Project Report on

AUTONOMOUS ANALYSIS OF SPACECRAFT DATA IN POWER DESCENT

By

Name-Angarika Das , RegNo-201600453

In partial fulfillment of requirements for the award of degree in
Bachelor of Technology in Computer Science and Engineering
(2020)

Under the Project Guidance of

INTERNAL GUIDE:
CHITRAPRIYA N.,
ASSISTANT PROFESSOR Gr-I,
SIKKIM MANIPAL INSTITUTE OF TECHNOLOGY

EXTERNAL GUIDE:
PRASHANT KULSHESTRA,
SCIENTIST/ENGINEER,
U.R.RAO SATELLITE CENTRE,
INDIAN SPACE RESEARCH ORGANISATION

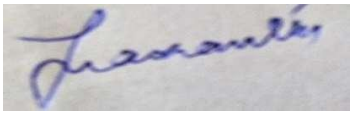
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SIKKIM MANIPAL INSTITUTE OF TECHNOLOGY

(A constituent college of Sikkim Manipal University)
MAJITAR, RANGPO, EAST SIKKIM – 737136

Project Completion Certificate

This is to certify that the below mentioned student(s) of Sikkim Manipal Institute of Technology has / have worked under my supervision and guidance from …13/01/2020…….. to …30/04/2020…… and has / have successfully completed the project entitled "…AUTONOMOUS ANALYSIS OF SPACECRAFT DATA IN POWER DESCENT…" in partial fulfillment of the requirements for the award of Bachelor of Technology in Computer Science and Engineering.

University Registration No	Name of Student(s)	Course
201600453	Angarika Das	B.Tech (CSE)



Prashant Kulshreshtha

Scientist/Engineer "SF"

OnBoard Software Development Group

U.R. Rao Satellite Centre(URSC), Bangalore.

Old Airport Road, Vimanpura Post, Bangalore-560017

Ph:080-25082346

PROJECT REVIEW CERTIFICATE

This is to certify that the work recorded in this project report entitled **"Autonomous Analysis Of Spacecraft Data In Power Descent"** has been jointly carried out by **Ms. Angarika Das (Reg. 201600453)** of Computer Science & Engineering Department of Sikkim Manipal Institute of Technology in partial fulfilment of the requirements for the award of Bachelor of Technology in Computer Science and Engineering. This report has been duly reviewed by the undersigned and recommended for final submission for Major Project Viva Examination.

Mrs. Chitrapriya N.,
Assistant Professor Gr-1,
Department of Computer Science & Engineering,
Sikkim Manipal Institute of Technology

CERTIFICATE OF ACCEPTANCE

This is to certify that the below mentioned students of Computer Science & Engineering Department of Sikkim Manipal Institute of Technology (SMIT) have worked under the supervision of **Mr. Prashant Kulsheshtra, EA Delivery Lead of Sprint Account** of **Ericsson India Global Service Pvt. Ltd., Kolkata** from 06 January 2020 to 08 May 2020 on the project entitled **"Web Application Development Using Spring Boot".**

The project is hereby accepted by the Department of Computer Science & Engineering, SMIT in partial fulfilment of the requirements for the award of Bachelor of Technology in Computer Science and Engineering.

University Registration No.	Name of Student(s)	Project Venue
201600453	Angarika Das	U.R.Rao Satellite Centre, Indian Space Research Organisation, Bengaluru

Prof. (Dr.) Kalpana Sharma,
Professor & HoD,
Computer Science & Engineering Department,
Sikkim Manipal Institute of Technology,
Majitar, Sikkim-737136

DECLARATION

I , the undersigned, hereby declare that the work recorded in this project report entitled "…Autonomous Analysis Of Spacecraft Data In Power Descent….." in partial fulfillment for the requirements of award of B.Tech (CSE) from Sikkim Manipal Institute of Technology (A constituent college of Sikkim Manipal University) is a faithful and bonafide project work carried out at "….U. R. Rao Satellite Centre, Indian Space Research Organisation, Bengaluru…." under the supervision and guidance of Mr. Prashant Kulsheshtra of U. R. Rao Satellite Centre, Indian Space Research Organisation, Bengaluru.

The results of this investigation reported in this project have so far not been reported for any other Degree / Diploma or any other Technical forum.

The assistance and help received during the course of the investigation have been duly acknowledged.

Name of Student- ANGARIKA DAS (Reg no-201600453)

ANGARIKA DAS

(Signature with date)

ACKNOWLEDGEMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and I am extremely privileged to have got this all along the completion of my project. All that I have done is only due to such supervision and assistance and I would not forget to thank them.

I respect and thank Mr. Prashant Kulsheshtra Sir, for providing me an opportunity to do the project work in U R Rao Satellite Centre, Indian Space Research Organization and giving me all support and guidance which made me complete the project duly. I am extremely thankful to Chitrapriya Ningthoujam Ma'am, Sikkim Manipal Institute of Technology for providing such a nice support and guidance. I pay my deep sense of gratitude to Prof. (Dr.) Kalpana Sharma, H.O.D, Computer Science & Engineering Department, Sikkim Manipal Institute of Technology to encourage me to the highest peak and to provide us the opportunity to prepare for the project. I am immensely obliged to my Major Project Coordinators, Mr. Sourav Paul, Mrs. Chitrapriya Ningthoujam, Mr. Santanu Kr. Mishra and Mr. Biswaraj Sen for elevating inspiration and kind supervision in completion of my project.

Last, but not least, I would like to thank my Computer Science & Engineering Department, Sikkim Manipal Institute of Technology and to all the faculty members for giving me continuous support and guidance that has helped me in completion of my project.

Name of the student-Angarika Das
(Reg No-201600453)

DOCUMENT CONTROL SHEET

1	Report No	CSE/Major Project/External /B.Tech/1/2020
2	Title of the Report	Autonomous Analysis Of Spacecraft Data In Power Descent
3	Type of Report	Technical
4	Author	Angarika Das
5	Organizing Unit	U.R. Rao Satellite Centre, Indian Space Research Organisation (ISRO), Bengaluru
6	Language of the Document	English
7	Abstract	Data Analysis project where the information of the descent of a spacecraft is explored and visualized for prediction.
8	Security Classification	General
9	Distribution Statement	General

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Onboard Computers are the brain of the satellite. It is also where the onboard software of a satellite runs. The task of the onboard software is to provide functions, procedure and services for different purposes in the spacecraft. Massive data have been collected through various space missions. To maximize the investment, the data need to be exploited to the fullest. In this project, I am trying to explore and visualize the dataset which contains the information of the descent of a spacecraft.

For example-The data during the descent of a spacecraft will contain important details which would tell us at what time was the Space Craft at a particular angle descending downwards. The autonomous analysis of this data will help to compare the new data with the previous one. This will highlight the trajectory of a spacecraft, as well as the deviation from its path. The comparison of the trajectories of various spacecraft's will be done and meaningful insights can be derived. This result analysis will help to identify errors and learn from the data.

