Group Id:-27

CHAPTER 1

Synopsis

1 Technical keywords

Data Science(DS), Data Analysis(DA), Machine Learning (ML), graphical user interface (GUI)

2 Project Definition

The concept of measuring happiness was introduced to the world by the government of Bhutan as Gross National Happiness which later was adapted by UN in the form of happiness index that measures the happiness and well- being of the population of the country. This analysis is very important for the specialists in the fields of financial aspects, brain research, overview examination, public insights, wellbeing, public arrangement, etc. This is very useful to evaluate the advancement of countries.

In this project, we are going to analyze the World and Indian happiness data using Python. We will answer the question- Are Indians Happy? We will discuss various parameters that contribute to providing happiness to the citizens like Log GDP per capita scores, Healthy life expectancy scores, Perceptions of corruption scores, Social support scores, Freedom to make life choices scores, and Generosity scores. We will also predict the future values of different parameters related to Happiness of a country. We also propose to generate the recommendations for India to make the country happier, in the ranking of Happiness-Index

3 Scope

Our project after completion can be applied to other fields such as

1. IN this project predict the future values of different parameters related to Happiness of acountry.
2. Generate the recommendations for India to make the country more happy in the rankingof Happiness-Index.
3. In this project the analysis of World Happiness data by extracting the knowledge about theHappiness-Index of a Country.

4 Objectives

The main objectives of the system can be given as

1. To perform the analysis of World Happiness data by extracting the knowledge about theHappiness-Index of a Country.
2. To predict the future values of different parameters related to Happiness of a country.
3. To generate the recommendations for India to make the country more happy in theranking of Happiness-Index.
4. To learn and implement the techniques of Multiple Regression.
5. To learn and implement the Naïve Bayes.

5 Review of Conference/Journal Papers and Relevant Theory

1. **Paper:-** A Comparative Analysis Of The Factors Affecting Happiness Index

   Year:-2020

   Author:- Parul Oberoi, Shalu Chopra, Yukti Sheth

   This paper proposes a study that has been carried out to find the reasons behind the country's low happiness index. The factors taken under study for this research are

   - Social Support
   - Freedom to make life choices
   - Physical well-being
   - Personal safety
   - Generosity

   The factor that was found to be most influential in determining the happiness of the selected sample through multiple regression was physical well-being and it was also observed that it was the same factor with the least average i.e., low value of physical well-being among the sample in area of study undertaken is causing low value of happiness index too.

2. **Paper: -** Future Prediction of World Countries Emotions Status to Understand Economic Status using Happiness Index and SVM Kernel

   Year:-2019

   Author:-B.Prashanthi, Dr. R. Ponnusamy

   In this paper, a supervised two-tier ensemble approach for predicting a country's BLI score was proposed. The work presented a cost-effective method of BLI prediction with a high degree of efficiency. The capability of the model to predict life satisfaction relied on the proper training features, chosen using a recursive elimination method with 10-fold cross-validation. The work combined three of the top four models, with simple averaging, to enhance the performance of the regression. This is forecasting the Better Life Index (BLI) score using machine learning based regression model that can influence the survival of future generations.

3. **Paper: -** Analyzing Happiness Index as a Measure Along With its Parameters and Strategies for Improving India's Rank in World Happiness Report

   Year: - 2019.

   Author: -sarah Ahtesham

   This paper elaborates the concept of Happiness Index as a measure and analyses various reasons for India to lose its position in the World Happiness Report. The author appropriately concludes the paper with suitable suggestions.

4. **Paper: -** A Data Analysis of the World Happiness Index and its Relation to the North-SouthDivide
   Year:- 2019
   Author:- Charles Alba

   In this paper, authors performed a detailed data analysis on the World Happiness Report with its relation to the socio-economic North-South Divide. In order to do so, they performed some extensive data cleaning and analysis before querying on the World Happiness Report. The results based on Hypothesis Testing determines the happiness of the Global North is greater than that of the Global South. Furthermore, queries presented in paper show that the mean happiness score for the Global North significantly outweighing that of the South. Likewise, the 10 'Happiest' nations all belong to the Global North whereas the 10 'least happy' nations belong to the Global South.

