AI & Automation Exam 2

05/06/2022

Mads Hegewald (EAAMHEG)

# Exercise 1

The first example of wordcloud is where I have used a file called “example.txt” in the directory “exercise1”, which is the transcript of the speech by Leonardo DiCaprio at the 2014 UN Climate Summit. First I read the txt file where I used .split() to get each individual word, followed by removing the comma “,” or dot “.” at the end of some of the words. After I added the words to a dictionary to keep track of the frequency of the words, but I did not want to add words such as “the”, “as”, “for” etc. So, I filtered these words out by using checking if the word was in the STOPWORDS from wordcloud. This is all done using the following code:

Text

Description automatically generated

Running the code shows the following wordcloud. One thing I noticed was how I wanted it to say “climate change” in one word but after some experimenting, I could not find a solid solution to that problem and stuck with splitting the words when a space between them occurred.

Text

Description automatically generated

For my second example I tried using a webtext from nltk.corpus called “pirate.txt” located at the same place as the example with “firefox.txt”. It seems like the text file contains a script of some sorts of the Pirates of the Caribbean movie. First we load in the words from webtext and use the nltk.FreqDist to create a dictionary with the frequencies of each word, this still contains all the stopwords mentioned previously. To sort these out I used a line from the example we had during one of the lessons, where we basically sort out all the words that are smaller than 3. This is done in the following code:

Text

Description automatically generated

After this we can simply just create a wordcloud object, where we generate the words from frequencies in the list “filter\_words” and we get the following wordcloud.

Text

Description automatically generated

# Exercise 2

## a)

## b)

## c)

# Exercise 3

## a)

## b)

## c)

# Exercise 4

## a)

## b)

## c)

## d)

## e)

# Exercise 5

## a)

## b)

## c)

## d)