1) INTRODUCTION

1.1) General Overview Of The Problem:

The dataset which contains the information of the descent of a spacecraft is huge in number. This data contains important observation of the trajectory which is studied for its analysis. The size of the dataset makes it hard for doing this task manually. The autonomous analysis of this data would make this job easier in order to do comparative analysis or detection of errors.

This autonomous system will help to achieve goals while operating independently. It will point out the important features in a huge dataset and display the trajectories of various spacecraft's. The decisions will be made better using the rich onboard data. The autonomous analysis will make performance better and reduce system cost.

Softwares and Libraries Used:

- Anaconda Distribution : Anaconda also known as the world's most popular data science platform is a free and open source distribution of the python and R programming languages for scientific computing applications. It is primarily written in Python and is available for Windows, macOS and Linux. It is developed by Anaconda, Inc and is extensively used in data science and machine learning projects.

- Jupyter Notebook : A spinoff from the IPython project, it is an open source web application in which we can create codes, equations, visualisations and texts. This notebook ships with the IPython kernel and hence we can write code in python but it also supports Julia and R. Jupyter Notebooks are a powerful and simple way to write and iterate on our Python code for data analysis. We can write lines of code and run them one at a time which saves us from the tedious task of writing and rewriting the lines of code. It is easier to make changes, edit and rerun the program again, all in the same window.

- Numpy : NumPy is a Python package and it stands for 'Numerical Python'. This library consists of multidimensional array objects and a collection of routines for

processing of array. Mathematical and logical operations on arrays can be performed using the NumPy library.

- **Pandas :** This software library is used for data manipulation and analysis. It is mostly used in form of dataframes for machine learning and also for various data manipulation operations. It does practical, real world data analysis for python and is used for loading databases , handling missing data to name a few.

- **Matplotlib :** Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib makes easy things easy and hard things possible. It develops publication quality plots with just a few lines of code and uses interactive figures that can zoom, pan, update etc. (source : https://matplotlib.org/)

-Seaborn : Seaborn is a library for making statistical graphics in Python. It is built on top of matplotlib and closely integrated with pandas data structures. Seaborn aims to make visualization a central part of exploring and understanding data. Its dataset-oriented plotting functions operate on dataframes and arrays containing whole datasets and internally perform the necessary semantic mapping and statistical aggregation to produce informative plots. (source: https://seaborn.pydata.org/)

- **Sklearn :** Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. The library is built upon the SciPy (Scientific Python) that must be installed before you can use scikit-learn. This stack that includes:

NumPy: Base n-dimensional array package

SciPy: Fundamental library for scientific computing

Matplotlib: Comprehensive 2D/3D plotting

IPython: Enhanced interactive console

Sympy: Symbolic mathematics

Pandas: Data structures and analysis

(source:https://machinelearningmastery.com/a-gentle-introduction-to-scikit-learn-a-python-machine-learning-library/)

-Tkinter: Python with tkinter is the fastest and easiest way to create the GUI applications. Creating a GUI using tkinter is an easy task. Tkinter is included with standard Linux, Microsoft Windows and Mac OS X installs of Python. The name Tkinter comes from Tk interface. Tkinter was written by Fredrik Lundh. Tkinter is free software released under a Python license. (source: https://en.wikipedia.org/wiki/Tkinter)

1.2) Literature Survey:

SL No.	Research	Findings	Relevance To Project
1.	"Analysis Of Big Data From Space" - The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Wuhan, China, VOLUME: XLII-2/W7 , ISSUE: 18–22 September 2017, J. Tan, B.Osborne.	Massive amount of space data is not exploited to maximize investment. No systematic discussion of methodologies to meet the demands of space data taking its features into account.	Systematical research on characteristics and applications of space data to make full use of it.

1.3) Problem Definition

The size of the spacecraft data makes it impossible to analyze it manually. This dataset is used for comparing different factors and attributes within the dataset. The main objective to achieve this problem is that these data contain massive amount of information which are not analyzed or automated. The various rows and columns of the raw dataset represent different factors and are not pre-processed. Hence, the data present cannot be used easily to predict the various factors of the space craft in descent with regard to a particular time. A good starting point would be to transform the data in a useful and efficient manner. The summarization of the main characteristics of the dataset would impact whether a particular statistical method can be used or not. This will explain the data in a minimum number of predictor variables. The problem for the dataset to be utilized in an autonomous manner for prediction and to identify the relationship between variables, the data needs to be analyzed properly with an appropriate model development. This task if done manually would be tedious and won't pave the way for simplification and effectiveness of the entire process, apart from being time consuming.

1.4) Analysis Of The Problem And The SRS

For the analysis and visualization of the given problem, it can be divided into four steps:

1) Data Wrangling
2) Exploratory Analysis of the data
3) Model Development
4) Creation of the model with GUI

SRS

1)Functional requirements:

1.1) Purpose: The purpose of this project is to make the analysis of spacecraft descent data autonomous which will save a lot to time and effort that is given to do it manually

1.2)Scope: This project will make visualisation and analysation of the spacecraft data easier to inspect . This will help in detailed analysis of the descent of the spacecraft simpler to read.

1.3)Indented Audience: The intended audience of this project will be the scientists and researchers working. They are constantly looking forward to do some work and so when they spend a lot of time in ananlysing data on their own, it takes a lot of time and effort. The analysation may or may not be correct if it is done manually. The autonomous analysation will give precise and accurate results from several files in a short duration.

2)Non-functional requirement

The requirements of this project include as mentioned earlier:

-Anaconda Distribution, Version- Anaconda3-2019.03-Windows-x86_64

-Python 3.0

-Jupyter Notebook

- Python packages

- Laptop edition- Windows 10 Pro

-RAM: 4.00 GB

-System Type- 64 bit operating system

1.5) Solution Strategy

To overcome the problems mentioned, a several number of steps were followed.

1) Data Wrangling : It is the method of processing the data which prepares the data for analyzing. With this process, the raw data is transformed into another format which makes it more appropriate. Data wrangling process helps in better decision making with the dataset. This is a very important step to follow before analysis of the data.

    a. Pre processing data in python : This step is done so that the data is formatted in such a way that more than one machine learning algorithms are executed in one dataset and the best out of them is chosen.

    b. Dealing missing values : Real world data often has missing values. These datasets have missing values for a number of reasons like data corruption, observations of data not being recorded etc. It is a part of data acquisition.

    c. Data formatting: For meaningful comparisons, data should be brought to a meaningful common standard of expression. Data formatting allows the data for easy aggregation and makes it more clear. In situations where the wrong data type is assigned to the feature, using data formatting, the data type is identified, corrected and converted to the correct data type.

d. Data normalization : Normalization refers to rescaling real valued numeric data to a similar value range, which will have a similar intrinsic influence on analytical model

e. Binning : This method is used to handle noisy data. The data is sorted and distributed into a number of bins or buckets.