6 System Architecture

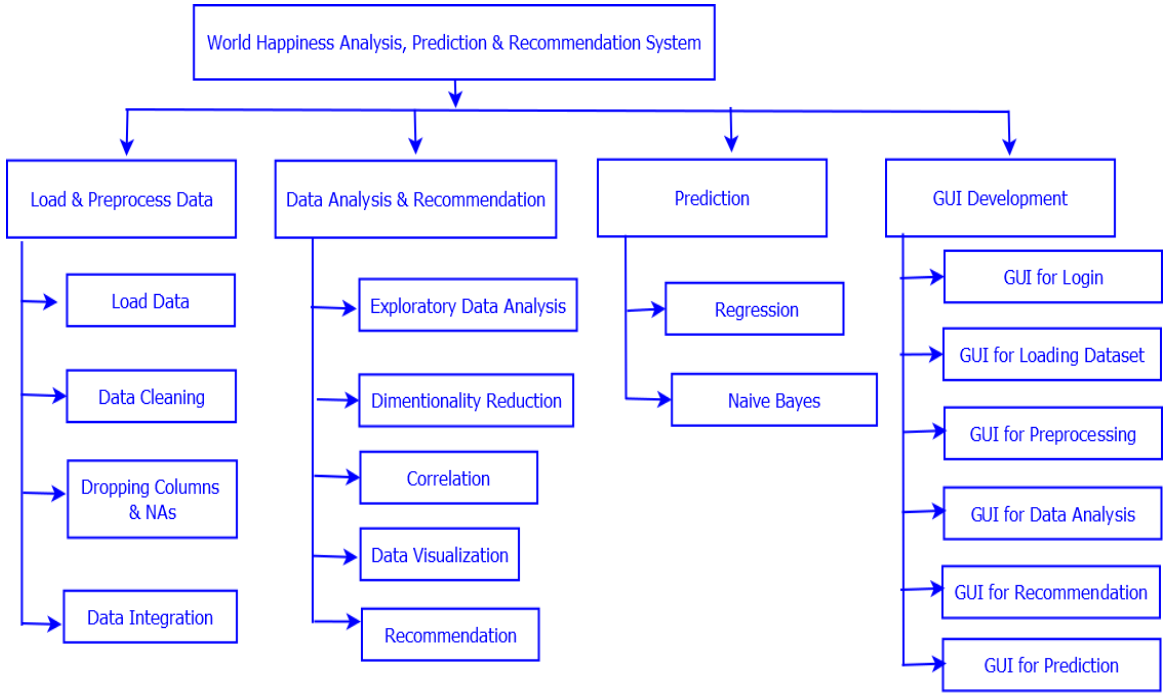

Figure 1: *System Architecture Diagram.*

7 Modules of Architecture

There modules in the project can be explained as

1. **Load & preprocess Data:-**

   1.1) Load Data

   This module will load the dataset from local computer to the system.

   1.2) Data Cleaning

   Many times the data is dirty. This module will preprocess the data towards removing the dirty components from the dataset.

   1.3) Dropping Columns & NAs

   In order to further process the data, the unnecessary columns will be dropped and NA entries will be handled by this module.

   1.4) Data Integration

   This module will generate new data from the existing data by merging and aggregating the existing data.

2. **Data Analysis & Recommendation** :-

This module will perform the detail analysis and derive many recommendations to the user.

2.1) Exploratory Data Analysis

The data explored to get the insights of data at abstract level.

2.2) Dimensionality Reduction

The dimensions of data are reduced to enhance the processing time and to focus the analysis on only the targeted data.

2.3) Correlation

The relationships between different parameters in the dataset, is analyzed using the method of correlation.

2.4) Data Visualization

This module display the graphical visualization of analysis and recommendation using the matplotlib & tkinter library.

2.5) Recommendation

Based on the analysis, this module recommends the conclusions based of certain predefined criterion.

