

Investigation Project Plan

Robyn Gebbie (2127777), Jesse Van Der Merwe (1829172)

School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg 2050, South Africa

Abstract — The EIE Investigation Capstone Project “Which Hand”, supervised by Dr V. Aharonson, will be undertaken during the second block of semester 2 of 2022. This report describes the project plan, work breakdown structure and Gantt charts, which have been generated using project management software Mind Genius and MS Project. Further, the required resources are described, and risks and ethical considerations included.

I. INTRODUCTION

The Electrical/Information Engineering Investigation Project requires students to undertake a significant investigation that involves research, design, implementation, and investigation of a system typical of a complex engineering problem. The project undertaken for this investigation is that of “Which Hand”, supervised by Dr V. Aharonson, which requires the implementation of computational and statistical analysis to provide quantitative insight to assess the severity of a subject’s tremor in order to determine the outcome of Parkinson’s disease and Essential Tremor treatment. This report will cover a comprehensive plan as to how this project will be implemented, taking into consideration any risks and ethical concerns that may arise during the execution of the project.

II. PROJECT BACKGROUND

Focused Ultrasounds (FUS) treatment has the potential to reduce the tremor in patients with Parkinson’s Disease (PD). The effectiveness of this treatment can be determined by analyzing traced spiral drawings done by PD patients in order to determine the tremor present in the drawings. In order to determine whether this treatment is decreasing the severity of PD in patients with the illness in a way that decreases the influence of human error and bias, a machine learning approach can be developed to determine whether a spiral drawing has been done by a patient with PD or simply by a patient that has Essential Tremor (ET). This model can be created with the use of Artificial Intelligence (AI) and will require use of image classification and preprocessing software.

III. PROJECT DESCRIPTION

This project will involve research and investigation into FUS and its effect on reducing the tremor in subjects who have PD or ET. The treatment is applied to one side of the brain and the result of the treatment will therefore only be visible in the treated side of the brain. The project involves the implementation of a model that minimizes the effect of human error and bias in the diagnoses process of PD.

The implementation of this project will require data in the form of traced spiral drawings drawn by subjects that have been diagnosed with PD and have the illness to varying severity, as well as subjects that do not have PD, but simply show signs of

ET. Each subject will need to trace the spiral with both their right and left hands in order for the effectiveness of the FUS treatment to be determined on the treated side of the brain. In order to get permission to use this data, a waiver must be submitted to Rambam Medical Center. A close consideration of data ethics must be carried out throughout this process.

Research will need to be done on the effectiveness of the FUS treatment and how it affects the tremor present in drawings drawn by patients. This research will be used in coordination with an analysis of the spiral drawings drawn by patients in order to determine whether the treatment is lessening the severity of PD in the side of the brain receiving treatment.

Research into effective methods of classifying the drawings as either being drawn by a subject with PD or ET will need to be done. This research will be largely focused on machine learning models that will be capable of accurately processing and classifying the spiral image data. This research will include investigation into appropriate methods to perform image processing and preprocessing in order for a dataset to be created for the implementation of the model.

The implementation of the AI model will require data preparation and preprocessing in order for the spiral image data to be used effectively. The implementation of this project will involve the development of a machine learning program that is able to accurately classify the PD and ET drawings and determine whether a drawing has been done by a subject that has PD. The model will help to determine the effectiveness of the FUS treatment on the PD patients.

IV. PROJECT MANAGEMENT

For the project to be delivered successfully and on time, a comprehensive plan needs to be developed. This plan needs to include the project scope, as well as any project deadlines with a comprehensive description of the requirements for each deadline. A work breakdown structure (WBS) has been generated for this project using [Mind Genius 20](#). This allows for the entire project to be broken down into manageable sections, whose overall lengths could then be estimated. Exporting this into [Microsoft Project](#) allowed for the further addition of relevant project data, including lengths of individual tasks as well as the order and reliance of tasks. Both these aforementioned project management software tools will help generate and control the scope of the project, as well as the success criteria. This will be expanded upon in the Methodology section below.