2) Exploratory Data Analysis : Exploratory Data Analysis or (EDA) is understanding the data sets by summarizing their main characteristics often plotting them visually. This step is very important especially when we arrive at modeling the data in order to apply Machine learning. Plotting in EDA consists of Histograms, Box plot, Scatter plot and many more. It often takes much time to explore the data. Through the process of EDA, we can ask to define the problem statement or definition on our data set which is very important.

a. Groupby: A groupby operation involves some combination of splitting the object, applying a function, and combining the results. We can create a grouping of categories and apply a function to the categories. It's a simple concept but it's an extremely valuable technique that's widely used in data science. Groupby concept is really important because it's ability to aggregate data efficiently, both in performance and the amount code is magnificent. The "groupby" method groups data by different categories. The data is grouped based on one or several variables and analysis is performed on the individual groups.

b. Analysis of variance : The Analysis of Variance (ANOVA) is a statistical method used to test whether there are significant

differences between the means of two or more groups. ANOVA returns two parameters:

F-test score: ANOVA assumes the means of all groups are the same, calculates how much the actual means deviate from the assumption, and reports it as the F-test score. A larger score means there is a larger difference between the means.

P-value: P-value tells how statistically significant is our calculated score value. If our price variable is strongly correlated with the variable we are analyzing, expect ANOVA to return a sizeable F-test score and a small p-value.

C. Correlation :

Correlation: a measure of the extent of interdependence between variables.

Causation: the relationship between cause and effect between two variables.

It is important to know the difference between these two and that correlation does not imply causation. Determining correlation is much simpler the determining causation as causation may require independent experimentation.

The Pearson Correlation measures the linear dependence between two variables X and Y.

The resulting coefficient is a value between -1 and 1 inclusive, where:

1: Total positive linear correlation.

0: No linear correlation, the two variables most likely do not affect each other.

-1: Total negative linear correlation.

Pearson Correlation is the default method of the function "corr". We can calculate the Pearson Correlation of the of

the 'int64' or 'float64' variables.

3) Model Development: A model can be thought of as a mathematical equation used to predict a value given one or more other values. It relates one or more independent variables to dependent variables. The model fitting the desired criteria is built using the results of this analysis.

a. Simple Linear Regression: It is the basic and commonly used type for predictive analysis. It is a statistical approach to modelling the relationship between a dependent variable and a given set of independent variables. The objective of a linear regression model is to find a relationship between one or more features(independent variables) and a continuous target variable(dependent variable).

b. Multiple Linear Regression: Multiple Linear Regression attempts to model the relationship between two or more features and a response by fitting a linear equation to observed data. These regression estimates are used to explain the relationship between one dependent variable and one or more independent variables. Mathematically, multipel regression estimates a linear regression function defined as:

$y = c + b1*x1+b2*x2+\ldots+bn*xn$

Where y = estimated dependent variable score, c = constant, b = regression coefficient, and x = score on the independent variable.

4) Creation of the model with GUI : After analysis and comparison, we know that the Multiple Linear Regression model is the best fit model for the prediction of the dataset.

Tkinter: The Tkinter module (“Tk interface”) is the standard Python interface to the Tk GUI toolkit from Scriptics (formerly developed by Sun Labs). Both Tk and Tkinter are available on most Unix platforms, as well as on Windows and Macintosh systems. Starting with the 8.0 release, Tk offers native look and feel on all platforms.

1.6)Preliminary User’s Manual

The preliminary user is the person analyzing the dataset. Initially, to get a proper understanding of what part this project is from, Onboard Software, Onboard Computers and Satellite Operations were studied. Data collected is raw in nature and no meaningful insights could be derived from the dataset. The dataset contains huge amount of information which was not interpreted. The first task was to research the initial steps of analysation of any data. A dummy dataset was implemented to experiment in the data transformation process. Similar steps were rechecked with the original dataset. The implementation process were carried out in that exact way. After proper analysis of the data, the prediction GUI had been created keeping in mind the main purpose of the project. The user of the manual will not know about the confidential datasets used in the project. However, the prediction with the dataset is something the user will definitely see.

1.7)Organization Of The Report

1.7.1)INTRODUCTION:

This section contains the meaning and purpose of the project in simple terms. It contains the general overview of the problem which is the basic explanation of the purpose of this project. The literature survey in this section shows the previous research works related to this project. The problem definition explains the project's expectations in brief. The Analysis of the Problem and SRS defines the functional and non-functional requirements of the project. The solution strategy gives a gist of how we would be continuing with the project and fulfill it's objectives.

1.7.2) DESIGN STRATEGY:

This section contains the architectural diagram and the flowchart. The architectural diagram or the block diagram represents the interconnection of the various modules of the solution strategy. The flowchart consists more of a detailed visualization of the implementation of the project.

1.7.3) TEST PLAN

This shows the input and output scenarios of the test cases used to show the running of the project. It consists of the input given by the user and the output generated with this input. If the testing is accurate, that would mean that the system has passed and is successful.

1.7.4) IMPLEMENTATION DETAILS

This section explains the pseudo code of the project and explains how the project was done.

1.7.5)RESULTS AND DISCUSSION

This section contains the screenshots of the results and application of the project.

1.7.6) SUMMARY AND CONCLUSION

This section focuses on the summary of achievements in the project as well as the main difficulties encountered and how they were tackled. It also discusses the limitations of the projects along with the future scope of work. It also records any special observations if any during the project.

1.7.7) GANTT CHART

It depicts the time wherein the various developments and levels of the project was accomplished along with the display of dates.

1.8.8) REFERENCES

The websites and links visited for the research work of this project is mentioned in this section.

2) DESIGN STRATEGY FOR THE SOLUTION

The design strategy is as follows:

