Functional Specification

Students:

Brendan Mannion (18395521)

Alan Hian Wu (18332283)

Date of Completion: 19 November 2021

Table of contents

1. Introduction

1.1 Overview

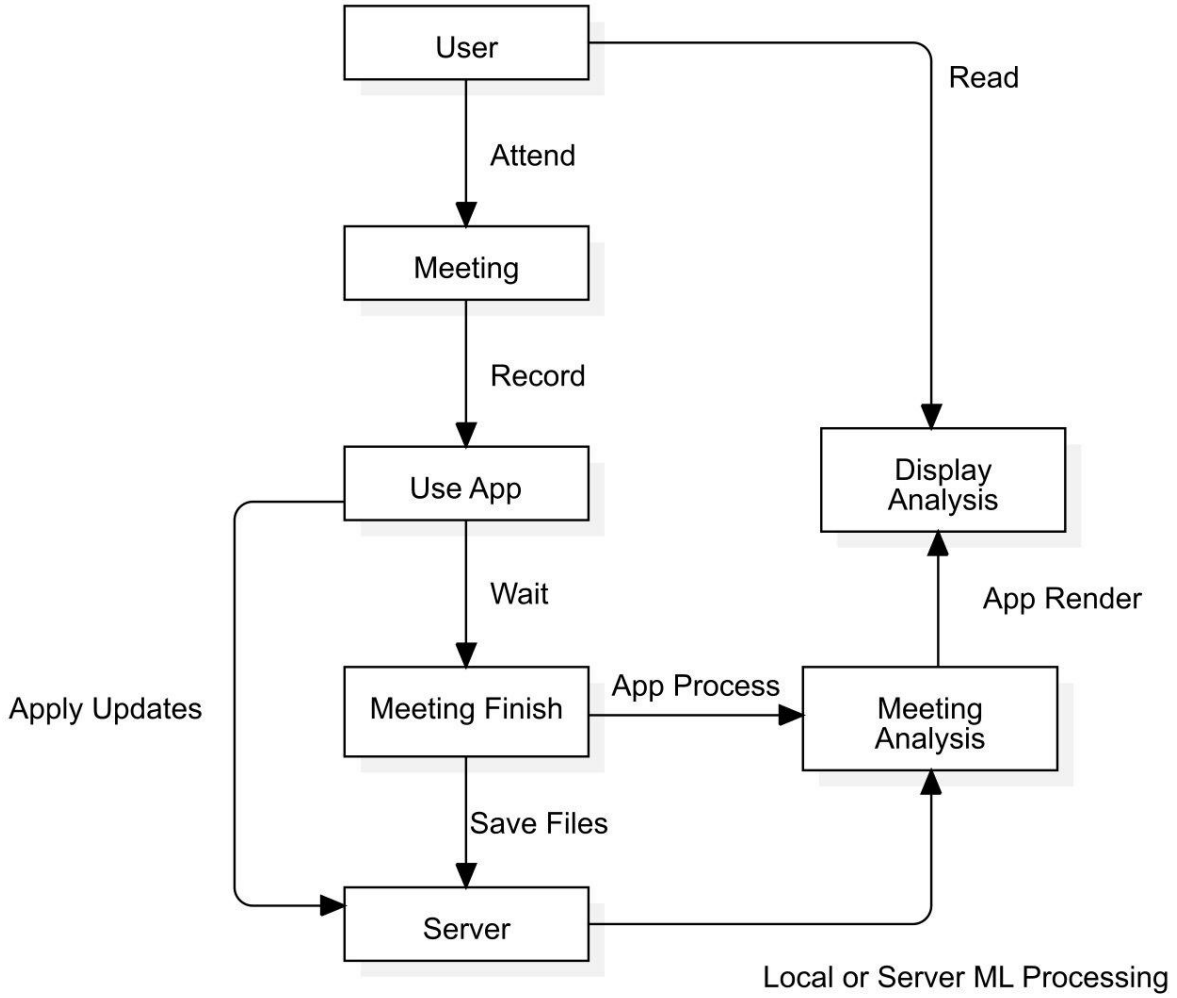

Fig 1

The aim of this project is to develop a system that provides insight and salient connections to the content of business meetings, using existing technologies, algorithms and Machine Learning models. The system transcribes recorded audio in real-time, using standard android phone hardware, then performs analysis which is displayed with graphics and charts. The online dashboard element of this system allows for administrative management and changes to users with authorisation.