V. METHODOLOGY

A detailed and complete WBS of the entire project consisting of 5 main components was created. This complete WBS can be found in appendix A. This WBS was then imported into MS Project, and the approximate hours per task allocated. This allowed MS Project to produce a Gantt chart and overall timeline. This chart can be found in Appendix A. According to the approximate hours per task currently allocated, the entire project will take 300 hours, as required by the CBO. The project should also be completed by mid-November. However, this current project plan allows for almost not error or variability. It is known that errors, mistakes, and delays are rampant in projects, and thus more efficient resource leveling, and task allocation must occur before the project can go forward. This includes better allocation of the hours between the two group members, as well as more streamlined order in which tasks are to be completed. Each of the five main components are detailed below.

A. Investigation

The first component of this project includes investigation (figure 1). To better understand the field in which this project operates in, a few hours should be spent researching the basics of the associated medical theory. This includes Parkinson's Disease (PD), Essential Tremors (ET), Static Spiral Tests (SST), the bias and inaccuracies of healthcare professionals, as well as the various methods of diagnosing PD/ET. Even though the project is simply to evaluate the effectiveness of the treatment, both PD and ET (neurological disorder) need to be researched in order to understand the similarities and differences. This will be vital if the functionality of differentiating between the two is to be implemented.

Next, research must be conducted in order to determine the best possible data preparation method in order to process the provided images, as well as to pre-process the resulting data. These provided images are pictures/scans of spiral drawings of patients that have PD or ET, as further expanded in the Required Resources section of this report. The way in which the data is prepared and split into training data, as well as data augmentation and hyper-parameter optimization must also be researched and carefully considered, as these aspects will greatly vary and affect the final results.

Finally, research must be conducted in order to determine the best possible machine learning/deep learning method to be implemented in order to

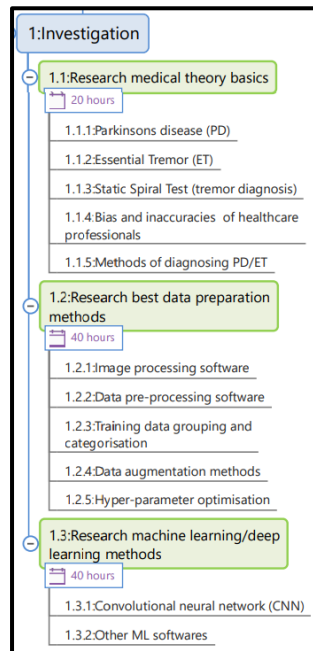


Figure 3: WBS - Investigation

generate the most accurate and useful results.

B. Design and Implementation

After the above research has been conducted, the design and implementation of each section can commence (figure 2). Once the most applicable image processing method is determined, this can be designed, coded, and implemented. Following, once the image processing has been complete, and the most applicable data pre-processing method is determined, the data-preprocessing software can be designed, coded and implemented. Following, once the data has been pre-processed and optimized, and the best machine learning method chosen, this software can be designed, coded and implemented in order to generate final results.

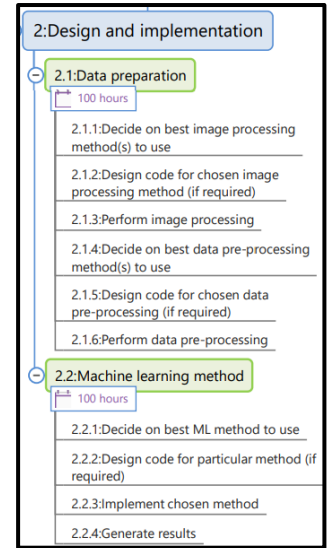


Figure 1: WBS - Design and Implementation

This section makes up the bulk of the project and thus 100 hours were initially allocated for each subsection. However, after loading the WBS into MS project and allocating more specific hours to each subtask, it was determined that perhaps 75 hours for data preparation implementation, and 215 hours for machine learning method implementation would be more appropriate. This includes factors such as both members of the project team being able to work on the first subsection concurrently, therefore halving the overall time required. These intricate details will be further evaluated and worked out during the first few days of the project in order to better determine a suitable schedule for coding and implementing of the software.