3. **Prediction:-**

   This module will calculate the happiness index and predict the future values of different parameters which are important in the happiness index.

4. **GUI Development** :-

   In this module, GUIs for different functionalities will be developed using the tkinter library in Python.

8 Algorithm

8.1 Regression

Regression Algorithms will be used to predict certain parameters of happiness which are dependent on one or more other parameters.

A multiple regression considers the effect of more than one explanatory variable on some outcome of interest. It evaluates the relative effect of these explanatory, or independent, variables on the dependent variable when holding all the other variables in the model constant.

8.2 Naive Bayes

In order to get a baseline accuracy rate for our data, we are using a Naive Bayes classifier. Specifically, we used the scikit-learn implementation of Gaussian Naive Bayes. This is one of the simplest approaches to classification, in which a probabilistic approach is used, with the assumption that all features are conditionally independent given the class label. As with the other models, we used the Doc2Vec embeddings described above. The Naive Bayes Rule is based on the Bayes theorem:

$$p\left(\frac{c}{x}\right) = \frac{p\left(\frac{x}{c}\right).p(c)}{p(x)}$$

Parameter estimation for naive Bayes models uses the method of maximum likelihood. The advantage here is that it requires only a small amount of training data to estimate the parameters.

9 System Requirement

9.1 Hardware Requirements

1. i5 processor
2. 4GB RAM

9.2 Software Requirements

1. **Platform/Operating System :-** Windows 10
2. **IDE :-** PyCharm.
3. **Programming Language :-** Python 3
4. **Includeed Python Packages:-** Tkinter, matplotlib, Tensorflow, numpy, pandas etc.