The need for the system arises from the fact that many meeting notes are lost or otherwise never revisited - This is apparent with discussions with others who describe their experiences in office meetings. This system can refine and merge these lost notes, so that users can get a more complete mental picture / birds-eye-view of their collective notes. The connections between these notes can be linked through queries on terminology, organisational units/roles, and commonalities that are found from user interactions.

The insights provided by the system can be supplemented by automatic intelligent searching and scraping of information available on the public web, and also from a private organisation dashboard.

The generated charts and graphs describing the underlying data can be shared to others, and can be configured by the user for their specific use-case, such as ‘show me the most common words’, or ‘show me the most mentioned person’). These graphics provide an intuitive experience which encourages the collaborative usage of the system, and the sharing of meeting topics and materials.

The motivation for this project is seeing the end result, and the potential to revitalise how discussions are taken in the office, which we believe will add a lot of value for our users. Seeing patterns in the data of meetings over time can give the necessary perspective required to progress further in meeting subjects, and to turn unstructured conversations into requirements, plans, assessments or other structured reports. We expect that this system will become more valuable the more that it is used, as the underlying software can understand the business context better when presented with more data. To summarise, this project will create a knowledge base from recorded conversations, to provide insights to users, supplemented through company data and public data from the internet.

The collaborative element of our system is important, as we foresee that continued usage of our system will only improve the AI/ML model as more users upload supporting content. We will attempt to utilise methods to encourage users to track their meeting conversations, keep their notes organised, and to incentivize them to review their notes regularly. We intend to implement a sharing function inherent with the Android app infrastructure, so that files such as RDF, XML and JSON can be both sent and received, and compiled into PDF reports which can then be sent as a synopsis. Additionally, the raw meeting minutes and recorded audio can be shared to contacts on a user’s phone, or shared to another app which can then process the data further. This collaborative element will both encourage continued use of the product, and encourage collaboration of users with others in a productive manner.

1.2 Business Context

General Business Context:

This project fills in knowledge gaps that arise in note-taking for business meetings, and provides a more reliable experience, by using existing technologies to ensure important details are kept and revised on. This system will allow the users to review and analyse their interactions with their fellow staff and other departments.

Hospital Context:

The business context here will be primarily targeted at the St James Hospital Quality, Safety and Improvement Directorate (QSID) (https://www.stjames.ie/services/qsid/) This department "aims to provide care that is safe, effective and accessible so that (their) patients achieve the best possible outcomes"

There are multiple staff on this team, alongside the overseeing QSID director. Team members work on multiple safety and risk management projects and have regular meetings to keep up to date with each other on their collaborations. Not all meetings are in an official capacity, and sometimes they are performed while walking around the hospital, i.e. outside of a meeting room. The app should sufficiently accomodate for such unofficial meetings considering that it is voice-based.

Use Case:

This system will allow the staff to:

- Summarise their meetings about these projects
- Review the meetings that a user may not have been to
- Review the meetings as a manager, to review contributions of each staff member
- Save time in disseminating meeting outcomes
- Analyse commonalities in meetings and repeating patterns

Special emphasis will be placed on general business requirements i.e. standards for reliability, security etc, including industry standards.

Part of the understanding of the Hospital Risk-Management context, is that often during meetings, each user may want to revisit topics mentioned in previous meetings, and fact-check future topics through cross referencing and company shared databases This system can provide both, with a best-fit solution designed for their specific role in their department.

The material gathered during meetings is supplemented by admin users using the online company dashboard, as they can upload many types of documents which can provide more context to the AI/ML model. The organisation dashboard can be updated and enhanced with additional relevant company documents / information.