C. Budget

After investigation and research has occurred, it will be clear as to whether any software required for this project will need to be bought, or whether such software resources are freely available already – either through free packages on the internet, or through the university. If such a purchase is required, it must be discussed with the project supervisor in order to generate and submit a purchase order to the School.

D. Validation (Success Criteria)

Once the final solution has been designed and implemented, it will need to be tested and validated in order to determine whether the solution acquired is accurate and/or useful. Both

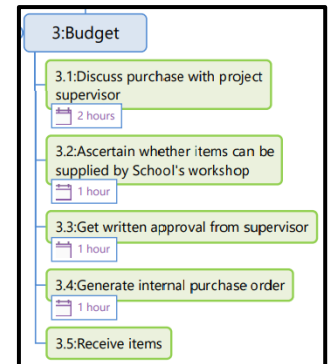


Figure 2: WBS - Budget

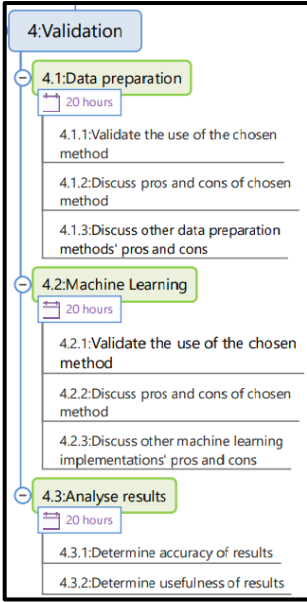


Figure 4: WBS - Validation

based on the particular test data used. This will help determine whether the solution is useful, as only a very high percentage accuracy will be considered acceptable in this particular application. The mis-classification between PD and ET might be considered acceptable since both PD and ET are movement disorders that result in tremor [10], however, effectively zero mis-classification between PD/ET and the control subjects is required.

E. Documentation

Finally, documentation must be planned for since this project involves not just the research, design and implementation of an applicable solution, but also the final submission of a project report. Further, each group member must keep a detailed Engineering Notebook throughout the project. Weekly meetings with the supervisor must be held, with minutes and agenda recorded at each. An ethics waiver must be filled in and approved (discussed further in the Ethics section below). A final group presentation must be prepared, and an individual interview conducted. All of these are important aspects of the project that must be considered and planned for accordingly.

the choice of data preparation method(s) used and machine learning method(s) used should be validated, with the pros and cons of each method thoroughly discussed and evaluated. Reasonings for not choosing a particular method should also be explained (if any).

As mentioned in the Ethics section below, this project is considered medium risk as the results of the implemented solution predict whether a subject has improved or worsened PD or ET. Thus, these final results must be analyzed in order to determine the accuracy of the results

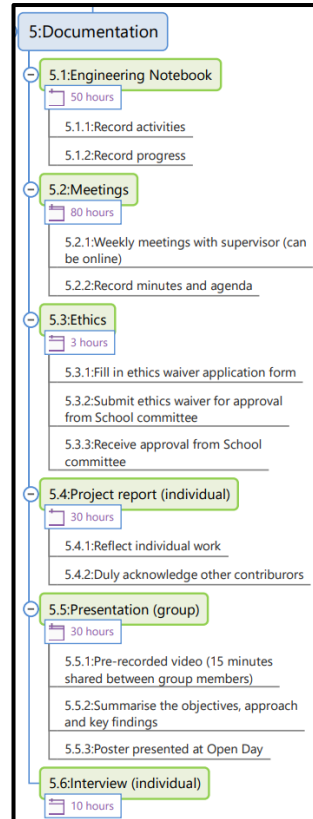


Figure 5: WBS - Documentation

VI. REQUIRED RESOURCES

This project will require the use of Artificial Intelligence (AI) to perform a statistical analysis of spiral drawings done by patients that have Parkinson's Disease. The drawings will be used to train a machine learning model to differentiate between drawings that have been drawn by patients that have Parkinson's and those who have an Essential Tremor (ET).

