EXAMENSARBETE Applications of Cloud-Based Cognitive Computing, Identifying Frequent Complaints from Social Media
STUDENTER Axel Ulmestig
HANDLEDARE Pierre Nugues (LTH)
EXAMINATOR Björn Regnell

Using Watson to decipher social media

POPULÄRVETENSKAPLIG SAMMANFATTNING **Axel Ulmestig**

In order to get insight into the possibilities of using the AI system Watson I have constructed an application that reads Twitter and tries to extract common complaints. This turned out to be rather difficult but gave some interesting insights into the workings of both Watson and social media.

Watson is an artificial intelligence (AI) system originally developed by IBM in order to compete in the American game show Jeopardy!. Since then, Watson has been made available for developers by allowing them to access it through the web. Since Watson is constructed for natural language processing, understanding language in the way that humans do, I wanted to do a project that utilizes that as much as possible. Additionally it would be more interesting to create something that is potentially commercially viable. Such a project would potentially be a good indicator of what direction this technology is going. With those two things in mind I decided to try to extract information from the social media platform Twitter. The things in social media that hold commercial value are mostly opinions regarding business and organizations. Both positive and negative opinions are interesting, but the negative ones are usually more urgent. This lead me to focus on the negative opinions and try to find the problems that were frequently brought up by the users. Twitter has also been extensively used for similar projects making it easy to compare this project with the accomplishments of other articles.

After the application was created it was compared against the performance of human test subjects who were instructed to extract complaints from a set of tweets. The results of the test were rather disappointing. The application and the test subjects had no overlap in what they perceived as expressed complaints. There are probably many reasons for why the results were as bad as they were. The algorithm that tries to identify complaints does essentially look for words that appear in a negative context, that is in the context of words that are deemed negative such as “bad”. The application also does not take such things as negations into account. These things makes the analysis rather blunt and makes it dependent on large quantities of data to function. This was not the case with the human test that only used 75 tweets that were a sample of real tweets.

However, Some of these problems also remained when running on much larger quantities of data. This came as a bit of a surprised but can be traced to the nature of Twitter. Popular tweets get “retweeted“, they get repeated with the exact same message worded in the same way. If such a message is picked up by the system and deemed as negative then it will have a big effect on the overall problem extraction.

Watson has, despite the poor results, been very helpful in this endeavour. This type of technology can enable other, potentially more successful, ideas to be relatively quickly implemented and tested. The fact that a lot of tools are easily available will most likely lead to an increased use of these types of services and applications.