1.3 Glossary

- **ML/Machine Learning**: a branch of artificial intelligence in which a computer generates rules underlying or based on raw data that has been fed into it
- **Transcription**: Converting captured audio into text representations of the words spoken
- **EARS**: Easy Textual Requirements
- **Word2Vec**: A natural language processing technique used to learn word association from a large corpus of text. It can detect similar words and suggest them.
- **Natural Language Processing**: Creating software which can understand and interpret text or voice data, and respond with text and speech of its own.
- **SJH:** St James's Hospital
- **QSID:** Quality, Safety and Improvement Department
- **Sharding:** Sharding is a method of storing data records across many server instances
- **Sprint:** Scrum Sprint is a repeatable fixed time-box during which a "Done" product of the highest possible value is created
- **AI/Artificial Intelligence:** The ability of a computer or other machine to perform those activities that are normally thought to require intelligence, and the branch of computer science concerned with the development of machines having this ability. Artificial Intelligence is utilising multiple ML models into one usable system.
- **Transcription:** Converting spoken audio into text
- **Online Dashboard:** The online element of the system, which will enable the user(s) to manage their meeting data and recordings

General Description

2.1 Product / System Functions

The product records and filters the audio of a user's meeting, transcribes the audio to text, analyses the text based on user requirements, then the text is summarised according to the analysis. Depending on the user configuration options, graphics and charts are created to display the data collected. In addition to the graphics and charts, the system will provide user-friendly access to data such as ‘most common words’, ‘most mentioned staff member’, ‘list of mentioned facts’, and ‘list of mentioned opinions’

The term ‘analysis’ in terms of the text summary in this case refers to the conversion of text elements into symbols which are then ranked based on importance both to the overall context, and the relationship with other text elements, and importance relating to commonly discussed topics (both in the provided documents in the company dashboard, and previous meeting minute notes). Types of analysis that the system will compose are descriptive analysis, predictive analysis, and key takeaways.

The term content and context analysis refers to the display of gathered contextual information, with the aim of providing a common understanding of the data which can be shared among staff members.

The intelligence of the system is aided by the introduction of a given context by the user, prior to the transcription and processing of the recorded meeting audio.

The dashboard element of this product will serve as the base from which to both improve the accuracy of the machine learning model, and provide more contextual information to better serve what the user intends to store as meeting notes. This will be achieved through the ability given to the user to store their company documents on the dashboard, including previous meeting minutes, decks etc.

It is through the interaction between these two systems that the user will gather the most benefit, primarily through augmenting their abilities on recording their thoughts and discussions with orders, ranked through time.

The system should be flexible enough to allow changes by the user, to alter the processing and filtering of the incoming data.

The system will encourage cooperation between users, allowing for social features like the sharing of content as described above in the introduction. We believe that with continued use of the system, more opportunities will arise for the AI/ML model to improve itself through more metadata, context, and background information.

2.2 User Characteristics and Objectives

The user for our application should have basic knowledge on how to operate a modern android based smart phone. Due to the office environment of the hospital staff, we would expect the primary users to have a basic knowledge of both a smartphone and a personal computer.

The objectives directed for the user would be to manage and organise their collective meeting notes in a secure and organised database, and keep them for the future for review in further meetings.

The users in a hospital setting would have a risk-averse nature due to the high-stakes of the business background and the health of clients, and there are special business standards applicable to the medical system which we will need to account for. They would be dealing with confidential information which would necessitate GDPR and security considerations in the software.

The target users, the SJH QSID staff may have additional characteristics / requirements which we can observe and gather from in a future meeting.
A table describing each role of the department:

User	*Role*
Standard User	Using both the dashboard for personal use, and the mobile app to record the meetings which they take part in. They can share their recordings and customised generated graphics with others within the company. Their expertise of the system will be sufficient to just record their sessions and share content.
Admin User	They usually will use the online dashboard, and may have some extra access permissions to add / remove users from meetings, review more data etc. They will be expected to have a reasonable expertise of the system.
Special User / Read Only User	This user would be one from outside the organisation, and would usually have read-only access. They may want access to specific information for a chosen context. They may not have expertise with the system but may have advice from the admin user.

2.3 Operational Scenarios

The user can expect to find an intuitive experience when using this product, provided they have the necessary technical knowledge to use it. We have designed this product to be as straight-forward to use as possible, while still allowing features for power-users. Included in our features would be the ability to customise the output and functionality of the machine learning algorithm.