10 Project Plan

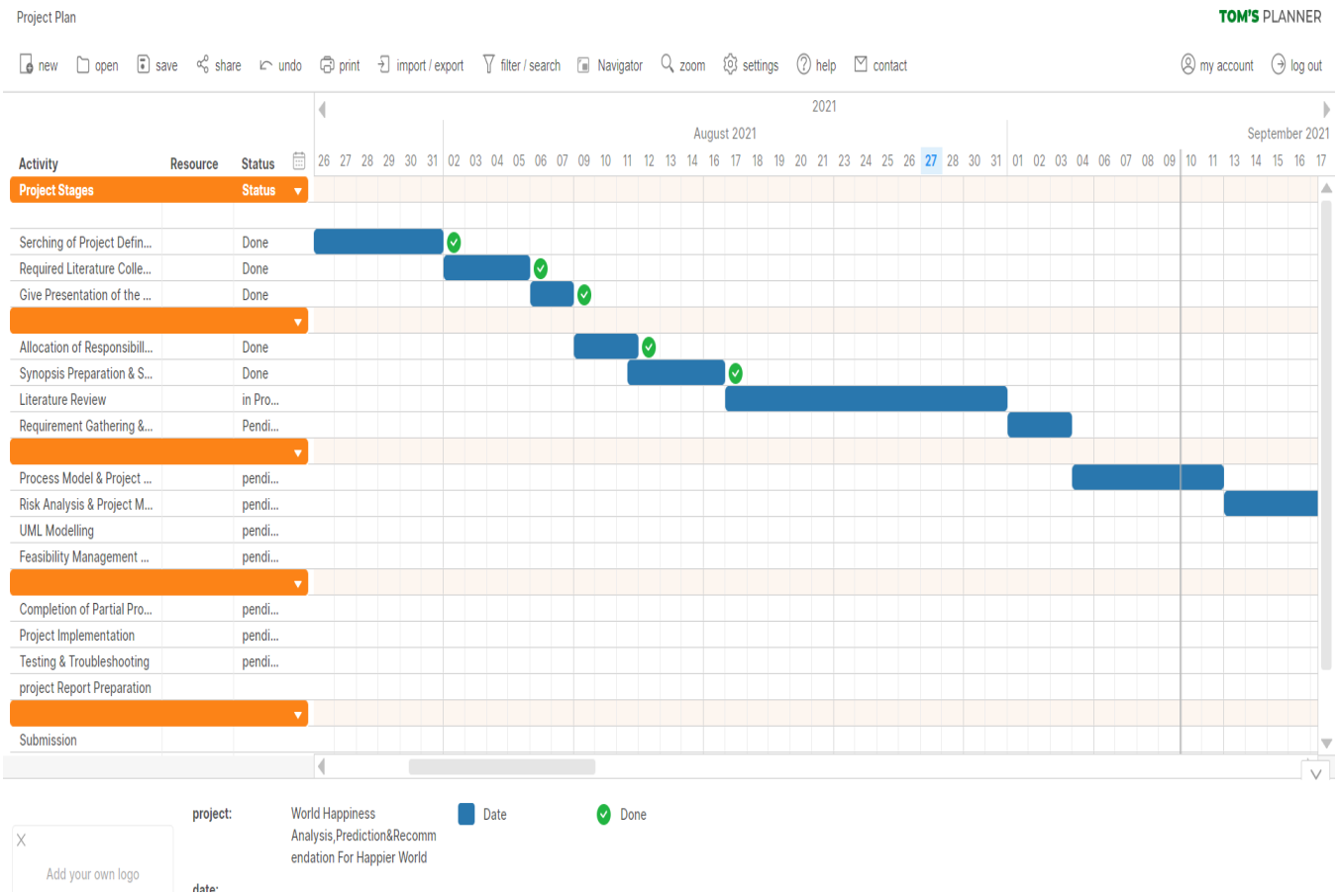

Figure 1: Project Plan (July – Aug)

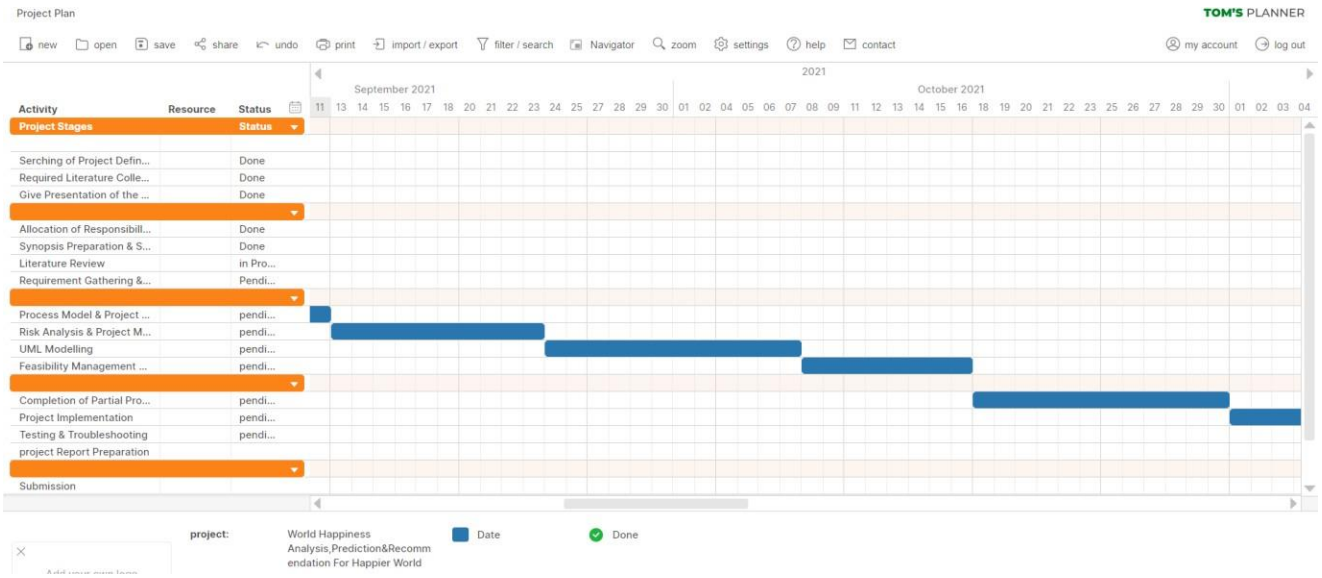

Figure 2: Project Plan (Sep - Oct)

