Name: M. Shebana

Mail Id: [shebanayaseen@gmail.com](mailto:shebanayaseen@gmail.com)

Academic Mail Id: [SI05202000188@smartinternz.com](mailto:SI05202000188@smartinternz.com)

Internship Title: Predicting Life Expectancy using Machine Learning - SB32873

Project ID: SPS\_PRO\_215

Project Title: Predicting Life Expectancy using Machine Learning

Duration: 30 Days

**PROJECT DESCRIPTION**

A typical Regression Machine Learning project leverages historical data to predict insights into the future. This problem statement is aimed at predicting Life Expectancy rate of a country given various features.

Life expectancy is a statistical measure of the average time a human being is expected to live, Life expectancy depends on various factors: Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors. This problem statement provides a way to predict average life expectancy of people living in a country when various factors such as year, GDP, education, alcohol intake of people in the country, expenditure on healthcare system and some specific disease related deaths that happened in the country are given.

**PROJECT REPORT**

**INTRODUCTION:**

Since ancient times, there are a lot of change in the behaviours and cultures of people in different places. According to their way of living, the health care and life expectancy of people varies among each other. These differences are may be based on various factors such as Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors.

OVERVIEW:

Life expectancy is a statistical measure of the average time a human being is expected to live. A typical Regression Machine Learning project leverages historical data to predict insights into the future. This problem statement is aimed at predicting Life Expectancy rate of a country given various features.

This problem statement provides a way to predict average life expectancy of people living in a country when various factors such as year, GDP, education, alcohol intake of people in the country, expenditure on healthcare system and some specific disease related deaths that happened in the country are given in a dataset.

PURPOSE:

The purpose of this project was the people from various places can easily predict their life expectancy by providing the inputs asked by the model. This software can be used by all people in the world because the training part of this model contains an inputs and predictions of more number of countries.

**LITERATURE SURVEY:**

There are so many organizations that are making research in the prediction of life expectancy. Many research papers dealing with the creation of this model under many algorithms such as Machine Learning, Deep learning and programming languages such as Python and Java script.

DATASET:

The data set for this model creation has been obtained from [www.kaggle.com](http://www.kaggle.com).There are totally 194 countries including both islands and territories. In this data set, they collected the factors such as Country, Year, Status, Life Expectancy, Adult Mortality, Infant Deaths, Alcohol, Percentage Expenditure, Hepatitis B, Measles, BMI, Under-Five Deaths, Polio, Total Expenditure, Diphtheria, HIV/AIDS, GDP, Population, Thinness 1-19 Years, Thinness 5-9 Years, Income Composition Of Resources, Schooling.

The time period of this all data collection was from 2000-2015. Totally 16 years data sets are available in this file. It also classifies the countries as developing and developed sectors. In some countries, the life expectancy level was 40 but in some countries the life expectancy rate was high as 80.

PROPOSED SOLUTION:

So many people were expecting to use a model of life expectancy prediction. In order to that, many institutions and companies are leading their team to build that model. In my project, I have proposed a solution to predict the life expectancy using machine learning. Machine Learning is the process of training the computer to think and decide solutions like human. The reason why I have chosen this architecture was only with the help of Machine Learning, deep understanding of the data and an ability to create a model can be done.

**PROCEDURE OF WORK:**

1. PROJECT PLANNING AND KICKOFF:

1. Understanding the project description and analyze the data and attributes in the given dataset.
2. Creating Github account-https://github.com/Shebana
3. Installing Slack and create account with the mail id [shebanayaseen@gmail.com](mailto:shebanayaseen@gmail.com)
4. Learning to use Zoho writer.

2. EXPLORE IBM CLOUD PLATFORM:

1. Creating IBM cloud account with the mail id [shebanayaseen@gmail.com](mailto:shebanayaseen@gmail.com)
2. Creating IBM academic initiative account with the mail id [si05202000188@gmail.com](mailto:si05202000188@gmail.com)
3. Create a Node-Red starter application.