Scenario :

1. When joining a meeting, the user would trigger the recording feature of the application, and would expect to get visual confirmation of the recording, such as a waveform animation.
2. During the recording process, the user would expect to see a running text display of the transcribed text, and additionally would expect to see a clear visual summary of key points brought up in the meeting.
3. When the data is sent to the online database for processing and storage, the user would have a reasonable expectation of the safety and security of their data, as well as the expectation of timely delivery of content.
4. During the usage of the system, the user would expect the standard design principles of other apps to apply, so that they can use prior knowledge of their application experience to navigate the system.
5. Despite differing network conditions (including a loss of network status), the system should continue to work through a base level of functionality, which would enable the user to work with the data stored locally on their device.
6. The system should expose some of it's inner workings for the purposes of customisation to the user, through common design elements which allow for altering the transcribing and parsing of the recorded audio text.

This scenario is shown in Fig 1

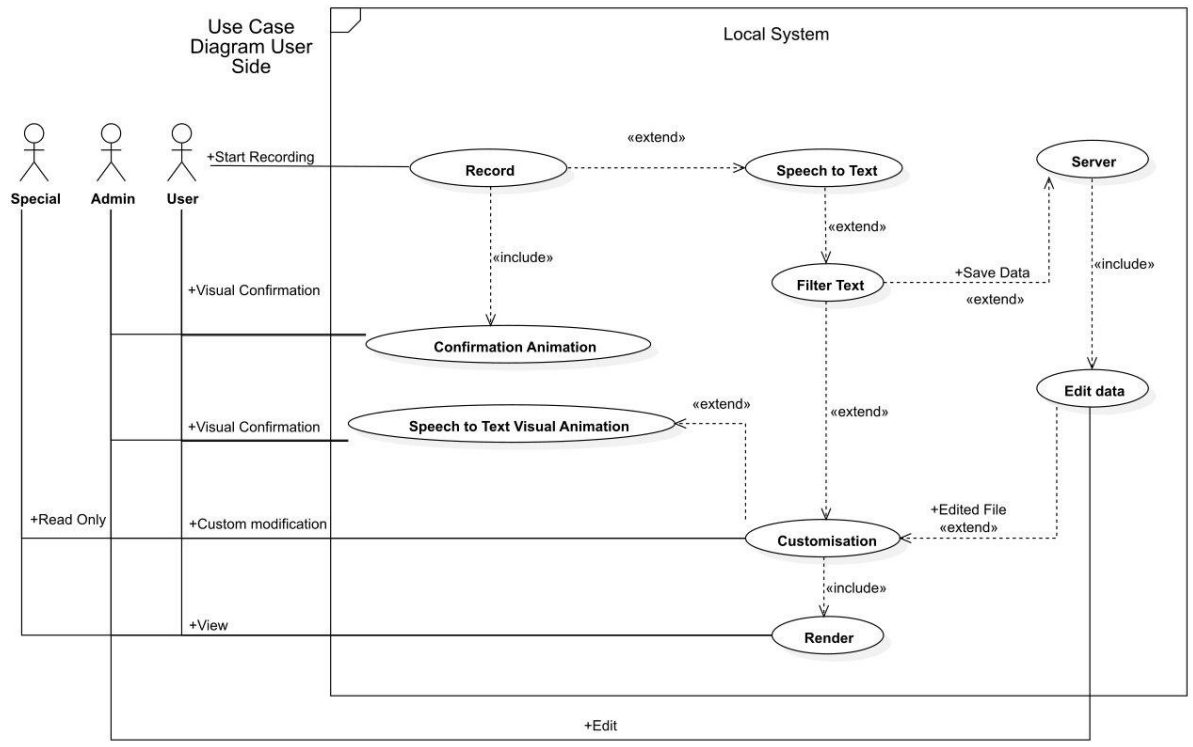

Fig 2

2.4 Constraints

1. The android system version may be a constraint, as specific functionality that were introduced in later versions may not be available in the previous.
2. The phone model itself may be a constraint, as the hardware may be lacking features such as a mic with noise cancelling.
3. The user's environment may be a constraint, as ambient environmental noise can affect the recorded audio quality
4. The user's technical ability may be a constraint, as they may have difficulty navigating an overly-technical product.
5. The user's network availability may be a constraint, as the system may need to function offline

3. Functional Requirements

For the below functional requirements, the EARS approach is used (see glossary), *they are ranked in order of importance* to the overall project.

