**Project Proposal**

**Candidate Number** – 164547

**Supervisor** – Dr Julie Weeds

**Working Title** – Fact-Checking Journalism

**Aims and Objectives**

The purpose of my project is to provide a tool to journalists which can be used to fact check aspects of their article/document against an external database.

Since ‘fake news’ is such a prominent topic currently, I’m motivated to create a tool which can help journalists mitigate false information in their article before being published. As an extension, I’d also like to provide this software to readers, so they can fact check articles themselves. However, time constraints will have to be taken into consideration.

It should be noted that my focus will be on providing a tool which journalists can manually use to check their facts as opposed to creating an automated service. I intend to do this by providing some sort of link to the external database concerning the fact in question. Automated fact-checking is an extension of my project.

Rather than scanning for all facts within a document - which may prove difficult and inaccurate – I’ll focus my tool on fact-checking a specific area of interest about named entities. This means searching for any instances where a person is stated to hold a job within an organisation. My tool would then provide a link to the external database displaying what role that person currently has.

**Relevance**

My project is relevant to my degree course as it will display a variety of skills that I’ve learnt so far and will hope to improve on. The focus is to incorporate Natural Language Processing (NLP) and Machine Learning (ML) techniques to perform the fact-checking scan. Additionally, I’ll be making use of the skills I’ve learnt in Python to code the back end and Java to code the front end. Since I’ve never used PyCharm before, that’ll be a useful skill to add to my repertoire along the way.

**Resources Required**

I’ll require a database to check my named entity facts against. Options that I’m looking at include DBpedia and Wikipedia. As for an API that I’m looking for to provide me with news articles to train and test my model with, options include LexisNexis, the Guardian API and Google’s API.

The software environment I’ll be using to code my backend is PyCharm, while I’ll most likely use the NetBeans IDE to code my front end.

**Weekly Timetable**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Mon | Tue | Wed | Thurs | Fri | Sat |
| 9:00-10:00 |  |  |  |  |  |  |
| 10:00-11:00 |  |  | Self-Study |  |  |  |
| 11:00-12:00 |  |  | Self-Study |  |  | Self-Study |
| 12:00-13:00 |  |  | Self-Study |  | Lecture | Self-Study |
| 13:00-14:00 |  |  |  |  |  |  |
| 14:00-15:00 |  |  |  | Self-Study |  |  |
| 15:00-16:00 |  | Supervisor Meeting |  | Self-Study |  |  |
| 16:00-17:00 | Self-Study |  |  | Self-Study |  |  |
| 17:00-18:00 | Self-Study |  |  |  |  |  |

The above timetable represents how much time I plan on devoting to my project per week. The timetable isn’t set in stone (for example, my supervisor meeting is subject to change from week to week), however I do plan on spending a minimum of 10 hours a week on self-study.

**Background Reading Bibliography**

|  |  |
| --- | --- |
| [1] | J. Thorne, “Automated Fact Checking: Task formulations, methods and future directions,” 20 August 2018. [Online]. Available: http://aclweb.org/anthology/C18-1283. [Accessed 11 October 2018]. |
| [2] | L. Graves, “Understanding the Promise and Limits of Automated Fact Checking,” February 2018. [Online]. Available: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/graves\_factsheet\_180226%20FINAL.pdf. [Accessed 04 October 2018]. |
| [3] | N. Hassan, “Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster,” 2017. [Online]. Available: http://ranger.uta.edu/~cli/pubs/2017/claimbuster-kdd17-hassan.pdf. [Accessed 07 October 2018]. |

**Interim Log of Supervisor Meetings**

|  |  |
| --- | --- |
| Date | Actions |
| 26/09/18 | Initial discussion about what my project would entail & first steps. |
| 02/10/18 | 10-minute Q&A about my project with supervisor + other students. |
| 09/10/18 | Update on background research & discussion on the software I’ll be using. |