A. Training Data

In order to train the AI model, sufficient data containing the spiral drawings done by individuals that have been diagnosed with Parkinson's Disease and those who do not have the illness will be required. This data will be provided by Rambam Medical Center and will be in the form of scanned images containing traced spiral drawings for each subject. Each subject will draw a clockwise and an anticlockwise spiral using each hand, totaling in 4 spirals for each patient. The tremor present in each spiral will be analyzed in order to determine the likelihood of the patient having Parkinson's Disease and the severity of the Parkinson's Disease for each patient that has been diagnosed and is undergoing treatment.

Data containing information about each subject's age, gender and date of PD diagnoses will also be provided. This data can be used to train a model to recognize the effect of these input characteristics on the severity of PD in each subject.

B. Image Processing Software

The project involves researching image preprocessing techniques that can be used to analyze scanned images of the spiral drawings. As a starting point, some image processing techniques will be considered, and further research will be done on these techniques during the implementation of this project in order to find the optimal method.

One of the most widely image processing and visual AI techniques is the use of a Convolutional Neural Network (CNN). CNN is a deep learning model that processes data in a grid format, making it an optimal method for image processing. This method has many supporting architectures, with some of the common ones being AlexNet, GoogleLeNet and VGG. These image processing architectures are supported by Python and can be implemented with the use of environments such as VSCode or Jupyter Notebook.

An alternative to CNN is a Graph Neural Network, this model can be implemented in Python with the use of packages such as Deep Graph Library, PyTorch Geometric and Spektral.

C. Image Preprocessing Software

It is required to research image preprocessing techniques in order to prepare the image data for the processing stage. The preprocessing resources are dependent on the machine learning model that is being implemented.

For most models, image preprocessing will require the resizing of all images to a 256x256 pixel dimension, image augmentation and test-time augmentation. These processes can be achieved with the use of a Python package and can be implemented using environments such as VSCode or Jupyter Notebook. Python has

several packages that support the image preprocessing implementation, some of which are Scikit-image and Keras.

VII. RISKS

Before beginning this project, it is important to consider any possible risks that implementing this project may pose. The risks are related to the type of data that is required to train the AI models and general risks that come with implementing an AI.

A. Data Privacy

The data required for this project contains personal medical information for various subjects. These people have given permission for their data to be used in research, however it is important to not to violate the conditions under which the data is allowed to be used. Violation of these conditions could pose harm or distress to the people whose information is contained in the data.

B. Bias

It is important that a fair AI model is developed that is inclusive of all social groups. If the AI model is fed training data that excludes a particular class of people, it will produce results that do not consider the full population demographic. The AI must therefore receive data that does not allow any form of unfairness to develop.

C. Unreliable Results

One of the greatest risks is that the resulting AI model will not accurately predict the severity of PD or the presence of PD in a test subject. This risk is related to the accuracy of the data preparation process and the implementation of the machine learning model. In order to ensure that this does not occur, it is important to perform extensive tests with data with known results in order to ensure that the model is producing the correct results.

D. Security

Due to the complexity of training an AI model, it is initially difficult to tell whether the data provided is accurate and has undergone the proper preprocessing. Bad data can be introduced into the model, resulting in incorrect results that do not meet the final requirements. The extensive testing of all stages of process can allow for cases like this to be caught and dealt with.

VIII. ETHICS

This project involves the use of data obtained through human research. This data contains information about a test subject's age, gender and date of Parkinson's Disease diagnoses, as well as spiral drawings done by the subject that can be used to track the severity of PD in the patient. The subjects whose data is available for this study have given permission for their information to be used for research.

In order to use the data, an ethics waiver must be submitted to the Rambam Medical Center. The facility will then provide permission for the data to be used for this project.