1. When triggered by the user, the system shall record meetings, with user provided context.

2. When triggered by the user, the system shall transcribe audio files from recorded meetings

3. When the meeting audio has been converted into a transcription, the server shall summarise, annotate, and try to understand this transcription

4. When the user uploads the data to the dashboard, the dashboard shall securely store the audio, data and document data for future use

5. When the user requests data in the future from the dashboard, the dashboard shall retrieve the relevant data based on a search query and return it to the user, supplementing this data with those gathered from the public internet

6. When the user requests to view the data located on the dashboard, the dashboard shall display this generated content from summarised data, and enable the user to customise the display

7. When the user shares the content with their contacts, the system shall use the platform native tools to upload and download the content

8. When the user enables configuration options, the system allows for customisation of the algorithms and compilation

Non-Functional requirements

- **Reliability:** The reliability of the system will be ensured by a regular system test performed at the end of every sprint cycle, these tests will include unit testing, performance test and accuracy test.
- **Performance:** The performance of the system is heavily subjected to the version of the local system it is operating on, we can not guarantee great performance on unsupported platforms and versions of Android.
- **Maintainability:** The system is designed to function with a server on the backend, with no access to the internet, the local application can only perform limited functions.
- **Scalability:** The system's database is designed to have unlimited horizontal scaling, the server node can be scaled via sharding
- **Usability:** The application is designed to be as user friendly as possible, to achieve this we will be using most if not all of the current conventions for UI design.

Quality issues

The following are the Quality issues we have considered that might cause problems to the overall system.

- Responsiveness of the platform
    - The user will expect a reasonable amount of time for page loading, audio transcription, graphics creation etc
- Timing of the audio transcription
    - The user will expect the transcription of audio to be processed within 60 seconds of the meeting ending, so that they can continue with their workday un-distracted.
- Server up-time
    - Critical meetings must be recorded, and we can't afford for the system to be down during these times.
- Compatibility with hardware and software
    - The user would expect the software to work on their device without issues, despite the chosen browser or Android phone

The rationale behind considering these quality issues is that if the user does not get their expected level of quality and performance, they may consider other systems/services.

4. System Architecture

UML Component diagram, simple diagram that shows the main components of the system and how they are related to each other.

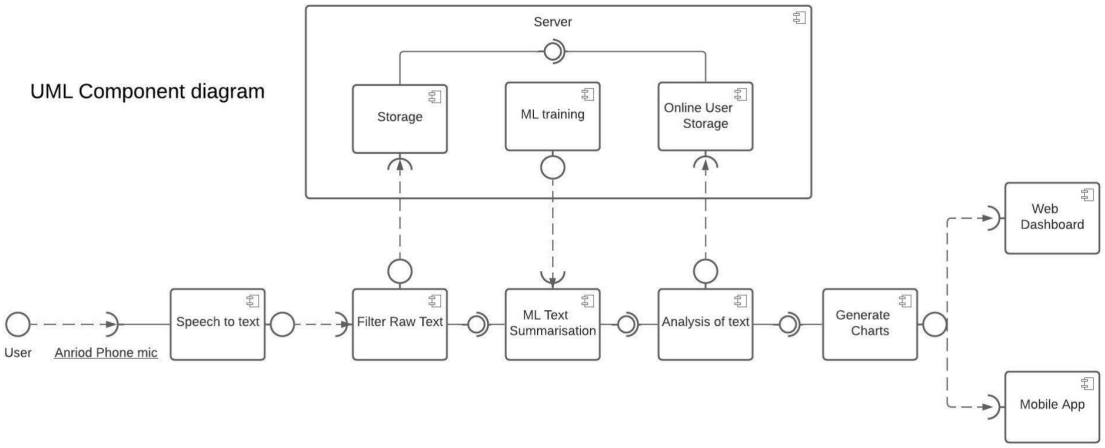

Fig 3

Intended devices and platforms:

- Android phone
- Computer with a supported browser

Hardware components

- Phone with at least Qualcomm Snapdragon 845 2.8Ghz
- Phone with at least 1 gb ram
- Phone with a Signal to Noise ratio of 79 db microphone
- Microphone with a Signal to Noise ratio of 79 db

5. High-Level Design

Main Class diagram, this diagram outlines the overall code structure for the system.

