Project Plan

UCD School of Computer Science

University College Dublin

June 12, 2023

Authors:
Kevin Singpurwala 16359146,
Jianhao Feng 22205710,
Natasha Suman 22200100,
Anju Krishna 22200540,
AhmedFaizel Asharafali 22204855,
Madhav Mohan 22204214

Contents

1 Project Objectives

By using sentiment analysis techniques on articles, the fake news detection system seeks to recognize and categorize content as either genuine or misleading. By comparing them with well-known labeled datasets or professional assessments, the system will assess the accuracy of its classifications. In the current information era, when preventing the spread of misleading information and fake news is a significant challenge, the goal of the fake news detection system is especially relevant. The system's goal is to complete categorization tasks reliably and accurately within a specific period of time, such as by meeting a predefined accuracy threshold. The underlying assumption is that fake news articles often employ emotional cues that differ from genuine news [1]. Sometimes news can be real yet still have extreme emotion in it. This can be used to identify behavioural patterns and emotional responses [2] [3].

2 Project Plan

Epics, User stories and tasks to realise those user stories have been created on Jira.

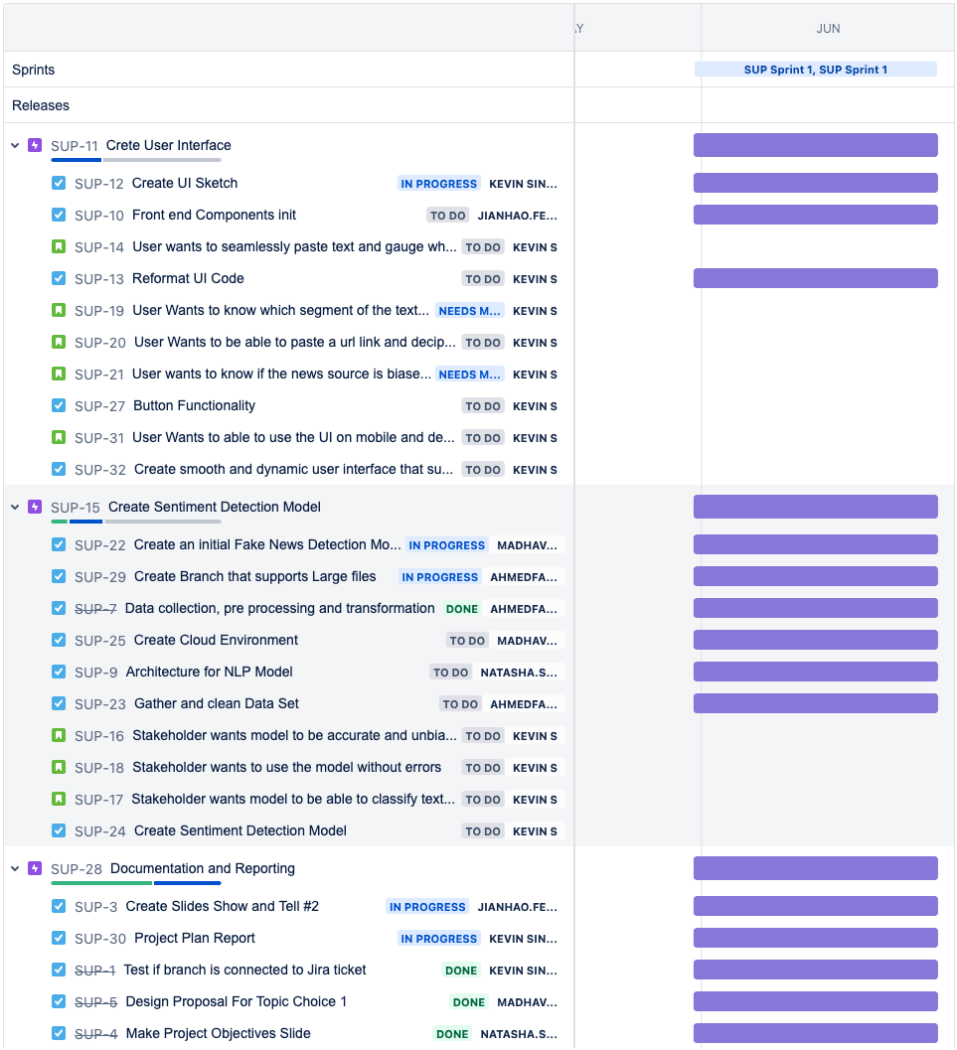

Figure 1: Sprint 1 Road Map

3 Roles

Table 1: Team Members and Roles

Member	Role
Kevin Singpurwala	Front-End
Jianhao Feng	Front-End
Madhav Mohan	Back-End
AhmedFaizel Asharafali	Back-End
Anju Krishna	Tester
Natasha Suman	Data Analyst/ML Engineer

The creation of a seamless user interface for our system using generative AI and large language models will be greatly aided by the front-end developers. They will guarantee smooth user-generated content engagement. The ML engineers will focus on creating complex models for producing AI-driven content. They will rigorously evaluate each model's performance once they have been trained and tested extensively in order to improve the accuracy and coherence of the output. Our data analyst will be accountable for gathering and preparing pertinent data from dependable sources. They will guarantee that the models have access to reliable and high-quality data for training and fine-tuning. Our devoted tester will carefully evaluate the system's usability, reliability, and performance to uphold the highest standards [3].

4 Architecture

Software Project Architecture Our front-end will be React based. Django and python will allow out database, model and back-end to interact with the front-end. [4] It is still possible that we make use of another framework in place of Django, however the general architecture remains unchanged.