A risk analysis can be done to categorize the risk at which the subjects are put in when participating in the research. The risk can be categorized as either having no risk, minimal risk, low risk, medium risk or high risk. In order for the research to have no risk, no human participants must be involved. High risk research involves research performed with human participants in which the participants are in risk of harm, this includes investigation into sensitive topics that may cause the participants emotional distress. This research can be considered to be medium risk research. The data used for this research has been obtained with the use of human participants, these participants have been diagnosed with PD and are required to trace a spiral with their left and right hands. The research has the potential to cause the subject emotional distress under the condition that the research results in the discovery that the subject has a severe or worsened stage of PD.

IX. CONCLUSION

This project is to be completed within the second block of the second semester of 2022. This plan serves as the basis from which to orientate the workflow and task allocation between group members in order to ensure a well planned and implemented project that is not rushed or left to the last minute. An ethics waiver must be completed, submitted and approved before use and manipulation of the data can occur. After research and investigation with regards to the required resources and methods to be used within this project, the design and implementation of a solution will begin. The methodology has been briefly mentioned and is to be further fleshed out as each section is tackled during the project. The solution is to be analyzed and validated, with the results and accuracy of such results being determined and thoroughly documented.

REFERENCES

- [1] N. Seedat, V. Aharonson, and I. Schlesinger, "Automated machine vision enabled detection of movement disorders from hand drawn spirals," in 2022 IEEE International Conference on Healthcare Informatics (ICHI), 2020: IEEE, pp. 1-5.
- [2] M. Tripathi, "Image Processing using CNN | Beginner's Guide to Image Processing", *Analytics Vidhya*, 2021. [Online]. Available: <https://www.analyticsvidhya.com/blog/2021/06/image-processing-using-cnn-a-beginners-guide/>. [Accessed: 24-Jul- 2022].
- [3] P. Walpita, "Convolutional Neural Networks For Artificial Vision", *Medium*, 2020. [Online]. Available: <https://priyalwalpita.medium.com/convolutional-neural-networks-for-artificial-vision-455be7c85d15>. [Accessed: 24-Jul- 2022].
- [4] J. Brownlee, "Best Practices for Preparing and Augmenting Image Data for CNNs", *Machine Learning Mastery*, 2022. [Online]. Available: <https://machinelearningmastery.com/best-practices-for-preparing-and-augmenting-image-data-for-convolutional-neural-networks/>. [Accessed: 24-Jul- 2022].

- [5] J. Leban, "Image recognition with Machine Learning on Python, Image processing", *Medium*, 2020. [Online]. Available: <https://towardsdatascience.com/image-recognition-with-machine-learning-on-python-image-processing-3abe6b158e9a>. [Accessed: 24- Jul- 2022].
- [6] R. Yadav, "Top 7 Image Processing Libraries In Python", *Analytics India Magazine*, 2022. [Online]. Available: <https://analyticsindiamag.com/top-8-image-processing-libraries-in-python/>. [Accessed: 24- Jul- 2022].
- [7] A. Fredrick, "Getting Started with Image Preprocessing in Python", *Engineering Education (EngEd) Program / Section*, 2021. [Online]. Available: <https://www.section.io/engineering-education/image-preprocessing-in-python/>. [Accessed: 24- Jul- 2022].
- [8] "Overview of Graph Neural Networks", *OpenGenus IQ: Computing Expertise & Legacy*, 2022. [Online]. Available: <https://iq.opengenus.org/graph-neural-networks/>. [Accessed: 24- Jul- 2022].
- [9] R. Teja, "Let's Talk About Graph Neural Network Python Libraries!", *Medium*, 2021. [Online]. Available: <https://towardsdatascience.com/lets-talk-about-graph-neural-network-python-libraries-a0b23ec983b0>. [Accessed: 24- Jul- 2022].
- [10] S. Singh, "Risk Assessment for AI projects", *Linkedin.com*, 2021. [Online]. Available: <https://www.linkedin.com/pulse/risk-assessment-ai-projects-sonu-singh>. [Accessed: 24- Jul- 2022].
- [11] Microsoft, "Microsoft Project", 2021. [Online]. Available: <https://www.microsoft.com/en-za/microsoft-365/project/project-management-software>. [Accessed: 24- Jul- 2022].
- [12] MindGenius, "MindGenius 20", 2021. [Online]. Available: <https://www.mindgenius.com/mindgenius-20/>. [Accessed: 24- Jul- 2022].

