C:/> ./detecting_political_bias_in_text.bat

~ Project_Student = "Richard Jones - 17011180"; Project_Supervisor = "Dr. Charles Day"

> mkdir ./Introduction

- In recent years, the atmosphere of UK politics has become a hotbed of unrest. Many have begun to challenge news reports and statistics as misleading or simply untrue. From this, more and more people are starting to realise they are not as good at detecting political bias as they once believed.
- This project aims to develop a system that uses deep learning techniques to detect political bias in text, and provide an interface for a member of the general public to analyse their own text for this sentiment.
- The aim of this project is to provide a platform where users can receive a generalised, unbiased opinion on the skew of text. By using a trained Artificial Intelligence, it is believed that this risk of a biased opinion can be reduced.

> cd ../Objectives

- Develop a deep learning system, trained on speeches from the UK parliament, that is capable of identifying whether a piece of variable-length text is more left-leaning or right-leaning ideologically.
- Provide an Interface for a user to interact with the system, allowing them to upload and analyse their own texts.
- Design and Build a Database to store the relative inputs and outputs of the system for the user to query separately.

> dir ./Project_Methodology

- This project will incorporate an AGILE methodology using the SCRUM framework [3]. Meaning all of the components will be completed iteratively.
- A backlog of the tasks that need to be completed for the projects success has been created, with necessary deadlines set for each of them.
- *The Gantt Chart* (Fig. 1) *outlines the content and deadlines of this product backlog.*

Section	Week																									
	Semester 1												Christmas					Semester 2								
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	1	2	3	4	5	6	7	8	9
Project Plan ***(Complete)***	■	■	■	■																						
Literature Review ***(Complete)***		■	■	■	■																					
Analysis and Design **(Complete)**				■	■	■	■	■																		
Neural Network Design **(Complete)**						■	■	■	■	■	■	■	■	■	■	■										
NN Training/Validating **(Complete)**								■	■		■	■		■	■											
NN Testing **(Complete)**										■			■			■										
Database **(Complete)**																	■	■	■							
UI Development																				■	■					
Final Testing																						■	■	■		
Evaluation																							■	■		
Project Report																								■	■	■
Documentation	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Current Week

Fig 1: Gantt Chart

> vim Project_Design.txt

> python ./NeuralNetwork.py

- As this project will be focussed around analysing the political sentiment of text, using a standard Multi-Layer Perceptron model for the Neural Network will not be viable. Instead, an **LSTM (Long Short-Term Memory)**[1] Model will be utilised *(pictured below)*.
- Every LSTM cell contains an input gate, an output gate and a forget gate. Which allows relevant or important information and characteristics of the text to be remembered, and irrelevant content to be forgotten. This has been proven to be effective in analysing sentiment across several scenarios