- BLOCK DIAGRAM

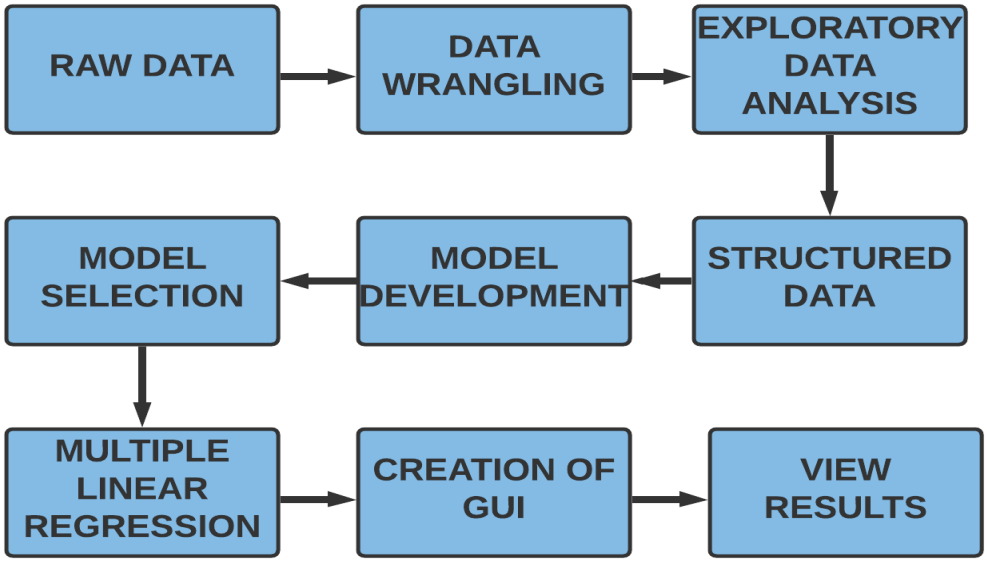


Figure 2.1

- FLOWCHART

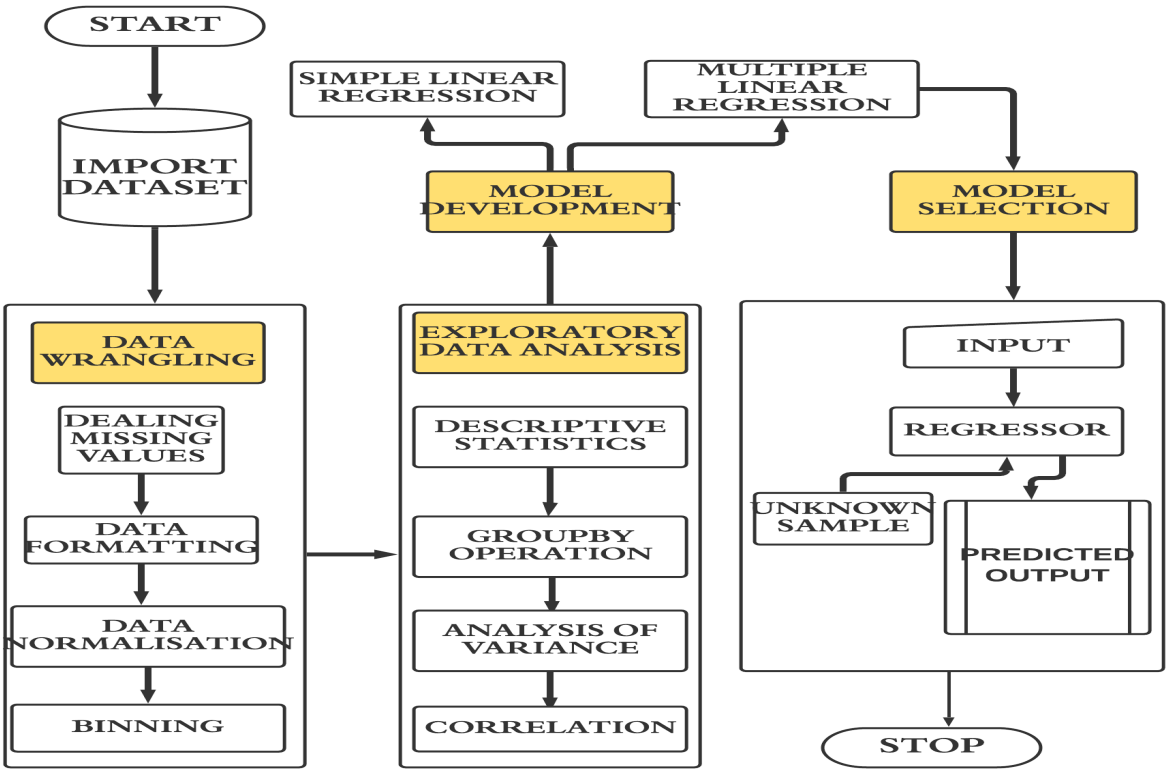


Figure 2.2

3) TESTING

SL No.	Input	Expected Outcome	Actual Outcome	Test Case
1.	0.048 463.28	-167.64005959	-134.112048	Pass
2.	0.056000 463.28094482	-167.64274655	-134.114198	Pass
3.	0.064000 463.27706909	-167.63696735	-134.109574	Pass

Figure 4.1)

4) IMPLEMENTATION DETAILS

1) Data Wrangling

STEP	**ALGORITHM:**
0	START
1	Import dataset
2	Read dataset
3	Create headers list
4	Replace headers and recheck dataset
5	Read datatype
6	Describe all columns in dataframe
7	Identify missing data df.isnull() 7.1) True -> Missing value 7.2) False-> Not missing value
8	Count missing value in each column
9	Replace missing data 9.1) Replace by mean 9.2) Replace by frequency 9.3) Drop row
10	Convert data type of missing data columns
11	Replace (original value) by (original value)/(maximum value) for data normalisation
12	numpy's linspace(start_value, end_value, numbers_generated function) for binning
13	Categorise data into 3 equal bins
14	Visualise bins using histogram
15	Save the new csv

2) Exploratory Data Analysis

STEP	**ALGORITHM:**

16 Import new dataset

17 Read dataset

18 Import Visualisation Packages

19 Plot linear relationship between variables

19.1) If scatter-plot is positive-> positive direct correlation

19.2) If scatter-plot is negative -> inverse/negative relationship

19.3) If scatter-plot is negative -> weak linear relationship

20 Visualise categorical variables using boxplot

20.1) If categories overlap-> not a good predictor

20.2) If categories are distinct -> good predictor

21 Calculate mean of data using groupby

22 Calculate Pearson Correlation of variables

22.1) 1: Total positive linear correlation.

22.2) 0: No linear correlation, the two variables most likely do not affect each other

22.3) -1: Total negative linear correlation

23 Calculate P-Value of variables

23.1) p-value is < 0.001: we say there is strong evidence that the correlation is significant.

23.2) the p-value is < 0.05: there is moderate evidence that the correlation is significant

23.2) the p-value is < 0.1: there is weak evidence that the correlation is significant

23.3) the p-value is > 0.1: there is no evidence that the correlation is significant

24 Calculation of function 'f_oneway' in the module 'stats' to obtain the F-test score and P-value of the grouped data.

24.1) Large score means large difference between means

3) Model Development

STEP ALGORITHM:

25 Load module for linear regression

26 Create Linear Regression Object

27 Fit Linear Model

28 Predict Output

29 Print Intercept

30 Print Slope

31 Fit Linear Model for multiple linear regression

32 Predict Output

33 Print Intercept

34 Print Slope

35 Import the visualization package: seaborn

36 Plot regression plot of simple linear regression

37 Plot residual plot of simple linear regression

38 Plot regression plot of multiple linear regression

39 Plot residual plot of multiple linear regression

40 Create function plotpolly(model, independent_variable, dependent_variable, Name):

40.1) x = dataframe1

40.2)y = dataframe2

40.3)fit polynomial polyfit()

40.4)display polynomial function

40.5)Import Pipeline module

40.6)Import StandardScaler in the pipeline

40.7)Calculate R^2 / R-squared of Simple Linear Regression

40.8)Calculate MSE of Simple Linear Regression

40.9)Calculate R^2 / R-squared of Multiple Linear Regression

40.10)Calculate MSE of Multiple Linear Regression

40.11)Determine a Good Model Fit with higher R^2 and smallest MSE value

4) Creation of model with GUI

STEP ALGORITHM:

