

Daniel Deloughry

-

R00115920

-

Assignment 2 Report

-

Data Driven Microservices

-

BSc (Honours) in Software Development

-

Cork Institute of Technology

For the first assignment of the module I created an application for analysing tweets containing a word to search for. I did this using three microservices. One for streaming the tweets, one for analysing each tweet's sentiment and lastly, one for printing the result to the browser via port 3000 on localhost. *Docker* compose was then used to run all the microservices together.

One of the aims of the second assignment was to edit the sentiment analysis so as a second data source could be easily added.

One of the first changes I made was using an *AWS* hosted *MYSQL* database instead of *MongoDB* and *RabbitMQ*. I used an online database as I had had considerable difficulty in containerising *MongoDB* and *RabbitMQ*, and it made deployment faster. My reason for choosing *MYSQL* over a NoSQL database was I found storing the sentences, source names and sentiment scores easier. Also, the *MYSQL* average (*AVG()*) function made finding average sentiments easier. I then edited the microservices in accordance with them using the database.

The data source I chose was the *BBC RSS* feed. For this I created a new microservice that uses *Feedparser* to stream the news headlines from the *BBC RSS feed* and save them to the database.

In the sentiment analysis microservice I edited it, so that it could analyse the sentiment for any sentence regardless of the source. In the database table contains a row with the sentence, data source name (e.g. *Twitter*, *BBC RSS feed*, etc), the time, and the sentiment score for that sentence. Using *TextBlob*, it simply calculates the sentiment polarity for every row where the sentiment score is null. If the score isn't null this means it is an old row and therefore its score has already been calculated. This service can now analyse sentences regardless of their source and it won't mix them up.

Calculating the average is done in the display microservice. It selects from the database using the *AVG()* function to calculate the average sentiment score. To differentiate between the sources, it simply selects based on the source name column. The average sentiment analysis score for each source can then be viewed in the browser.