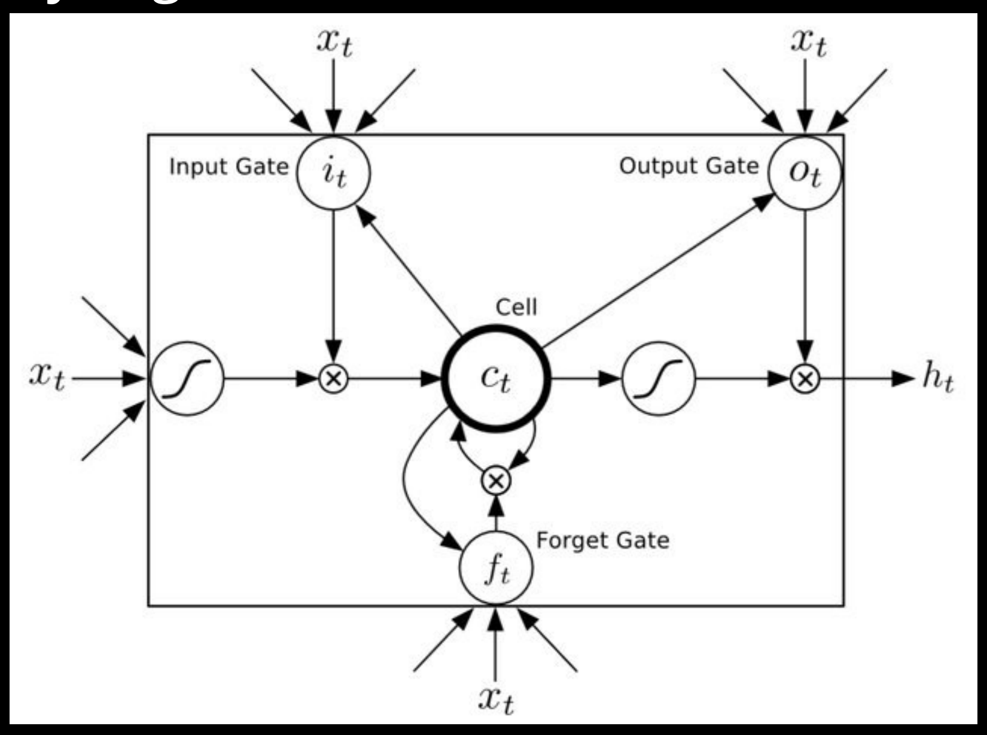
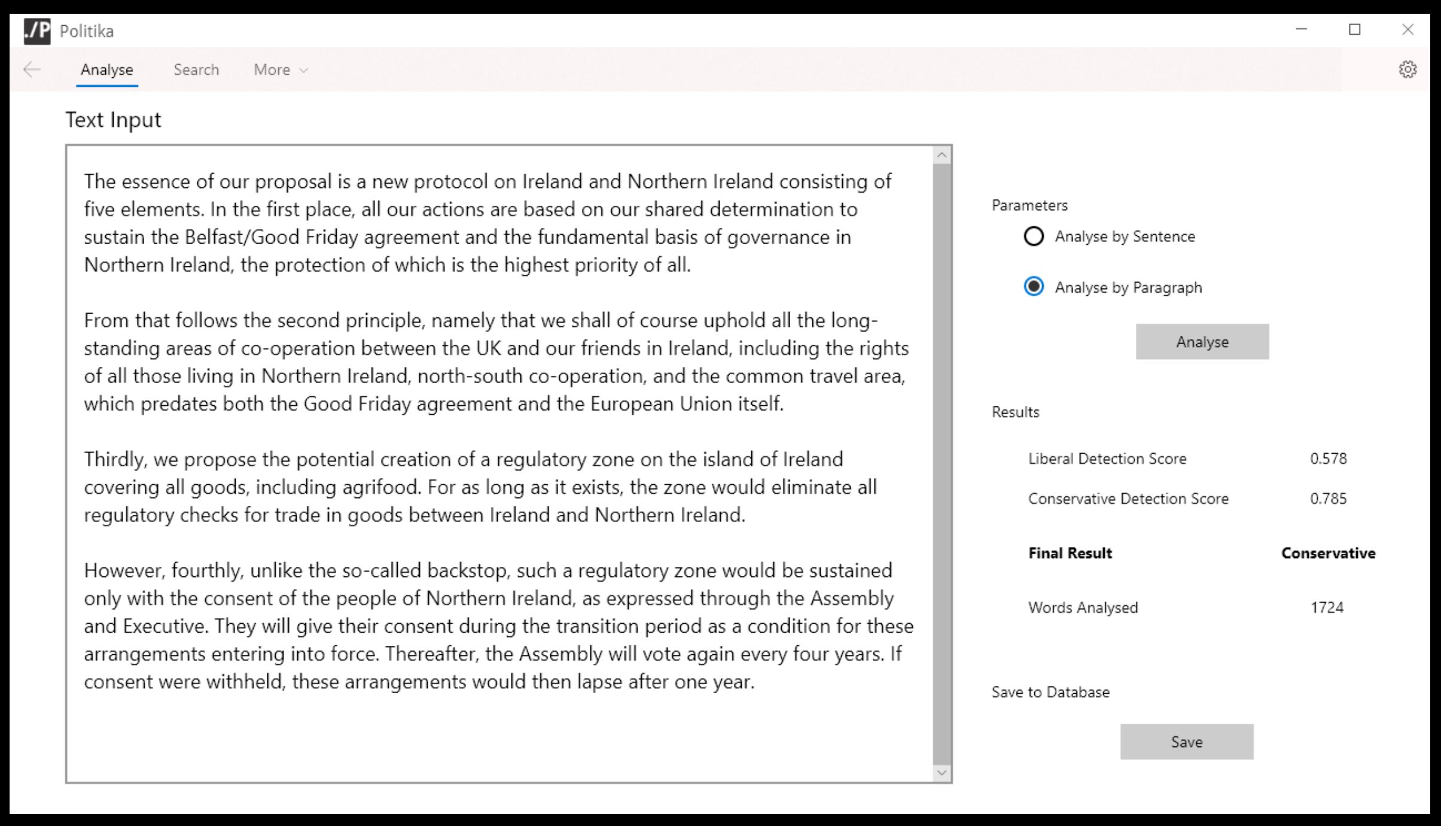


Fig 2: A diagram of an LSTM Cell

> The Neural Network has been developed in python using Google's TensorFlow Library.

