 occurrences.
- No. 4 : Menschen with 37 occurrences.
- No. 5 : Guten with 29 occurrences.
- No. 6 : halt with 27 occurrences.
- No. 7 : is with 26 occurrences.
- No. 8 : via with 24 occurrences.
- No. 9 : Kinder with 24 occurrences.
- No. 10 : Danke with 23 occurrences.

Top 10 list of most frequent words in wrongly classified tweets:

[('mal', 97), ('hab', 49), ('einfach', 45), ('Menschen', 37), ('Guten', 29), ('halt', 27), ('is', 26), ('via', 24), ('Kinder', 24), ('Danke', 23)]

\*\*\*\*\*  
\*\*\*\*\*

### Sheet 6 Exercise 1.2:

----->Creating Top 10 list of most frequent wrongly classified hashtags..

\*\*\*\*\*  
\*\*\*\*\*

Most frequent wrongly classified hashtags in tweet list and number of incorrect occasions:

- No. 1 : #Corona with 9 occurrences.
- No. 2 : #ibes with 9 occurrences.
- No. 3 : #twitch with 5 occurrences.
- No. 4 : #coronavirus with 4 occurrences.
- No. 5 : #twlz with 3 occurrences.
- No. 6 : #B117 with 3 occurrences.
- No. 7 : #Wetter with 3 occurrences.
- No. 8 : #covid19 with 2 occurrences.
- No. 9 : #nowplaying with 2 occurrences.
- No. 10 : #w3101 with 2 occurrences.

-> Tweets mit dem Hashtag #Corona werden mit 82 Prozent Wahrscheinlichkeit falsch klassifiziert.

- > Tweets mit dem Hashtag #ibes werden mit 100 Prozent Wahrscheinlichkeit falsch klassifiziert.
- > Tweets mit dem Hashtag #twitch werden mit 100 Prozent Wahrscheinlichkeit falsch klassifiziert.
- > Tweets mit dem Hashtag #coronavirus werden mit 100 Prozent Wahrscheinlichkeit falsch klassifiziert.
- > Tweets mit dem Hashtag #twlz werden mit 100 Prozent Wahrscheinlichkeit falsch klassifiziert.
- > Tweets mit dem Hashtag #B117 werden mit 100 Prozent Wahrscheinlichkeit falsch klassifiziert.
- > Tweets mit dem Hashtag #Wetter werden mit 100 Prozent Wahrscheinlichkeit falsch klassifiziert.

List of most often wrongly classified hashtags in tweets list:

Hashtag name, total occurrences, incorrect occurrences, probability of wrong classification

```
[('#Corona', 11, 9, 0.8181818181818182), ('#ibes', 9, 9, 1.0), ('#twitch', 5, 5, 1.0),  
('#coronavirus', 4, 4, 1.0), ('#twlz', 3, 3, 1.0), ('#B117', 3, 3, 1.0), ('#Wetter', 3, 3, 1.0)]
```