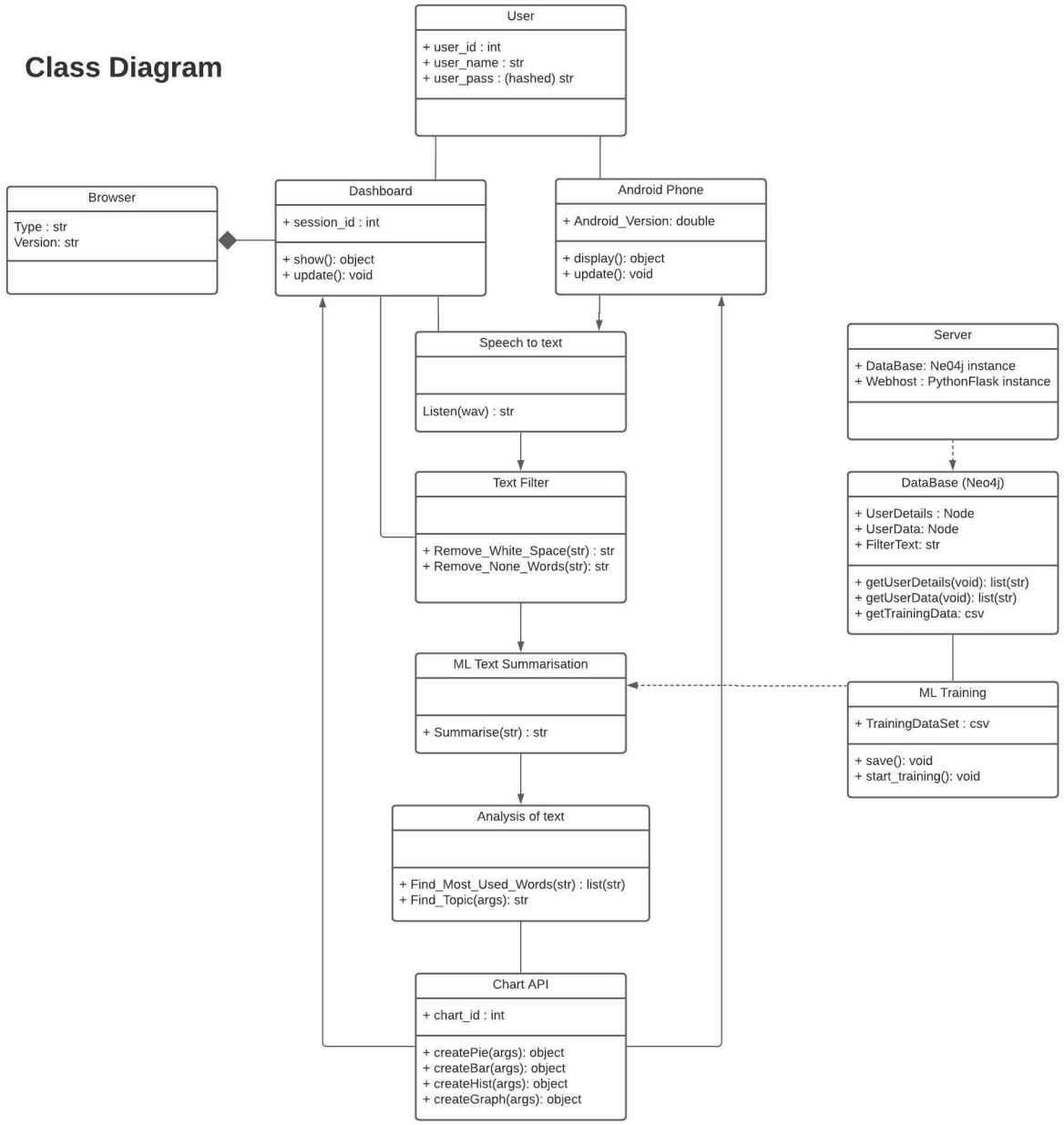

Fig 4

Sequence diagram for base use case, this diagram show how the system will respond when a user visits the Web Dashboard