> javac ./UserInterface.java

- The UI will be programmed in Java using the Swing GUI Library.
- It will provide a interface for the user to analyse their own texts, as well as query previously analysed texts through a search function.
- It will also be responsible for passing information between the relevant python scrips that will analyse all inputted text.
- A wireframe *(Fig. 3)* was created visualise the expected design of the UI.



Text Input

The essence of our proposal is a new protocol on Ireland and Northern Ireland consisting of five elements. In the first place, all our actions are based on our shared determination to sustain the Belfast/Good Friday agreement and the fundamental basis of governance in Northern Ireland, the protection of which is the highest priority of all.

From that follows the second principle, namely that we shall of course uphold all the long-standing areas of co-operation between the UK and our friends in Ireland, including the rights of all those living in Northern Ireland, north-south co-operation, and the common travel area, which predates both the Good Friday agreement and the European Union itself.

Thirdly, we propose the potential creation of a regulatory zone on the island of Ireland covering all goods, including agrifood. For as long as it exists, the zone would eliminate all regulatory checks for trade in goods between Ireland and Northern Ireland.

However, fourthly, unlike the so-called backstop, such a regulatory zone would be sustained only with the consent of the people of Northern Ireland, as expressed through the Assembly and Executive. They will give their consent during the transition period as a condition for these arrangements entering into force. Thereafter, the Assembly will vote again every four years. If consent were withheld, these arrangements would then lapse after one year.

Parameters

○ Analyse by Sentence

● Analyse by Paragraph

Analyse

Results

Liberal Detection Score 0.578

Conservative Detection Score 0.785

Final Result **Conservative**

Words Analysed 1724

Save to Database

Save

Fig 3: User-Interface Wireframe for the analysis section of the program

SQL> CREATE DATABASE articleDB

- The Article Database will be developed using the SQLite Database Management system following the schema of this Logical UML Diagram *(Fig. 4)*.