40 Import necessary libraries

41 Import tkinter module

42 Import dataset

43 Create Pandas dataframe

44 X = input variables

45 Y = output variables

46 Splitting the Data set into Training Set and Test Set

47 Create instance of regression model

48 Fit model on training dataset

49 root = Create window for GUI

50 Canvas1 = Create canvas for GUI

51 Print intercept result in GUI

52 Print coefficient result in GUI

53 Label1 = Create label for 1st input in GUI

54 Label2 = Create label for 2nd input in GUI

55 Create function values ():

55.1) 1st input variable value1.get()

55.2) 2nd input variable value2.get()

55.3) Print prediction result

56 Create button to call function values

57 Plot 1st scatter graph

58 Plot 2nd scatter graph

59 Exit window

5)RESULTS AND DISCUSSIONS

1) Histograms representing Binned prices in Low, Medium, High

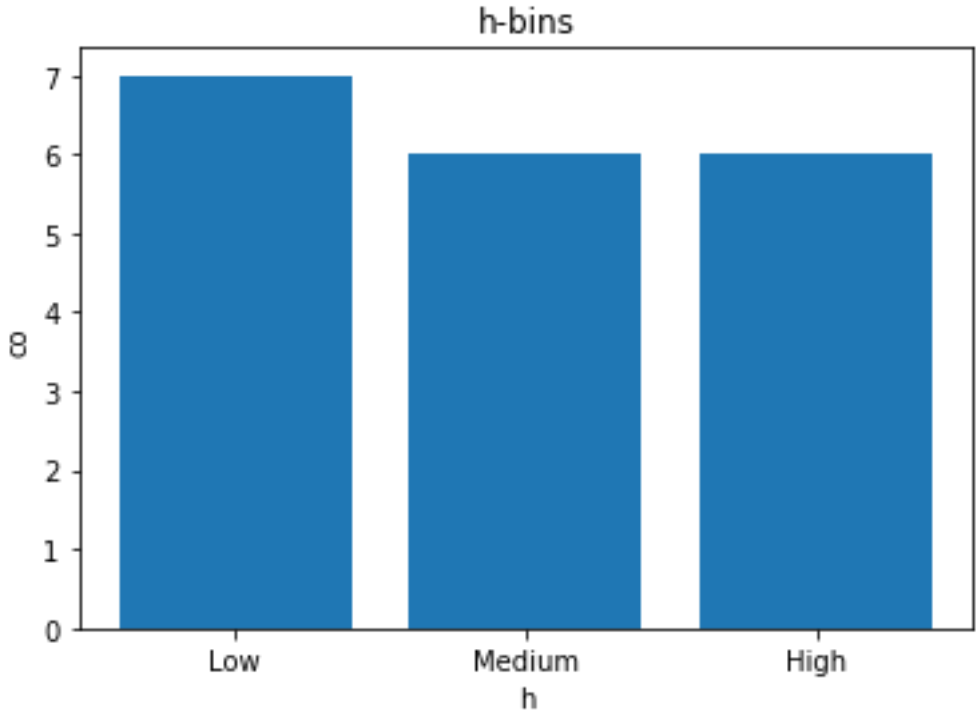


Figure 5.1)

2) Regression plot with ‘pr’ and ‘time’- Negative linear relationship

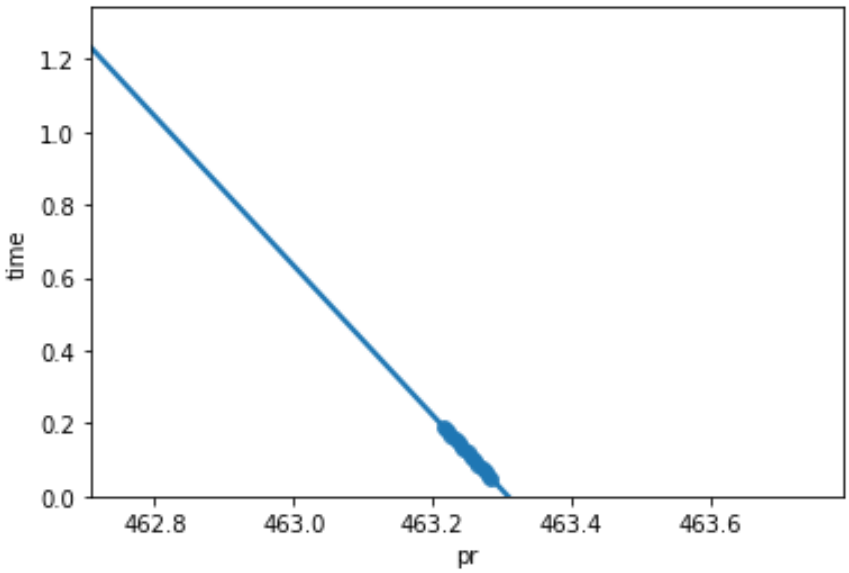


Figure 5.2)

3) Regression plot with ‘symboling’ and ‘time’- Positive linear relationship

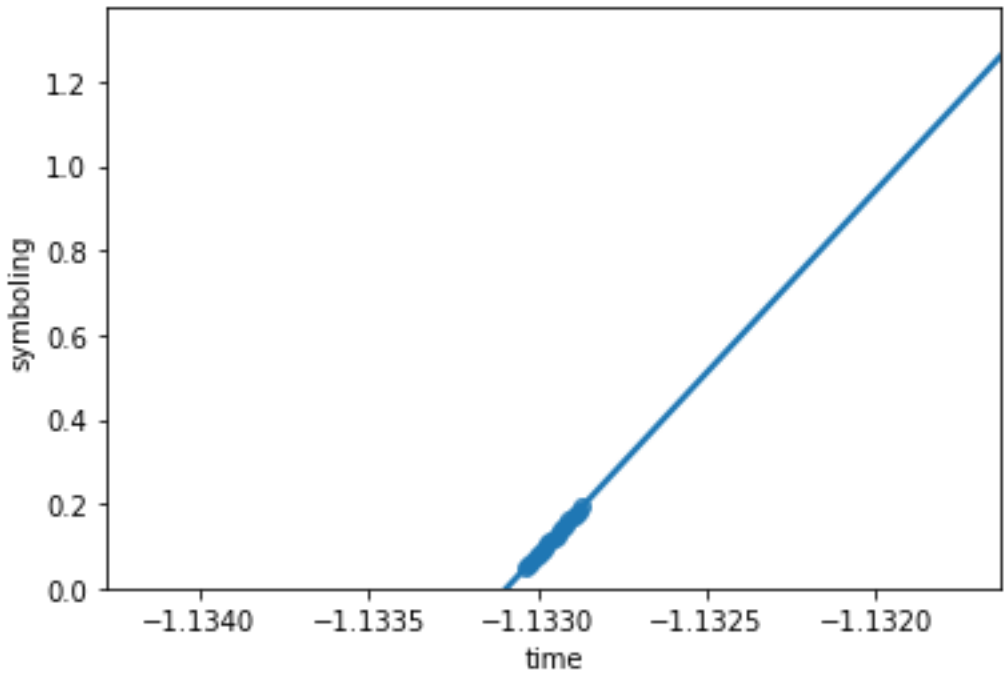


Figure 5.3)

