* Project Summary

The project is titled "Projecting life expectancy using machine learning". The

skills required are Python,IBM cloud and Watson studio.Life expectancy is the key metric for assessing population health. Broader than the narrow metric of the infant and child mortality which focus solely at mortality at a young age, life expectancy captures the mortality along the entire life course. It tells us the average age of death in a population. Now this is an effort to link it with Machine

Learning. The document provided contains an array of countries with statistics

related to GDP, Diseases, Mortality etc.

* Project Requirements

This project requires a basic idea of what life expectancy is. It is going to estimate the life expectancy using the data given as an input to a machine

learning model. Coding knowledge of python is required for the same. In

2017, Indian life expectancy stands at 69.16 years. If the trends are true, it will

definitely decrease. It also requires a working knowledge of Python, a basic idea

how to use a cloud service. Watson studio can be accessed with the help of

tutorials provided by the website.

.

* Functional Requirements

1.Collect data.

2.Make a test and train set (basic strategy in creating a dataset)i.e the csv excel

file provided in references.

3.Use the data made available to public for health research

Collect enough data to make two different sections train and test data. The train

data is used to tell how exactly the model should respond to the input data. It is

drawn with different kinds of parameters acting on it. The test and train data

should not be same as it would be unfair and would lead to biasing. One good

effective strategy would be taking a part of train data and use it as test data.

4.The results can be tested by giving the appropriate inputs to the model(both

python based and using Auto AI).

* Problem and Solution

The problem is to find the life expectancy when some statistics are given

according to the data. The solution is to use an AutoAI model to predict the same.

* Advantages

The main advantage is the accuracy. I have achieved 80 percent accuracy by

using this model. There is also an UI created for anyone to use easily. It is done

by using a node red flow.

* Disadvantages

The disadvantage is that Country and Status columns are deleted because it is

not wise to use strings in node red . Instead a switch is used for Status field.

* Application

It can be used by the general public to find life expectancy to certain conditions

applied.

* Flowchart

The following steps where done to make this model in the same order

1.Create a cloud account.

2.Create services required(Watson, Node Red)

3.Data refining

4.AutoAI model

5.Model Deployment and testing

6.Linking the model to Node Red

7.Finally, using the UI

* Technical Requirements

1.The databases in csv format

2.Data science can be used to implement the project. A dataframe can be

created and python language can be used to edit, create, delete or format. It

should be kept in mind that python is a crucial tool in this regard. The data is

already provided in the form of a csv file.

* Software Requirements

The software requirements suggested by the organization is

1.Python IDE

2.Excel

3.IBM Cloud

4.IBM Watson Studio

5.Machine Learning Instance

6.Node Red Instance

* Project Deliverables

The end product would be a code , data set and a working Machine Learning model to predict life expectancy with the following inputs.

1.Year

2.Status(Developing or developed)

3.Adult Mortality

4.Infant Deaths

5.Alcohol

6.Percentage Expenditure

7.Hepatitis B

8.Measles

9.Under five deaths

10.Polio

11.Total Expenditure

12.Diphtheria

13.HIV/AIDS

14.GDP

15.Population

16.Thinness in 1-19years

17.Thinness in 5-9years

18.Income composition of resources

19.Schooling

LIfe expectancy is predicted using all these fields provided by the user.

* Project Team

This project is to be done individually so I make up the team. I have created Slack, Github, IBM cloud accounts and intend to make efficient use of them .

Additionally I have posted by code on Git repository.

* Project Schedule

I was assigned this project on 15th of May,2020. It is mentioned that a month

time is given to complete the same.

* Conclusion

In the end, the product is a finished ML model which can predict the life

expectancy as close as possible to the given data.

* Source Code

The corresponding source code is provided in the github repositary.

* Bibliography

The references which were suggested to us in the smartbridge workspace were

used . Youtube vides and tutorials were provided for each and very task.