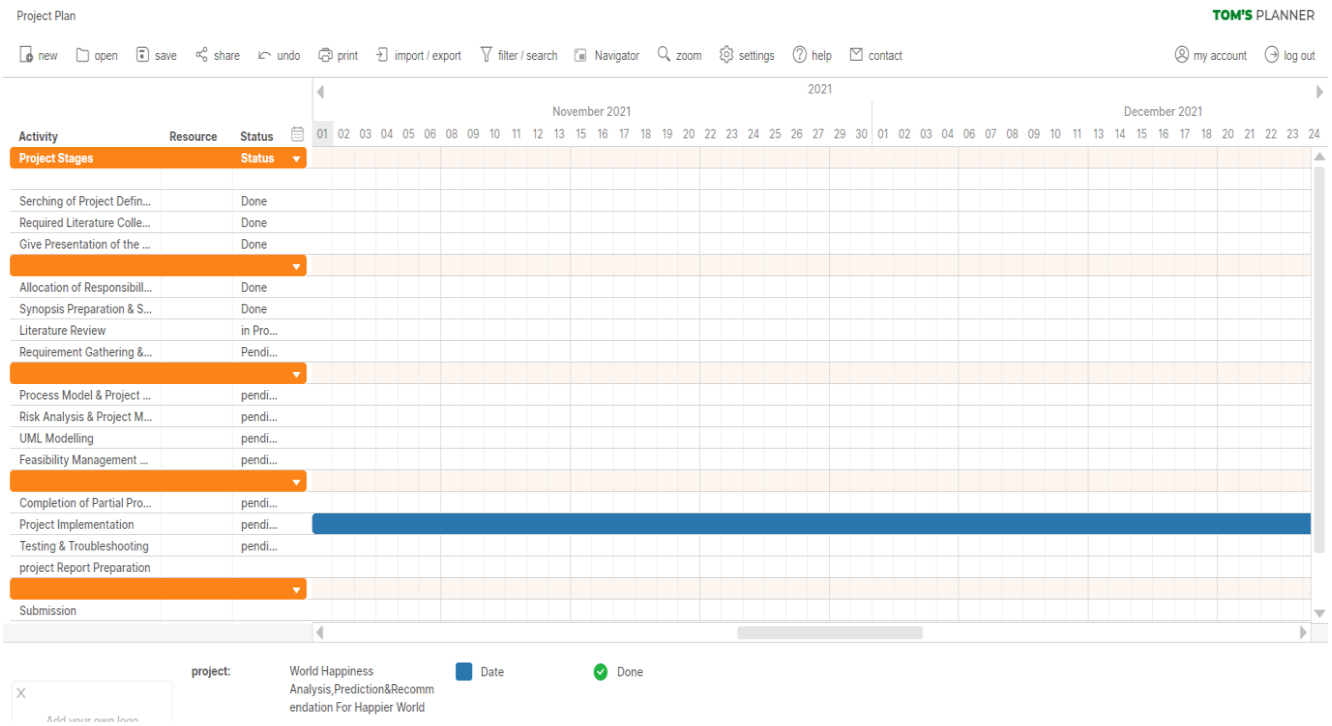

Figure 3: Project Plan (Nov - Dec)