3. EXPLORE IBM WATSON SERVICES:

1. Exploring IBM Watson use cases.
2. Learning about IBM Watson Machine Learning.

4. INTRODUCTION TO WATSON STUDIO:

1. Learning to build own Machine Learning model using IBM Watson.
2. Automate the Machine Learning Model

5. PREDICTING LIFE EXPECTANCY WITH PYTHON:

1. Collecting Data set from [www.kaggle.com](http://www.kaggle.com)
2. Creating IBM Watson services
3. Create a jupyter notebook and import data from Object storage.

6. PREDICTING LIFE EXPECTANCY WITHOUT PYTHON:

1. Created Node-Red model and integrated with Machine Learning model.

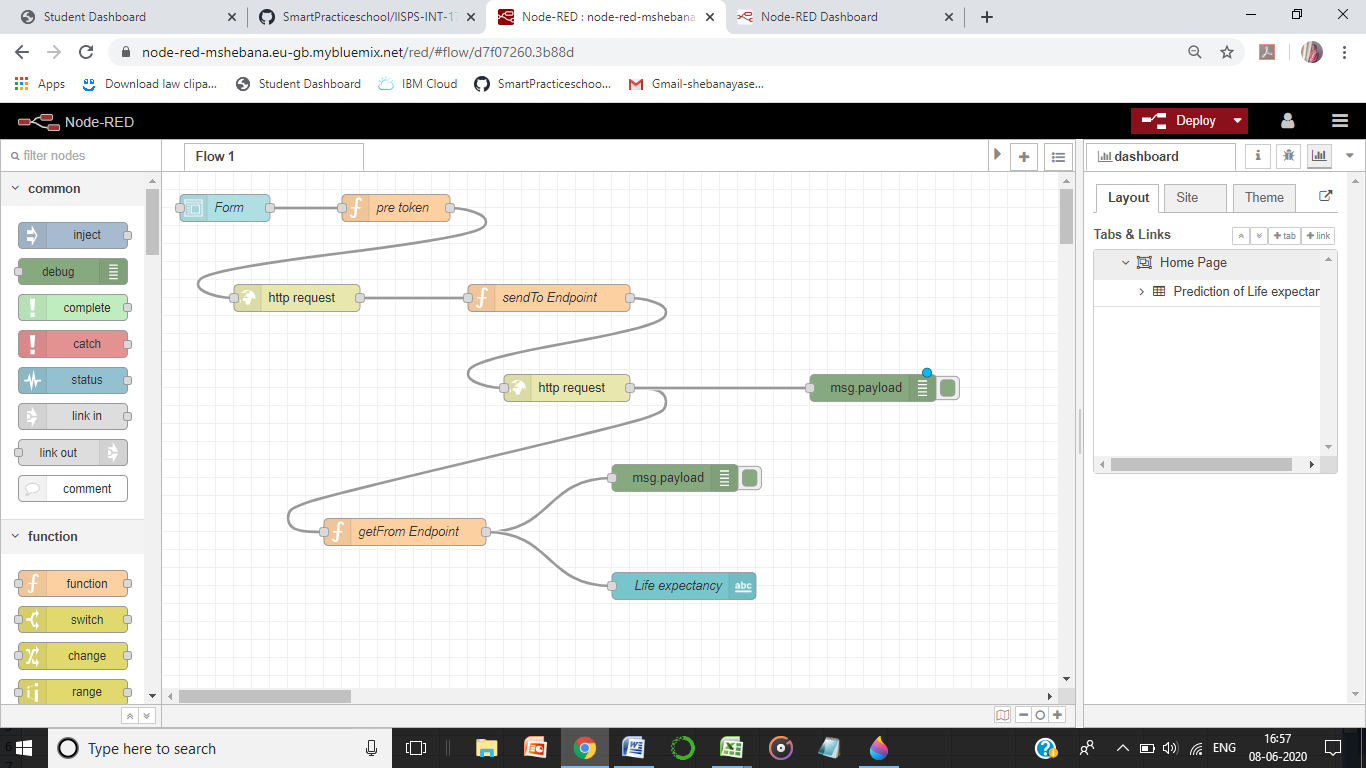
**EXPERIMENTAL EXPLANATION:**

1. JUPYTER NOTEBOOK:

<https://github.com/SmartPracticeschool/llSPS-INT-1710-Predicting-Life-Expectancy-using-Machine-Learning/blob/master/LifeExpectancyPrediction.ipynb>