4) Boxplot representing effect of h-binned with prices

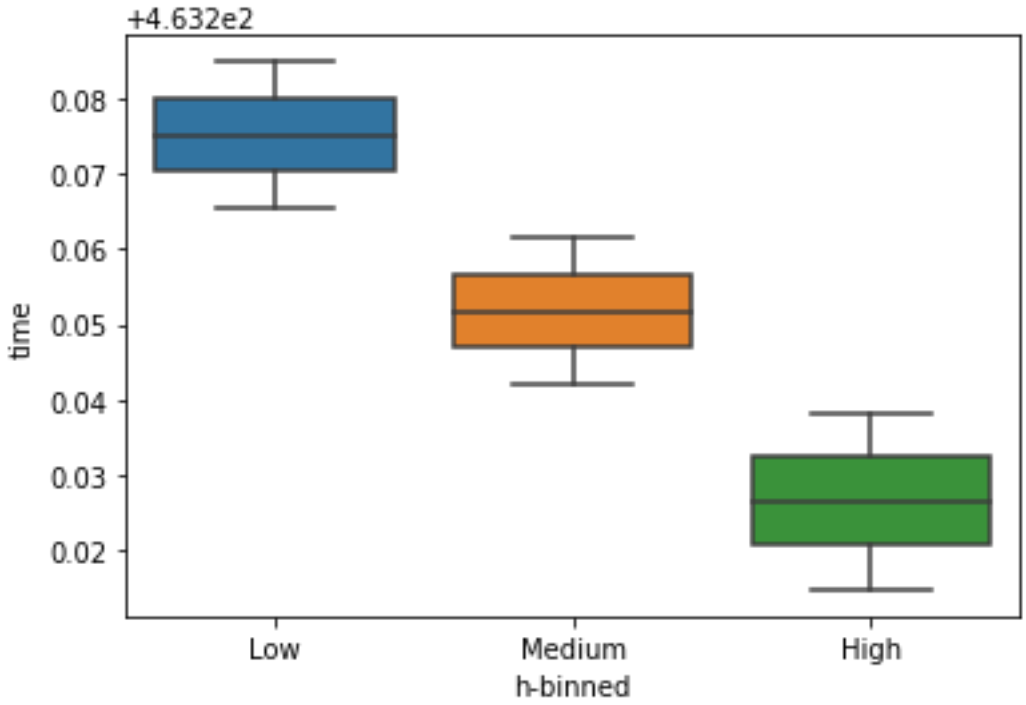


Figure 5.4)

5) Scatter plot for ‘Time’ over ‘Symboling’

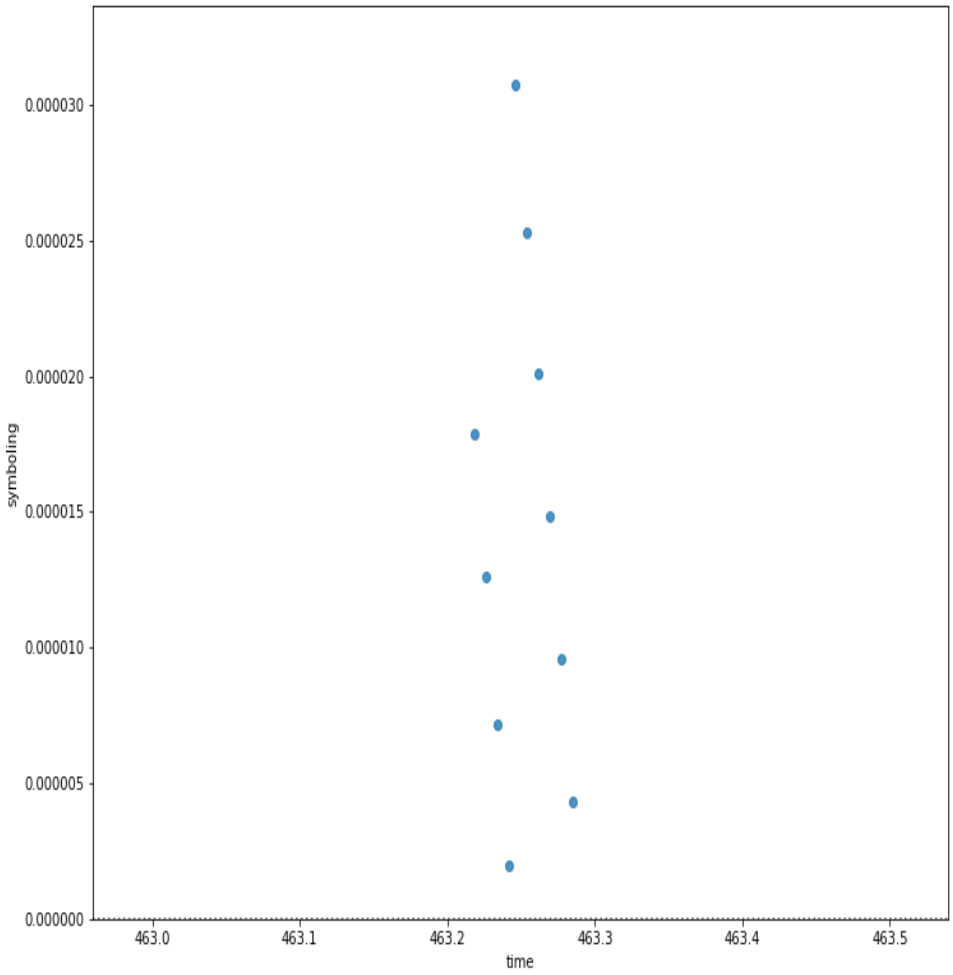


Figure 5.5)

6) Weak correlation between ‘time’ and ‘n-l’