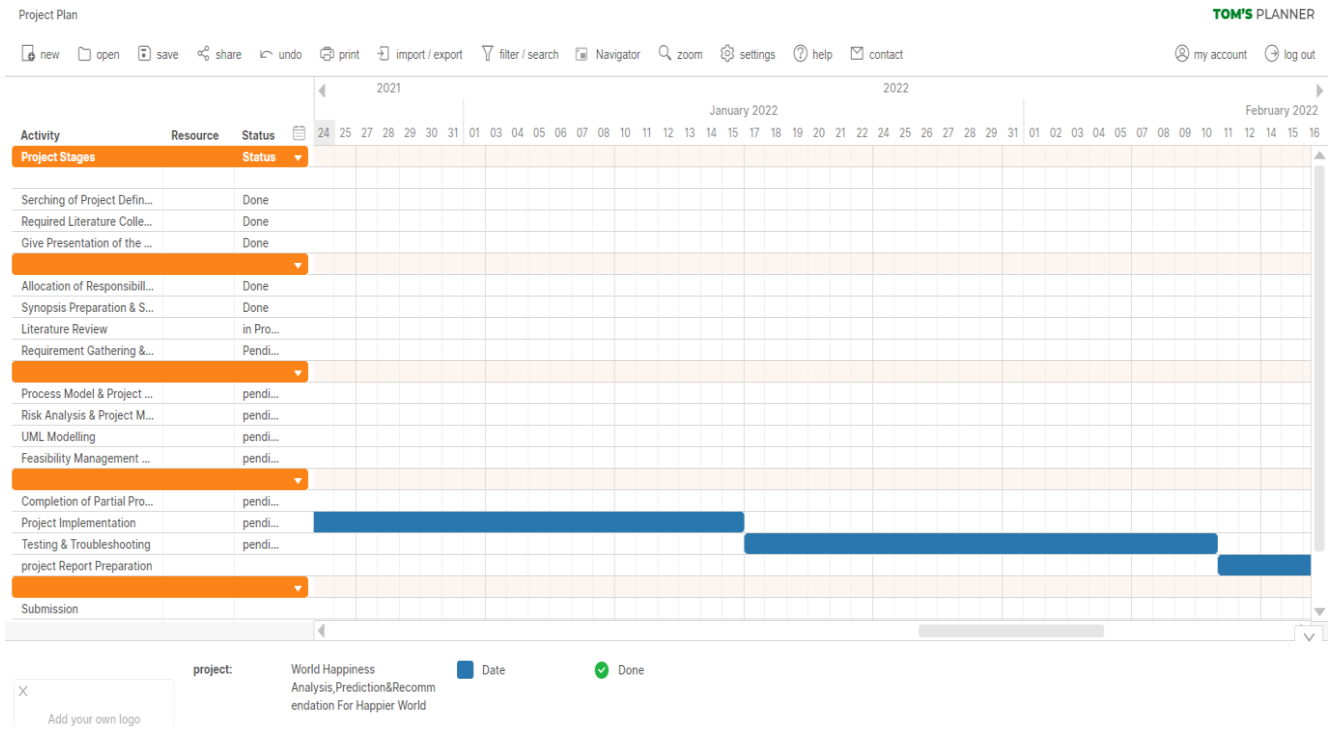

Figure 4: Project Plan (Dec - Jan)

