DS L1T20 – CAPSTONE PROJECT III

Raman Rahjiv Sewjugath

Compulsory Task 1

A. Text Classification
B. Automatic summarisation
C. Question Answering

Compulsory Task 2

I work in a research institute and one of the projects we work on is to extract oil from algae to process into biofuel. Now the amount of data we have dates back from 2010. We have multiple ponds (7 - 10) at 3000L each, and to simply put it, we document the ponds health in terms of physiochemical and morphological changes during every run (and these are run multiple times a year). So basically next to our readings we not down various changes such as presence of worms, colouration, Extracellular polymeric substances (usually occurs when the pond is very old), when chemical nutrients have been added and at what time periods, when the pond is about to die etc.

So these keywords over the years become lost, due to the fact that they are too much to work with because there is no way of extracting the information from such huge amounts of data.

We can use NLP to flag certain keywords through the database, finding out at what time periods the water colour changes to certain hues and we can relate that to pond stress, nutrient stress, or pond crash. The relationship between what chemicals initiate certain physical responses.

We also have a few species of microalgae growing in the same ponds, so we could also use NLP to identify when a certain undesirable species starts to take over, giving scientists ample time to implement measures to prevent growth of the specified species. This is very important because some species produce very little oil which is not feasible and it's crucial to know.

So in just, we can use NLP to monitor pond health, pond crash and relate to reasons for the crash, monitor species growth and flag growth of undesirable species, and a lot more. I am fairly new to NLP and I'm sure with more experience will come more ideas for its use.

Compulsory Task 3

Grammarly uses NLP along with artificial intelligence and deep learning to provide intuitive suggestions, corrections, synonyms etc. thereby improving the overall writing calibre.

The author explains that NLP is used to teach machines how to understand the human language in order to be useful. This is done by feeding high-quality training data which human researchers have organized and labelled in a way that machines can understand. They further explain that they teach AI the correct ways of using the language (punctuation, words,

grammar etc.) in a sentence as well as the incorrect way so that it knows how to correct the mistake.

To further facilitate NLP and machine learning, adjustments are made to the algorithm when users hit “ignore” when a sentence is flagged for correction by the AI. In this way it becomes more accurate and efficient in its purpose.

Reference: https://www.grammarly.com/blog/how-grammarly-uses-ai/