

Criterion B: Record of Tasks

| Task Number | Planned Action | Planned outcome | Time estimated | Target completion date | Criterion |
|-------------|---|--|----------------|------------------------|-----------|
| 1 | Brainstorming for projects related to NLP or artificial intelligence | Google docs with list of potential projects narrowed down to the two categories or a merge of both | 2 days | 01/02/2023 | A |
| 2 | Discussing with advisor and deciding on suitable client for project | Narrowed down project ideas to focus on the use of NLP in podcast project and found client to work with | 1 hour | 06/02/2023 | A |
| 3 | First meeting with client to gain a deeper understanding of the client's issue and solution | Gained a more tangible understanding of the issue and the general approaches of how it could be tackled | 1.5 hours | 15/02/2023 | A |
| 4 | Research into NLP and web scraping, understanding the required knowledge to create successful solution to | Watched a full tutorial on NLP, including bag of words, word vectors, regexes, BERT model, and more. Watched another video on using beautiful soup for web | 4.5 hours | 20/02/2023 | A |

| | | | | | |
|---|---|---|------------|------------|-----|
| | issue | scraping. | | | |
| 5 | Create a list of main success criterias needed for solution and simple illustrated outline of how the solution will proceed | Listed out the main functions of the product required for the solution and created a simple flowchart of the solution | 2 hours | 6/03/2023 | A,B |
| 6 | Second client interaction to see if the rough solution and criteria established matches with client requirements | Client requests the function of having a single input be changed to allow several inputs that output a suitable recommended podcast to allow for more flexibility and specificity. Furthermore, the client requests that the product extends its functions to include features like a summarizer, and a simple and convenient user interface to easily access information about the podcast that is recommended. | 45 minutes | 10/03/2023 | A |
| 7 | Complete Criterion A | Finished writing out the issue raised by the client, our consensus | 3.5 hours | 6/04/2023 | A |

| | | | | | |
|----|--|---|-----------|------------|---|
| | | on its solution, and the libraries, functions, subsequent research and methods required to tackle the issue. | | | |
| 8 | Create a modified flowchart illustration of product | Drew a flowchart that illustrates the steps taken to allow multiple inputs to output recommended podcasts with relevant features requested such as a word cloud, transcript and more. | 1 hour | 15/04/2023 | B |
| 9 | Research into the libraries and modules needed to complete these steps | Did research into podcast recommendation systems, word clouds, and text summarization using NLP. | 4 hours | 24/04/2023 | B |
| 10 | Implementing research techniques and building a system that incorporates NLP into the podcast recommendation system. | Tested out the use of NLTK, BERT, and Gensim for text summarisation of podcast descriptions. | 4.5 hours | 1/05/2023 | B |
| 11 | Perform a | A hypothesis test is | 10 hours | 11/06/2023 | |

| | | | | | |
|----|---|---|---------|------------|---|
| | hypothesis test to prove that euler distance of vector representation of keywords can be used to statistically measure relevance of inputs. | conducted after data of random inputs and similar inputs are collected for over 2000 input results. | | | B |
| 12 | Figure out that cannot replace nan with zero or else it will skew the values when doing hypothesis testing | Instead replace with average values of similar and non-similar centroid from previous sets of inputs | 30 min | 13/06/2023 | B |
| 13 | After inputting missing values, test for skewness of the distribution to see if it is normal, and if the hypothesis is correct. | The graph displayed from the centroid data is normally distributed suggesting that the hypothesis is correct. | 5 hours | 20/06/2023 | B |
| 14 | Planning a simple flask interface connecting to backend using flask which | I have successfully drafted a plan to connect the inputs from a form to the backend | 45 min | 22/06/2023 | B |

| | | | | | |
|----|--|---|---------|------------|---|
| | incorporates HTML and CSS in design | | | | |
| 15 | Creating a simple user interface to fill out inputs through HTML and rendering it using flask and confirming that input has been received by connecting interface to new page that displays the input results. This will demonstrate the successful use of flask in connecting different HTML pages. | I have successfully written an HTML form that is able to take in the three user inputs and using the post method transfer the data to a new HTML page that is able to display those inputs. | 2 hours | 29/06/2023 | C |
| 16 | Creating function that validates the user inputs | I have successfully created a function that checks if the user inputs return legitimate podcast results from the google podcast api. If successful, then the function would display the | 1 hour | 24/06/2023 | C |