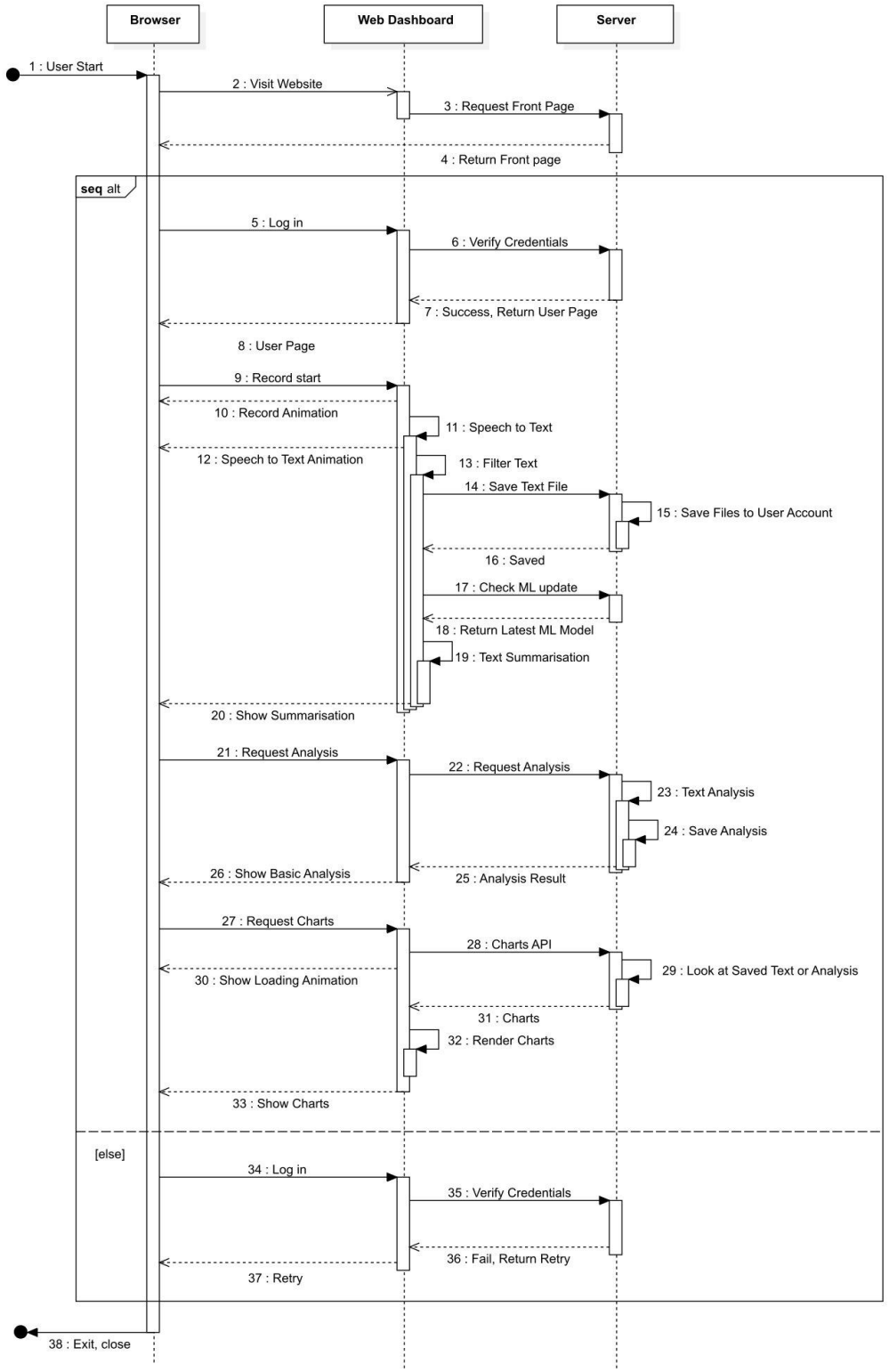

Fig 5

6. Preliminary Schedule

We decided to use the Agile approach for this project, here is the plan laid out in a gantt chart, this details the development plan for 1 Agile cycle. Each cycle is planned for 1 month for a planned 5 month development cycle.

Since we are taking the Agile approach it is hard to estimate exactly what the tasks are for each month ahead of time, but the general plan is the following.