Appendix A

1. Work Breakdown Structure
2. Gantt Chart

EIE Investigation: "Which Hand"

19 Sep

1: Investigation

1.1: Research medical theory basics

20 hours

- 1.1.1: Parkinsons disease (PD)
- 1.1.2: Essential Tremor (ET)
- 1.1.3: Static Spiral Test (tremor diagnosis)
- 1.1.4: Bias and inaccuracies of healthcare professionals
- 1.1.5: Methods of diagnosing PD/ET

1.2: Research best data preparation methods

40 hours

- 1.2.1: Image processing software
- 1.2.2: Data pre-processing software
- 1.2.3: Training data grouping and categorisation
- 1.2.4: Data augmentation methods
- 1.2.5: Hyper-parameter optimisation

1.3: Research machine learning/deep learning methods

40 hours

- 1.3.1: Convolutional neural network (CNN)
- 1.3.2: Other ML softwares

2: Design and implementation

2.1: Data preparation

100 hours

- 2.1.1: Decide on best image processing method(s) to use
- 2.1.2: Design code for chosen image processing method (if required)
- 2.1.3: Perform image processing
- 2.1.4: Decide on best data pre-processing method(s) to use
- 2.1.5: Design code for chosen data pre-processing (if required)
- 2.1.6: Perform data pre-processing

2.2: Machine learning method

100 hours

- 2.2.1: Decide on best ML method to use
- 2.2.2: Design code for particular method (if required)
- 2.2.3: Implement chosen method
- 2.2.4: Generate results

3: Budget

3.1: Discuss purchase with project supervisor

2 hours

3.2: Ascertain whether items can be supplied by School's workshop

1 hour

3.3: Get written approval from supervisor

1 hour

3.4: Generate internal purchase order

1 hour

3.5: Receive items

4: Validation

4.1: Data preparation

20 hours

- 4.1.1: Validate the use of the chosen method
- 4.1.2: Discuss pros and cons of chosen method
- 4.1.3: Discuss other data preparation methods' pros and cons

4.2: Machine Learning

20 hours

- 4.2.1: Validate the use of the chosen method
- 4.2.2: Discuss pros and cons of chosen method
- 4.2.3: Discuss other machine learning implementations' pros and cons

4.3: Analyse results

20 hours

- 4.3.1: Determine accuracy of results
- 4.3.2: Determine usefulness of results

5: Documentation

5.1: Engineering Notebook

50 hours

- 5.1.1: Record activities
- 5.1.2: Record progress

5.2: Meetings

80 hours

- 5.2.1: Weekly meetings with supervisor (can be online)
- 5.2.2: Record minutes and agenda

5.3: Ethics

3 hours

- 5.3.1: Fill in ethics waiver application form
- 5.3.2: Submit ethics waiver for approval from School committee
- 5.3.3: Receive approval from School committee

5.4: Project report (individual)