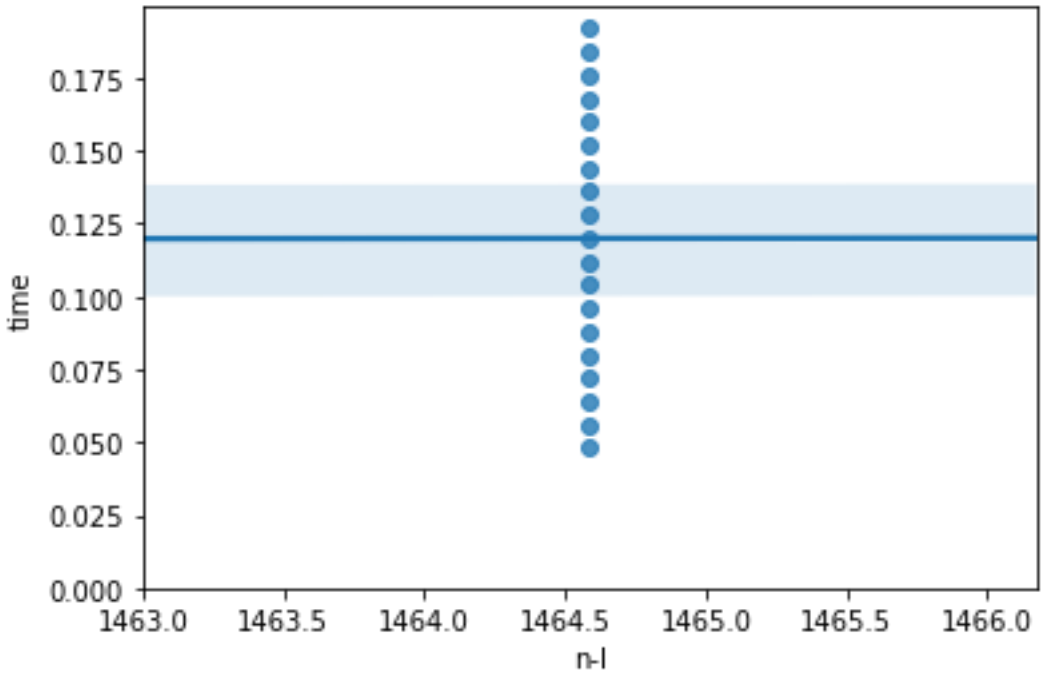


Figure 5.6)

7) Multiple Linear Regression Plot:

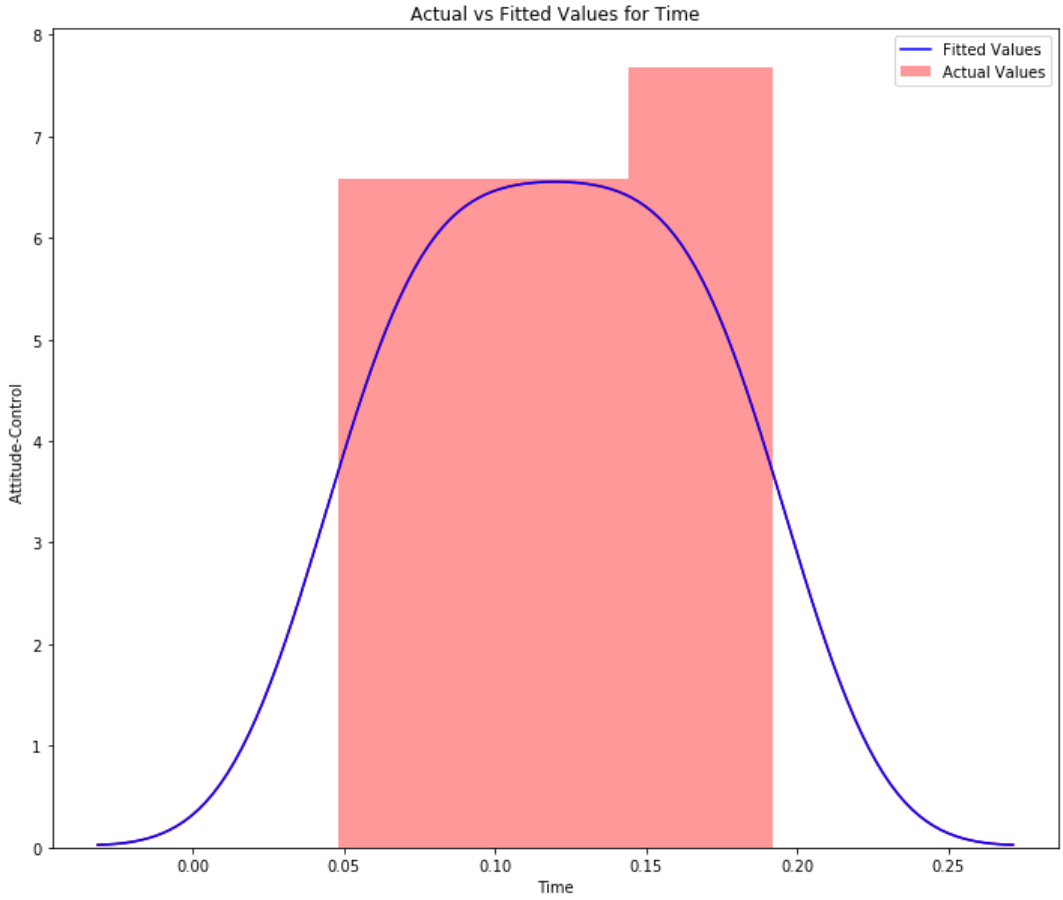


Figure 5.7)

8) GUI for multiple Linear Regression:

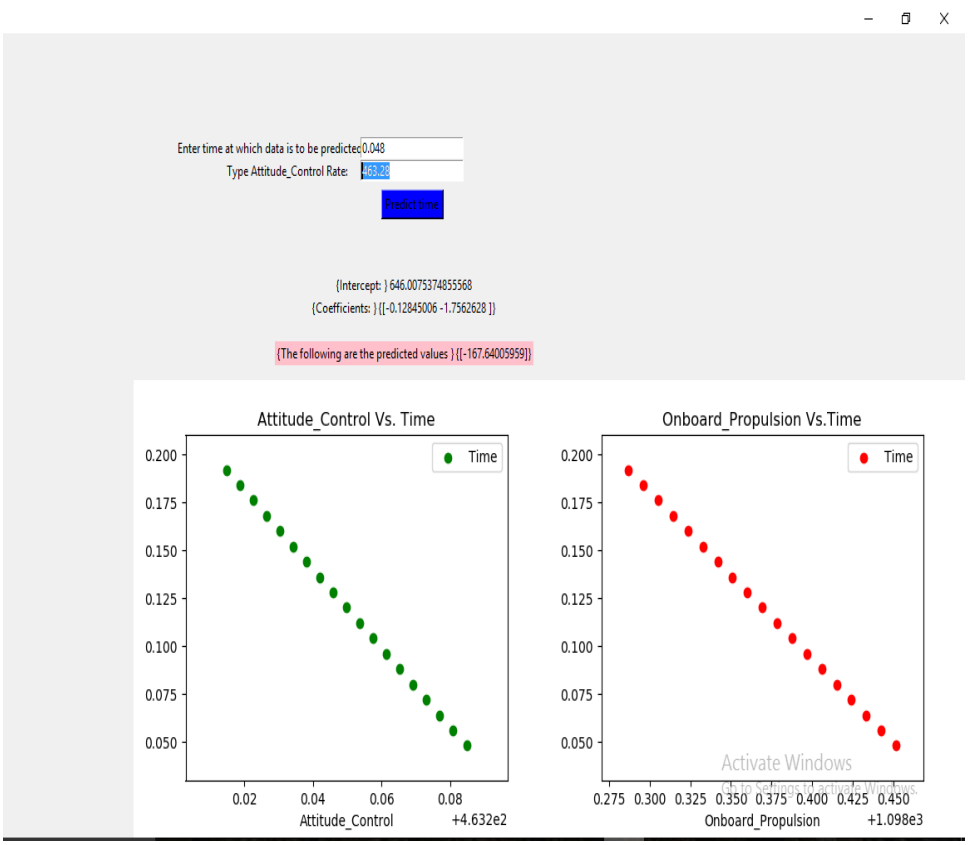


Figure 5.8)

6) SUMMARY AND CONCLUSION

The distribution plot of Linear Regression and Multiple Regression technique shows how the model predicts the time of descent of space craft based on several variables eligible to predict the time.