- First month, we plan to implement the core systems, mainly addressing Text to Speech solutions, servers and database solutions
- Second month, we plan to start training our first ML model to summarise the transcribed text
- Third month, we will create the Charts and algorithms for analysis of the summarised text
- Fourth month, we plan to refine the front end of the web interface and phone app
- Fifth month, will mainly be allocated to testing and quality assurance

Fig 6.1

Link to Full Gantt Chart:

https://docs.google.com/spreadsheets/d/1fawKvLeJ6axhogWClpQ4i4AFTlq3EuKY/edit?usp=sharing&ouid=109100529714017224924&rtpof=true&sd=true

Resource Requirements:

1. At least 2 personal computers for development team
2. 1 server for web hosting
3. Dedicated GPU for neural network training
4. At least 1 Android phone for development team

Full Gantt Chart Screenshot

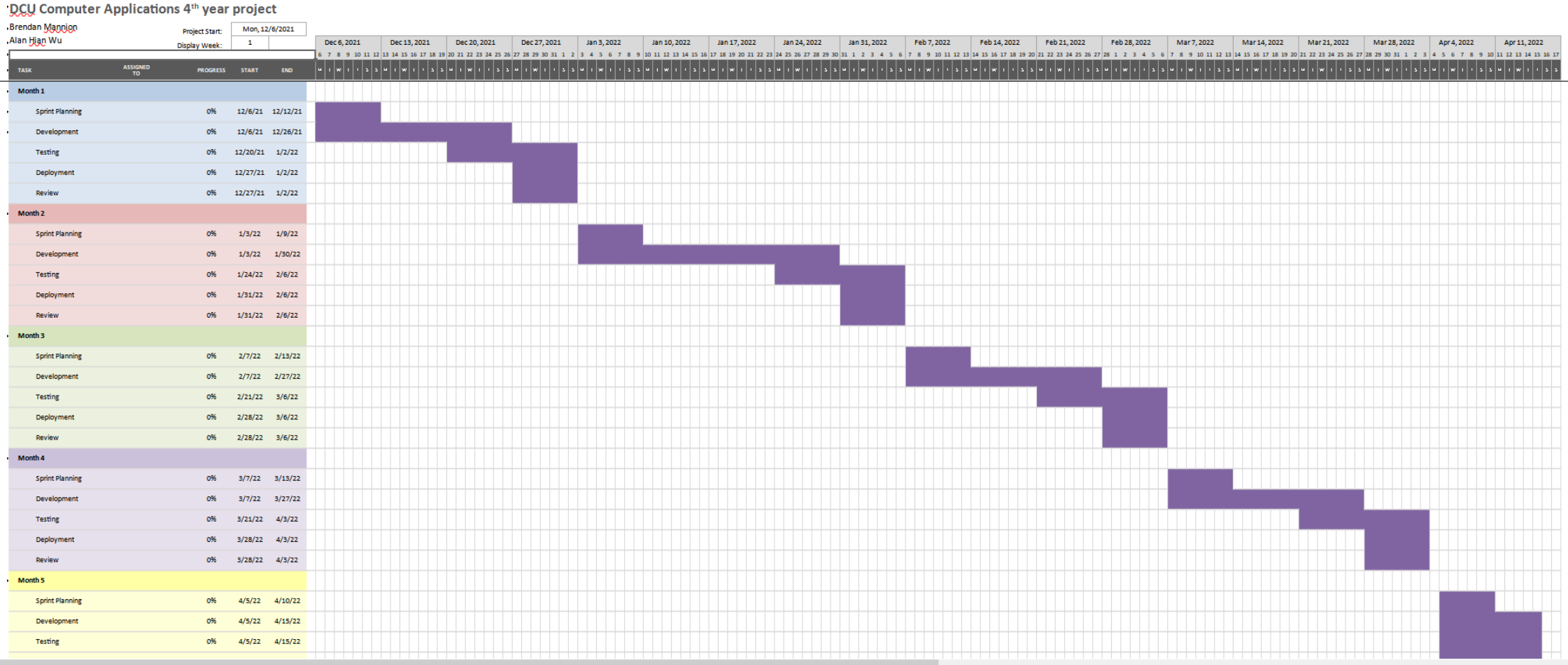

Fig 6.2

7. Appendices

https://www.tensorflow.org/tutorials/text/word2vec

https://www.ibm.com/cloud/learn/machine-learning

Figures