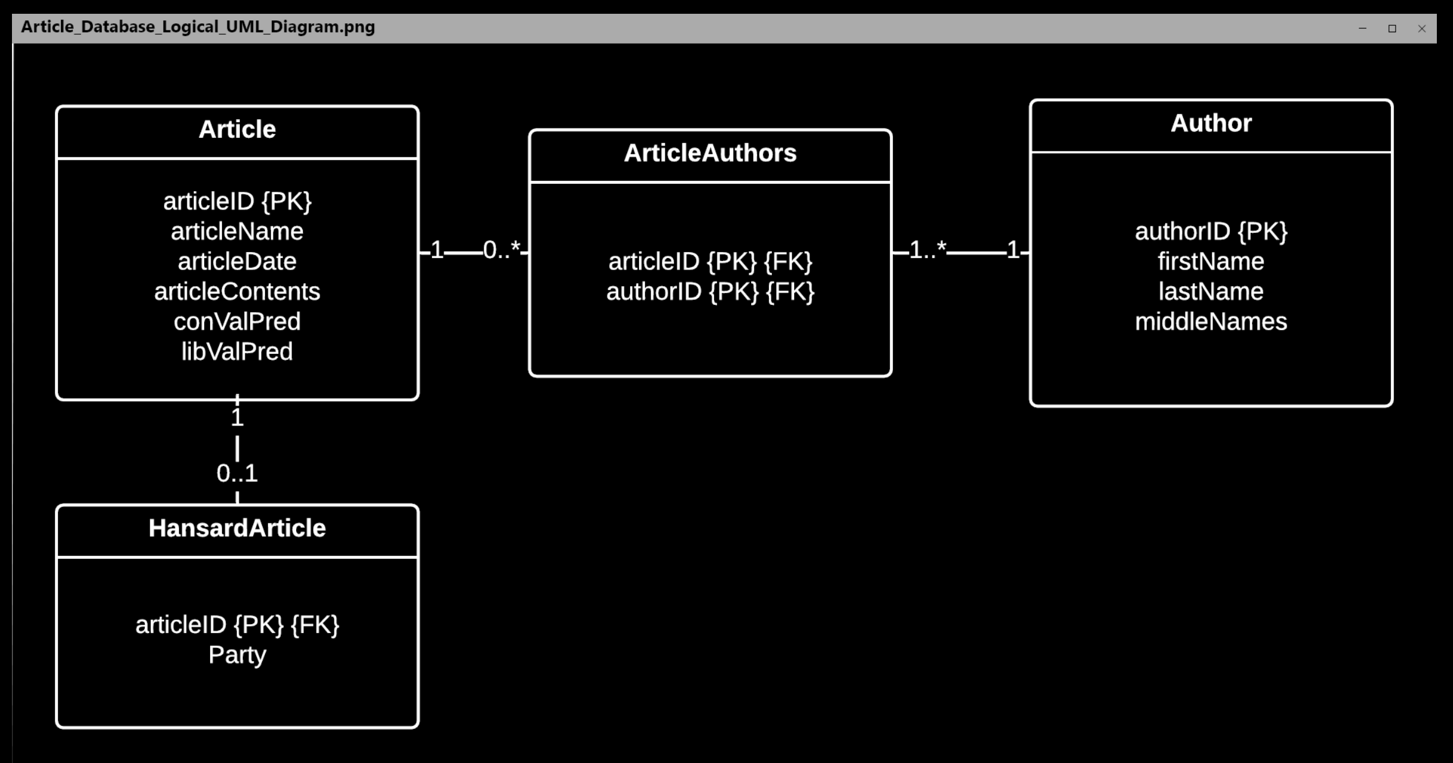


Fig 4: Logical UML Diagram for the Article Database

> ./Development.exe

> head ./Current_Progress

- Retrieved and Processed Training, Validation and Testing Datasets from the Hansard Parliamentary Debates Archives.
- Created a TensorFlow Input Pipeline that: Retrieves the collected debate data, loads it into a TensorFlow dataset, and builds a text-encoder vocabulary file for use in the training of the Neural Network.
- Developed and have begun training/optimising the Neural Network.
- Coded the analysis section of the User-Interface and developed a method to communicate with the Neural Network.
- Created and populated the Article Database based on the designed UML Schema.

> echo Neural_Network_Performance

- Currently, attempts at training the Neural Network have not been as optimal as anticipated.
- As Fig. 5 shows, though there is a slow increase in training accuracy, the validation accuracy is constantly shifting or "exploding".
- The cause of this is possible due to too large of an initial LSTM layer size, which could've immediately over-trained the Neural Network, leading to the sub-prime results here.
- Currently, work is being carried out using much smaller LSTM-Layer sizes to see if performance is improved.

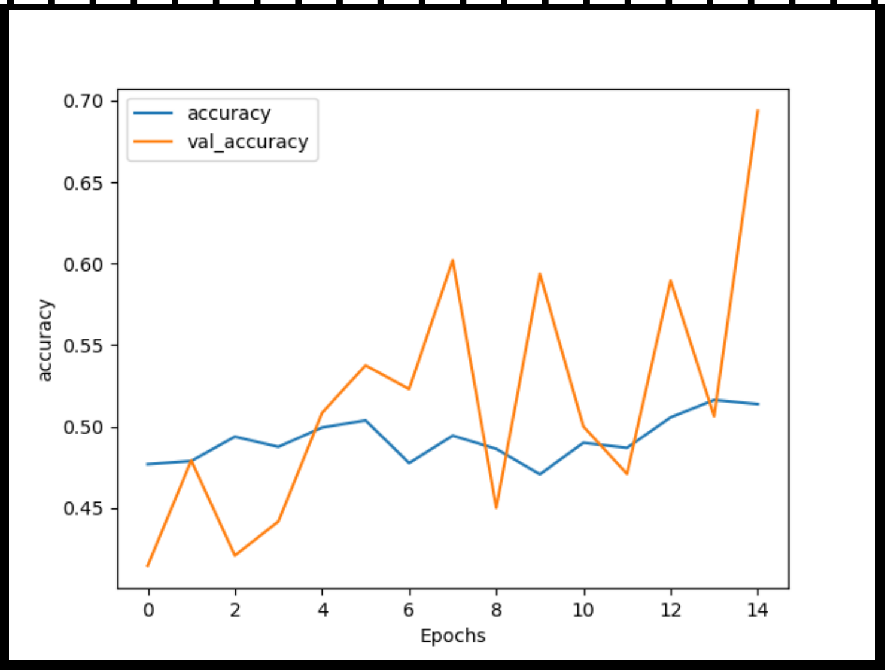


Fig. 5

> TaskKill ./Remaining_Tasks

- Develop the search functionality between the User -Interface and the Article Database.
- Finish optimising the Neural Network to a sufficient standard.
- Test the Program's functionality and usability through feedback sessions.
- Finalise Project in a Report.

> git blame ./References : [1]Hochreiter, S. and Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), pp.1735-1780. ; [2]Commons.wikimedia.org. (2020). File:Peephole Long Short-Term Memory.svg - Wikimedia Commons. [online] Available at: https://commons.wikimedia.org/wiki/File:Peephole_Long_Short-Term_Memory.svg [Accessed 16 Feb. 2020]. ; [3]Scrum.org. (2020). What is Scrum?. [online] Available at: https://www.scrum.org/resources/what-is-scrum [Accessed 16 Feb. 2020].