|  |  |  |
| --- | --- | --- |
| **Date** | **Team members present** | **Points discussed, names of contributors and any relevant material** |
| 14/04/2023 | * Kevin Uturgaury   (3246031) | (Min 0 - 30) Today I started with Exploratory Data Analysis and visualizations in Google Collab and started adding them to the Report Template that was provided to me by lecturer.  I mounted the drive and imported various libraries for the Exploratory Data Analysis.  I performed df.head(), df.tail() and df.shape() methods and added these to the Report. |

# Kevin Uturgaury U3246031 LogBook

|  |  |  |
| --- | --- | --- |
| **Date** | **Team members present** | **Points discussed, names of contributors and any relevant material** |
| 15/04/2023 | * Kevin Uturgaury   (3246031) | (Min 0 - 30) Today I continued with the EDA and performed the df.describe(), df.corr(), df.columns methods in the Google Collab and added them to my Report.  (Min 30 - 60) I peformed the df.nunique(), df.info() methods to show unique values and info about the dataset in the EDA and added this information to the Report.  (Min 60 - 90) I then proceeded to add figures to the EDA that visually display information relevant to the dataset and added these figures to my report. |

|  |  |  |
| --- | --- | --- |
| **Date** | **Team members present** | **Points discussed, names of contributors and any relevant material** |
| 21/04/2023 | * Kevin Uturgaury   (3246031) | (Min 0 - 30) Today I continued with the EDA and identified the outliers in the data. I drew up a graph and added it to the Report.  (Min 30 - 60) I asked and answered my first question in the EDA section which was “What are the 2 classes for this dataframe?” and then I added it to my Report.  (Min 60 - 90) I then to asked and answered my second question, which was “What percentage of people survived and died?” and then added it to my Report.  (Min 90 – 120) I then added the heatmap that shows correlation between the attributes to the EDA and Report. I then used ProfileReport from pandas\_profiling to create a report and ask more question based on the stats provided by the report. I also asked and answered my 3rd question “Which sex comprised the majority of the passengers?” and added it to my Report. |

|  |  |  |
| --- | --- | --- |
| **Date** | **Team members present** | **Points discussed, names of contributors and any relevant material** |
| 22/04/2023 | * Kevin Uturgaury   (3246031) | (Min 0 - 30) Today I asked and answered my 4th question: “From which port did the majority of passengers depart from?” and I added it to my Report.  (Min 30 - 60) I asked and answered my 5th question: “To which Passenger class did the majority of passengers belong to?” and I added it to my Report.  (Min 60 - 90) I then started with the Predictive Data Analytics Stage and performed the pre-processing and normalization upon the dataframe and added it to my Report. I stared with encoding the object columns of my dataframe into integer values and added it to my report.  (Min 90 – 120) I then normalized the data and added the normalized data visualization to my report. |

|  |  |  |
| --- | --- | --- |
| **Date** | **Team members present** | **Points discussed, names of contributors and any relevant material** |
| 23/04/2023 | * Kevin Uturgaury   (3246031) | (Min 0 - 30) Today I prepared the data for model building.  (Min 30 - 60) I then proceeded to zip the encoded values together.  (Min 60 - 90) I then split the data into training and test data sets for model evaluation and compared the different analytical models to see which one is the best. I added the results to my report.  (Min 90 – 120) I then used “Performance Evaluation Metric 1 – Classification Report” to generate a report on the best classification model. |

|  |  |  |
| --- | --- | --- |
| **Date** | **Team members present** | **Points discussed, names of contributors and any relevant material** |
| 25/04/2023 | * Kevin Uturgaury   (3246031) | (Min 0 - 30) Today I used the “Performance Evalutaion Metric 2 – Confusion Matrix” and added it the results to my report.  (Min 30 – 60) I used the “Evaluation Metric 3 – Roc-Auc Curve” and added the results to my report.  (Min 60 – 90) I used the “Evaluation Metric 4 – Prediction Report” and added it to my report. |

|  |  |  |
| --- | --- | --- |
| **Date** | **Team members present** | **Points discussed, names of contributors and any relevant material** |
| 25/04/2023 | * Kevin Uturgaury   (3246031) | (Min 0 - 30) Today I used the “Performance Evalutaion Metric 2 – Confusion Matrix” and added it the results to my report.  (Min 30 – 60) I used the “Evaluation Metric 3 – Roc-Auc Curve” and added the results to my report.  (Min 60 – 90) I used the “Evaluation Metric 4 – Prediction Report” and added it to my report.  (Min 90 – 120) I then proceeded to perform the code Implementation and Deployment Stage by using the best performing algorithm as a desktop Tkinter software tool to make predictions about who will survive. |
| **Date** | **Team members present** | **Points discussed, names of contributors and any relevant material** |
| 29/04/2023 | * Kevin Uturgaury (3246031) | (Min 0 - 30) Today I finished off my project presentation template. |