| | | | | | |
|----|--|---|----------|------------|---|
| | | <p>user_inputs on a new HTML page routed using flask.</p> <p>Otherwise, the user will have to re-enter the inputs on the form.</p> | | | |
| 17 | <p>Compartmentalising all the backend functions into a main function that can access all the relevant features after calling itself, making the process automatic.</p> | <p>I have gathered all the relevant functions needed to produce the final similar keyword matching the user inputs and the relevance of the keyword to those inputs. The two results can thus be utilised for frontend.</p> | 1 hour | 09/07/2023 | C |
| 17 | <p>Connecting the frontend to the backend; after processing the inputs and checking they are legitimate, then they can be processed through all the backend features to return the most similar keyword that can match to various podcasts for the</p> | <p>After input is checked, I connected it to the main backend function that was able to process those inputs and output the relevant data to be displayed. This was successfully rendered on a new HTML page called podcast_recommendation.html</p> | 1.5 hour | 20/07/2023 | C |

| | | | | | |
|----|--|---|------------|------------|---|
| | user. | | | | |
| 18 | <p>Since the processing of the inputs takes a long duration, I needed an intermediary page that displays a loading spinner whilst processing is done in the background. So, I researched ways to use flask to render a loading screen.</p> | <p>This took quite a while to research as I was still unfamiliar with flask. What I needed was a function that was able to call the main backend function whilst loading the spinner. After the backend is done, it will bring me to a new page that renders the results. To this end, I finally found the right function that was able to perform this. I then implemented it in my HTML document which is also connected to flask to route to different functions when rendering the different HTML interfaces.</p> | 2 hours | 25/07/2023 | C |
| 19 | <p>I needed to find a basic spinner that can be used to check if the loading page works</p> | <p>I searched on github and found a spinner that is written in css. Because it is just a simple spinner, it doesn't need to be written in Javascript used for more complicated animation effects. Using css, this</p> | 30 minutes | 27/07/2023 | C |

| | | | | | |
|----|--|--|--------|------------|---|
| | | can be easily implemented as a stylesheet in my loading HTML file. Upon display, a loading spinner was rendered! | | | |
| 20 | I needed to find a professional bootstrap template with prewritten CSS and HTML code that fits the specific theme of a podcast recommendation website: this would be minimal in design, clean, and user-friendly | Searching for a while, I found a bootstrap template that fits the prerequisite criteria for the front end that allows user to easily input data in a form-like template. | 30 min | 07/08/2023 | C |
| 21 | Learn CSS and HTML to change the format of the form | I learned how to tweak the formatting of the form using HTML and CSS, allowing only three input boxes positioned in the middle of the form | 30 min | 20/08/2023 | C |
| 22 | Learn how to connect the database to my | I successfully connected the SQLite database to my | 45 min | 24/08/2023 | C |

| | | | | | |
|----|---|--|-----------|------------|---|
| | python script to increase efficiency of the program. | python file such that previous user inputs with their corresponding keywords can be accessed and directly processed. | | | |
| 23 | Taking client suggestions of implementing auto-suggestion features to the homepage form to increase user-friendliness of website. | Searched for tutorials on implementing auto-complete functions through the use of javascript and python from data in the database. | 2 hours | 05/09/2023 | C |
| 24 | Successfully connecting the database such that frontend renders auto-suggestions in real-time when user types. | The database is successfully connected such that input records are auto-suggested to the user in real-time when the user types. | 1.5 hours | 10/09/2023 | C |
| 25 | Display a datatable of previous inputs and corresponding keywords in the previous inputs page. | I came across a tutorial that allows you to create an interactive data table based on your database. I decided to implement it. | 1 hour | 20/09/2023 | C |

| | | | | | |
|----|---|--|-----------|------------|---|
| 26 | Successfully render the datatable with all of its corresponding functionalities properly working in the previous inputs page. | I have successfully done that by importing the necessary libraries, and writing a javascript code for the datatable. | 1.5 hours | 02/10/2023 | C |
| 27 | Create a search query function that is able to return a wordcloud of the keywords to a corresponding input query. | I successfully created a search box that is able to redirect to a new page with the word cloud when the user submits an input in the data table. | 3 hours | 11/10/2023 | C |
| 28 | Complete documentation for C | I have successfully outlined important functionalities of the program from the code and completed documentation for C. | 1 month | 12/11/2023 | C |
| 29 | Complete documentation for criterion B. | Criterion B is successfully finalised. | 3 weeks | 13/12/2023 | B |
| 30 | Complete documentation for criterion E. | Criterion E is successfully finalised. Client feedback is written. | 2 weeks | 30/12/2023 | E |
| 31 | Complete demonstration | I have successfully completed the | 3 days | 03/1/2023 | D |

| | | | | | |
|----|--------------------------------------|--|-------|-----------|---|
| | video for criterion D. | demonstration video and edited it such that each success criterion is hit. | | | |
| 32 | Final interaction with client | Verified that all the success criteria have been hit. Discussed directions for future development | 1 day | 04/1/2023 | E |
| 33 | Deliver final version of the product | Project is submitted. | 1 day | 05/1/2023 | E |