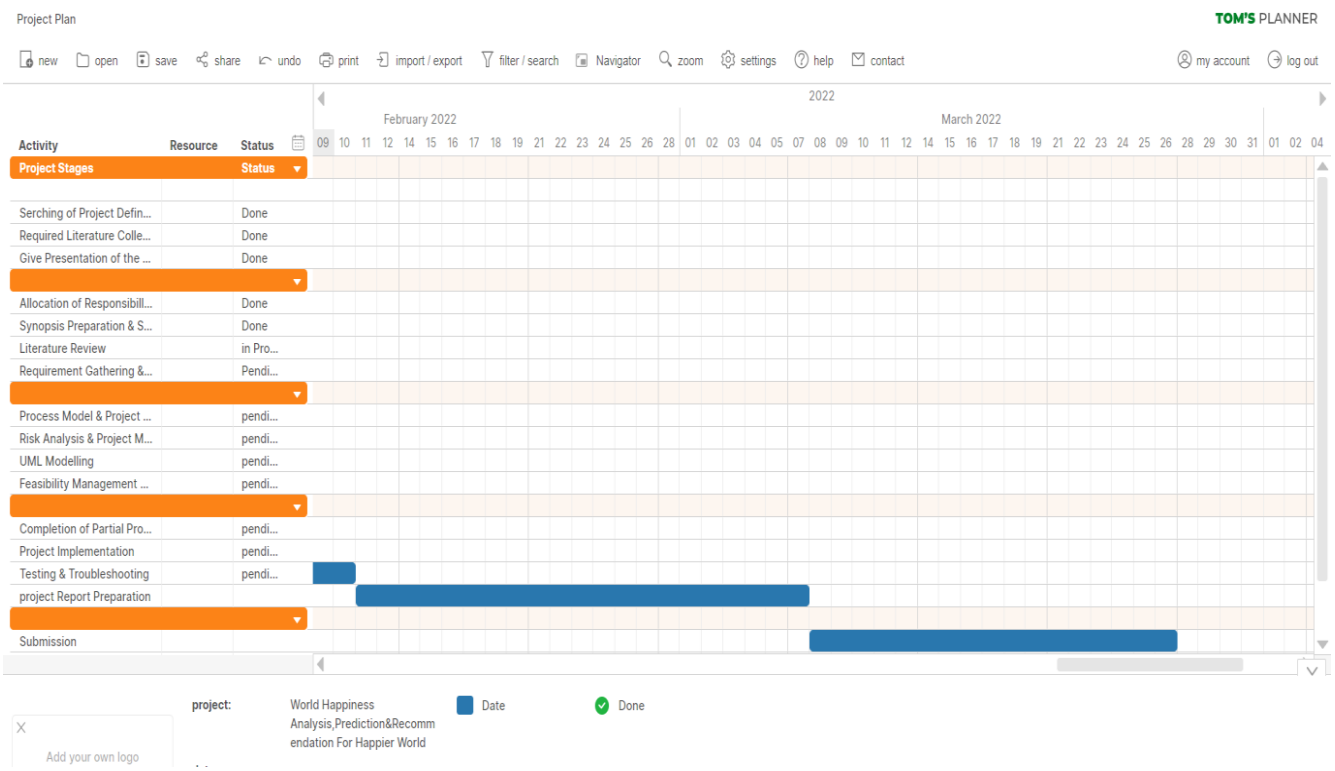

Figure 5: Project Plan(Feb - Mar)

11 Expected Outcomes

The system is a GUI based desktop applications that will run on desktop systems. The user will be able to upload data and analyze the different aspects of World Happiness. Also there will be features to draw the conclusions and thus recommendations from the system, which will be very useful to take strategic decisions for happier country. The systemcan also predict the future values of important parameters related to happiness.

12 References

1] Parul Oberoi, Shalu Chopra, Yukti Seth, "A Comparative Analysis Of The Factors Affecting Happiness Index", International Journal of Scientific & Technology Research Volume 9, Issue 03, March 2020.

2] B. Prashanthi, Dr. R. Ponnusamy, "Future Prediction of World Countries Emotions Status to Understand Economic Status using Happiness Index and SVM Kernel", International Research Journal of Engineering and Technology (IRJET), Volume: 06 Issue: 11, Nov 2019.

3] Sarah Ahtesham, "Analyzing Happiness Index as a Measure Along With its Parameters and Strategies for Improving India's Rank In World Happiness", 10.21917/ijms.2020.0161.

4] Charles Alba, "A Data Analysis of the World Happiness Index and Its Relation to the North-South Divide", Undergraduate Economic Review, Volume 16, Issue 1, Article 6, 2019.

5] Rajesh K Pillania, India Happiness Report 2020, Retrieved from http://pillania.org › wp-content › uploads › 2020/09.

6] World Happiness Report 2021 - Amazon AWS, https://happiness-report.s3.amazonaws.com

7] World Happiness Report – ResearchGate, https://www.researchgate.net › publication

Group Member Names

Sr.No.	Roll NO	Name Of Student	Signature
1.	87	More Amrat Vishwanath	
2.	95	Muthe Mayuri Madhukar	
3.	105	Pathan Tabassum Faruk	
4.	117	Shelar Prajakta Ishwar	

Prof .C.V.Longani
[Project Guide]

Dr.A.B.Pawar
[Project Coordinator]

Dr.D.B.Kshirsagar
[H.O.D Computer Dept]