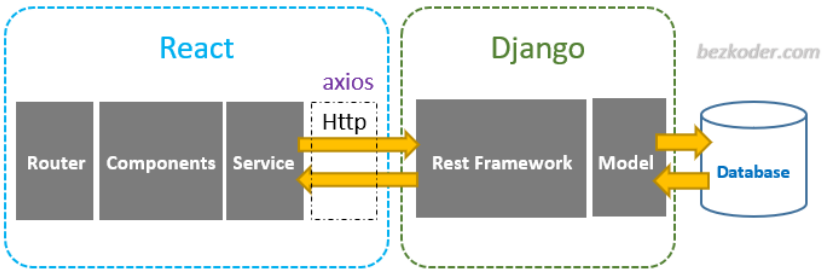

Figure 2: Our Software Architecture

Cloud Environment We will avail of Azure's cloud environment to tune our model after our minimal viable product(MVP) is created. We are in the process of obtaining a licence for Azure, from Microsoft. Azure will allow us to deploy our model to the cloud.

5 Data Plan

We will make use of data from multiple sources. SemEval , social media sites such as Twitter, our own generated data, and Kaggle.

There are numerous challenges, from limiting bias in the training data set to labelling an adequate amount of "fake" and "genuine" text data. Deciphering sentiment from training and test data is then another hurdle to overcome. The reliability and integrity of the chosen data sets is of utmost importance. The data set is the first step of building our model. Without the right data set, we can not create a competent model.

6 Github

Our team makes regular commits on Github. These are reflected in the Jira tickets and our shared google drive.

Table 2: Team Members and Github Commit Evidence

Member	**Commit hash(es)**
Kevin Singpurwala	c3e91b07d8964d6bee534a8f1c0ebf8ed2290a2e
Jianhao Feng	448aa714f54ad62dc26fc459dd3093a5de1fc365
Madhav Mohan	1a0332d2bdefc3287718261d133bcc8f967939c8
AhmedFaizel Asharafali	87b5a3f05f9b2dd4cbcc34cec9a15b66783016e6
Anju Krishna	b2c24b92e18aaa2a4167de150d421c3861cb40de
Natasha Suman	af081466fd13fdd6e4e3c0b2c398e6ec9b83421e

7 Team Management

As seen in section 2, we are availing of Jira for Project Team Management. Sprints have been set up on a 2 week bases. It is decided at the start of every sprint in a session called "Sprint Planning" what the objectives of said sprint are. Then task prioritisation occurs so that we can take on the most important core tasks first to aid in the development of our MVP. We meet with our Microsoft mentors every 2 weeks. This allows us to update them on our progress and improve our system based on their guidance and feedback.

The minutes of our meetings are uploaded to both Github and our google drive. All tasks are created and assigned using Jira. This is depicted in Figure 3. Code reviews ensures code quality and facilitates communication between team members.

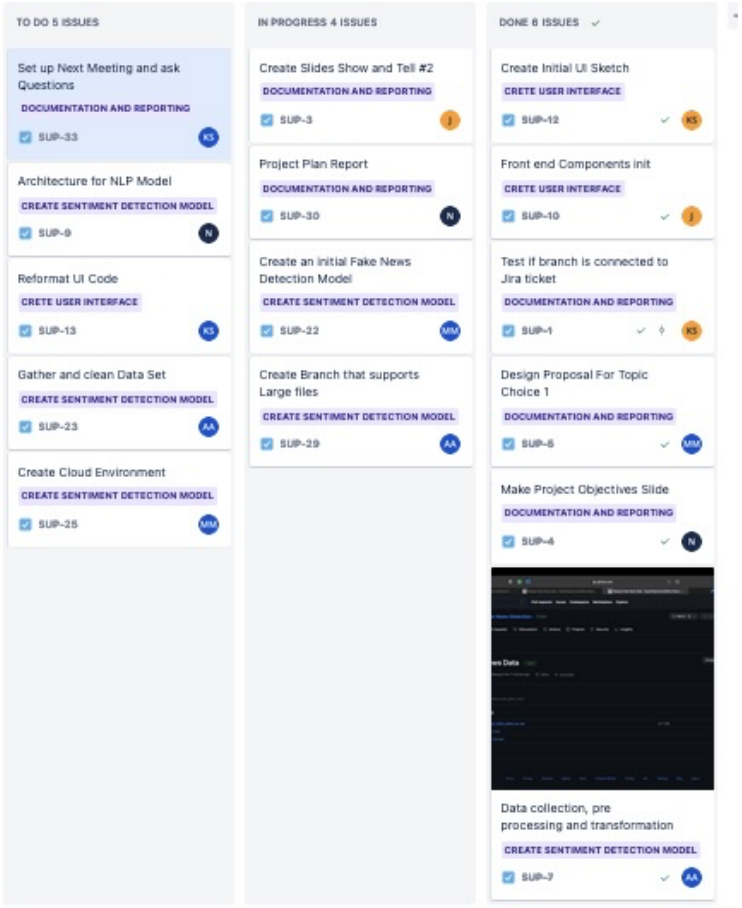

Figure 3: Jira Board

References

[1] M. Fernández-López and M. Perea, "Language does not modulate fake news credibility, but emotion does," *Psicológica Journal*, vol. 41, no. 2, pp. 84–102, 2020.

[2] J. Dupuis, "Who do you say i am?" *Theological Studies*, vol. 55, no. 4, p. 782, 1994.

[3] K. Singpurwala, J. Feng, N. Suman, A. Krishna, A. Asharafali, and M. Mohan, *Microsoft pitch document*, Team: Super Saiyans, Jun. 2023.

[4] Django Software Foundation, *Django*, version 2.2, May 5, 2019. [Online]. Available: https://djangoproject.com.