30 hours

- 5.4.1: Reflect individual work
- 5.4.2: Duly acknowledge other contributors

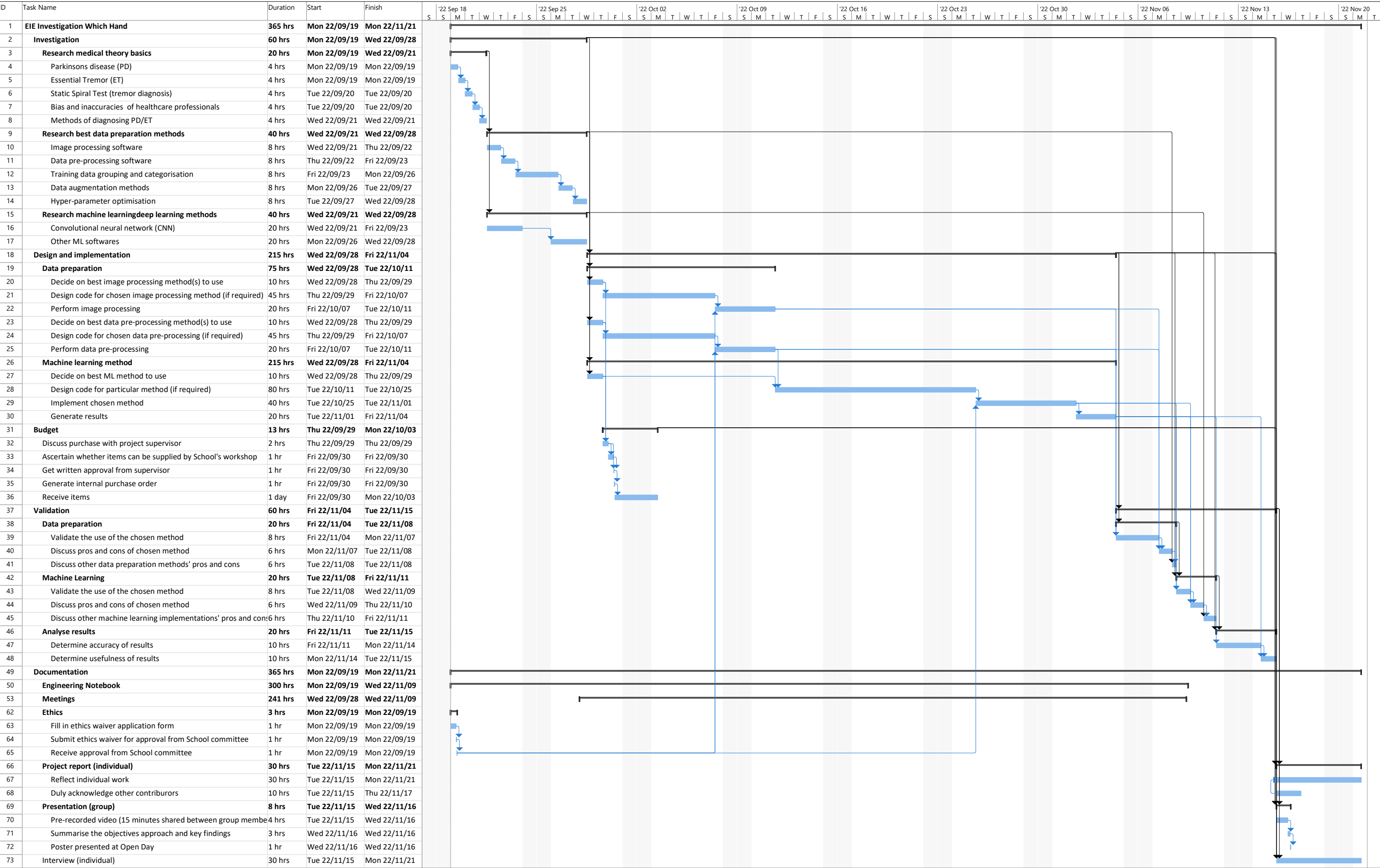
5.5: Presentation (group)

30 hours

- 5.5.1: Pre-recorded video (15 minutes shared between group members)
- 5.5.2: Summarise the objectives, approach and key findings
- 5.5.3: Poster presented at Open Day

5.6: Interview (individual)

10 hours



Project: WhichHand_WBS_With

Date: Mon 22/07/25

Task

Split

Milestone

Summary

Project Summary

Inactive Task

Inactive Milestone

Inactive Summary

Manual Task

Duration-only

Manual Summary Rollup

Manual Summary

Start-only

Finish-only

External Tasks

External Milestone

Deadline

Progress

Manual Progress