Comparing these models, we conclude that the MLR model is the best model to be able to predict time from our dataset. This result makes sense, since we have a lot of variables, and we know that more than one of those variables are potential predictors of the final time of descent.

6.1) SUMMARY OF ACHIEVEMENTS

The autonomous analysis of the dataset has been achieved in this project. Predictions can be done in a matter of seconds. Also, it is easy to use as a GUI has also been implemented for the ease of a user. This saves a lot of manual work and also it saves time from manual analysis.

6.2) MAIN DIFFICULTIES ENCOUNTERED

There were several difficulties encountered. Initially the structure of the dataset was very raw and was not understandable. It was already huge in size and on top of that very complex to understand. On the other hand, data analysis methods were to be researched to be used in the dataset. Creation of GUI in python was another difficult task.
These were tackled by transforming the data and cleansing the data. The main important characteristics of the dataset reduced the size of the dataset. Data analysis methods were researched and implemented on the data to derive insights. Tkinter helped in creation of GUI in python.

6.3) LIMITATION OF THE PROJECT

The limitations are:

- The prediction could be made more accurate with more available data. Due to unavoidable circumstaces, the entire dataset was not available.
- Objective of this analysis could be more than prediction. More advanced form of data analysis could be performed in this dataset to explore the data.

6.4) FUTURE SCOPE

- GUI can be advanced with more options. It could be made more attractive with more simplification for the user to be able to play with the data.
- This could be applied in trajectory analysis of Spacecraft for successful unmanned operations. Futuristic operations can be carried out with the continuation of this project.

6.5) SPECIAL OBSERVATIONS

The visualisation of the data to the user can be made more detailed for more research. This project has a lot of scope and can be useful to various sectors of the Space Department. It can also pave way for far more technologically advanced work.

7)GANTT CHART

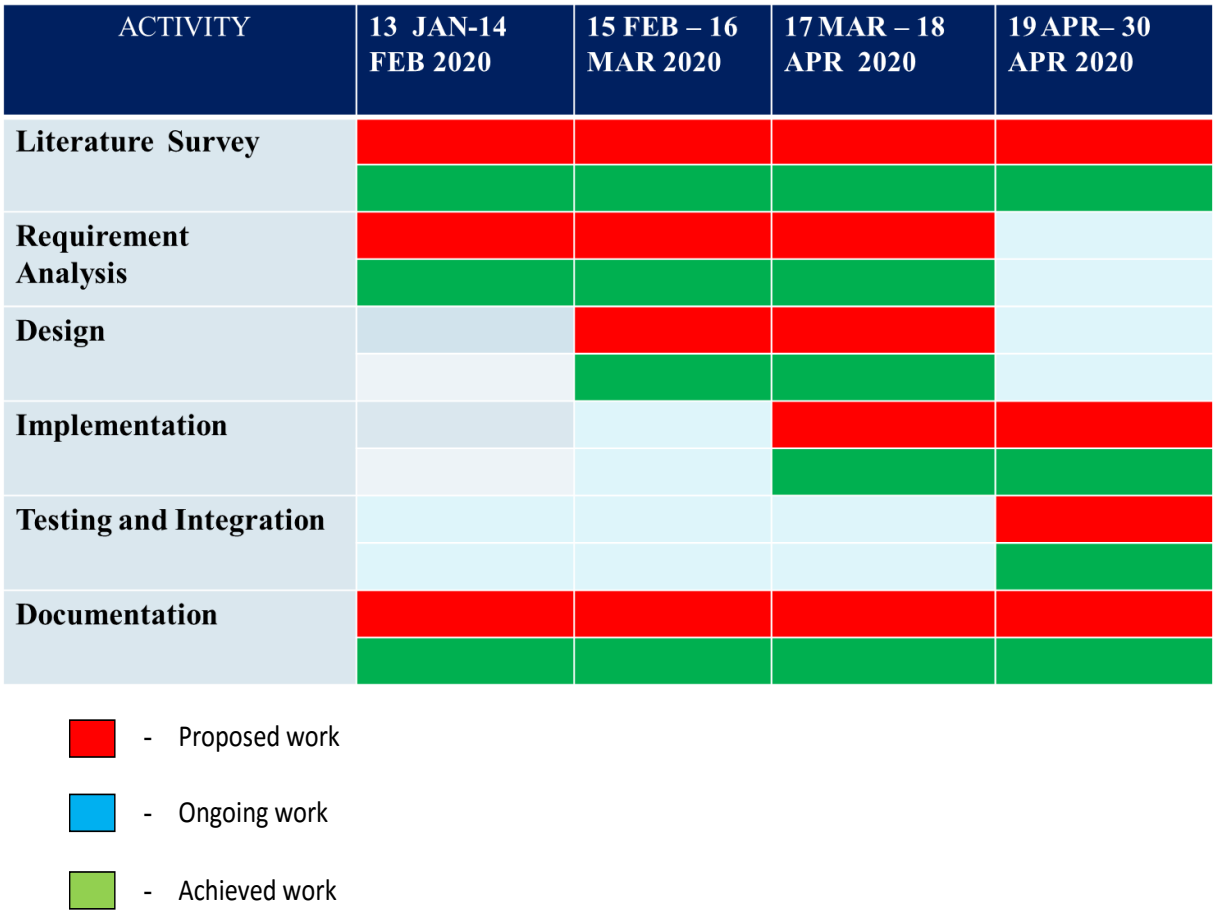


Figure 7.1)

8)REFERENCES

1) *Name-Onboard Computers, Onboard Software and Satellite Operations: an introduction , Author-Jens Eichkhoff [page no-3,4,5,6,21], Last Accessed Date: 14 Mar, 2020, Edition-Aug 2016,Publisher-Springer Science & Business Media*
2) *https://www.nasa.gov/sites/default/files/atoms/files/nac_tie_aug2018_tfong_tagged.pdf, Last Accessed Date: 12 Mar, 2020*
3) *https://www.coursera.org/learn/python/home/welcome, Last Accessed Date: 1 April, 2020*
4) *https://codingthesmartway.com/the-machine-learning-crash-course-part-2-linear-regression/ , Last Accessed Date: 15 April, 2020*
5) *https://towardsdatascience.com/summarizing-goodness-of-fit-6b11395c769a, Last Accessed Date: 18 April, 2020*
6) *https://dzone.com/articles/python-gui-examples-tkinter-tutorial-like-geeks, Last Accessed Date: 19 April, 2020*
7) *https://likegeeks.com/python-gui-examples-tkinter-tutorial/*
8) *https://www.coursera.org/learn/machine-learning-data-analysis, Last Accessed Date: 31 May, 2020*

THANK YOU