**PROJECT SPECIFICATION - Project (SEGM) 2017/18**

|  |  |
| --- | --- |
| **Student:** | **Ben Ewen** |
| **Date:** | **20/10/17** |
| **Supervisor:** | **Jing Wang** |
| **Degree Course:** | **Computer Science** |
| **Title of Project:** | **Using natural language processing and deep learning to predict sentiment in online reviews** |

#### Elaboration

|  |
| --- |
| Natural language processing (NLP) aims to allow computers to process human language in a meaningful way such as responding to questions, converting data into sentences that can be understood by humans, or sentiment analysis. Additionally, as deep learning algorithms have matured, the ability to have a computer analyse human language has become more autonomous and relies less on supervision by a human.  Sentiment analysis, in essence, is trying to understand the sentiment behind a sentence. For example the sentence: “We went to the cinema, it was affordable but the movie was rubbish!” can be easily interpreted by a human. We know that the price was good, but the movie wasn’t great. This is valuable information to the cinema, and sentiment analysis would be able to automate the analysis of these reviews.  This project will look at the various ways sentiment analysis can be performed via deep learning algorithms. |

#### Project Aims

|  |
| --- |
| * Learn natural language processing concepts in relation to sentiment analysis * Learn deep learning concepts and how these can be used for sentiment analysis * Create a sentiment analysis tool to categorise online reviews * Research and evaluate the performance of existing solutions for sentiment analysis * Evaluate my own sentiment analysis tool by comparing its performance with existing solutions |

#### Project deliverable(s)

|  |
| --- |
| I will develop a deep learning algorithm that will analyse large data sets of human language and interpret the sentiment behind them. I will devise an evaluation method that will compare the performance of my own algorithm compared to existing ones. This algorithm will then be used to create a prototype sentiment analysis tool that can be used by online review businesses to categorise their reviews by sentiment. |

#### Action plan

|  |  |
| --- | --- |
| **Task** | **Deadline** |
| Find a Project Supervisor | 06/10/17 |
| Complete Project Specification and Ethics Form | 20/10/17 |
| Background Research   * Natural Language Processing   + Word Embeddings     - Word2Vec * Deep Learning   + Logistic Regression * Types of Neural Networks   + Convolutional Neural Networks   + Recursive Neural Networks   + Recurrent Neural Networks & LSTM * Existing products for sentiment analysis * Traditional NLP algorithms | 10/11/17 (Overall) |
| Information/Literature Review | 01/12/17 |
| Produce Deliverable   * Experimental Development * Development of Learning algorithm * Development of front-end system | 14/12/17  01/02/18  01/03/18 |
| Provisional Contents Page | 09/02/18 |
| Draft Critical Evaluation | 09/03/18 |
| Sections of a Draft Report | 09/03/18 |
| Evaluation   * Accuracy * Personal Reflection * Future work | 14/03/18  01/04/18  01/04/18 |
| Submit Body of Project to TurnItIn | 11/04/18 |
| Submit Project | 12/04/18 |

#### Ethics

|  |
| --- |
| Based on the discussion with my project supervisor, this project has a low risk in terms of unethical practices. I will be analysing openly available data – none of the data used will be personally identifiable. |