2. NODE-RED APPLICATION:

<https://github.com/SmartPracticeschool/llSPS-INT-1710-Predicting-Life-Expectancy-using-Machine-Learning/blob/master/Node%20red%20Application-Life%20Expectancy.html>

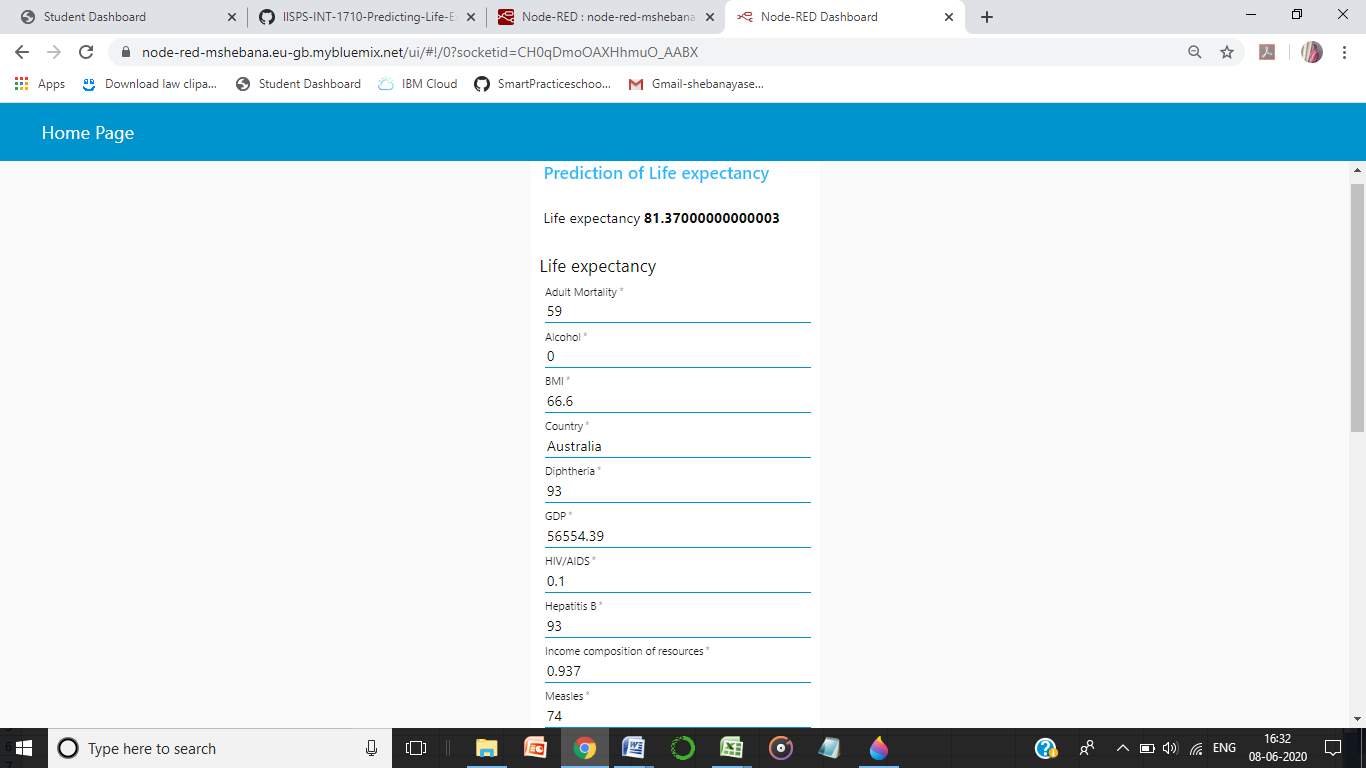


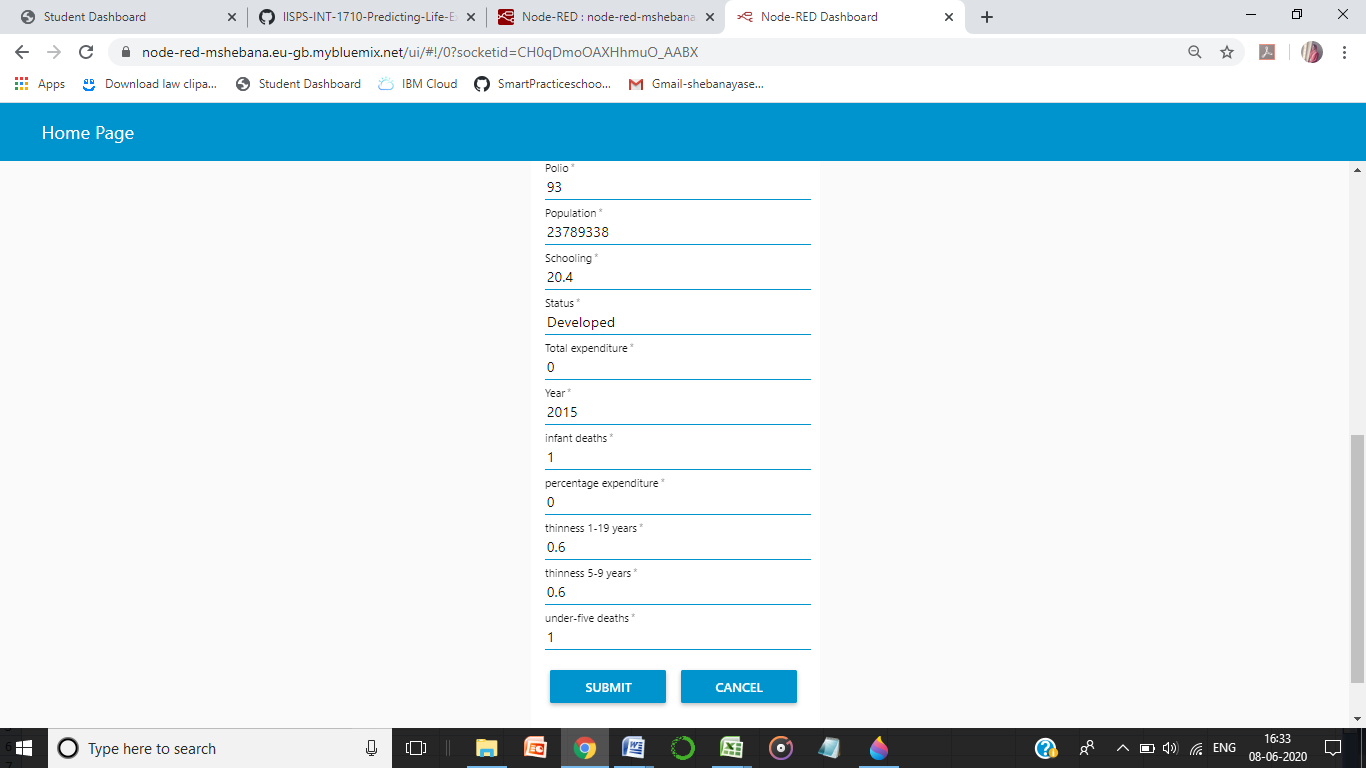
3. LIFE EXPECTANCY DATA SET:

<https://github.com/SmartPracticeschool/llSPS-INT-1710-Predicting-Life-Expectancy-using-Machine-Learning/blob/master/Life%20Expectancy%20Data.csv>

RESULT:

By using IBM and smart bridge platform, the model for predicting Life expectancy in different places has been created.





While giving the inputs for the country Australia on the year 2015, the life expectancy value 81.370 has been predicted.

CONCLUSION:

* This model can be used to predict the life expectancy of people in different places.
* This model contains various factors such as Country, Year, Status, Life Expectancy, Adult Mortality, Infant Deaths, Alcohol, Percentage Expenditure, Hepatitis B, Measles, BMI, Under-Five Deaths, Polio, Total Expenditure, Diphtheria, HIV/AIDS, GDP, Population, Thinness 1-19 Years, Thinness 5-9 Years, Income Composition Of Resources, Schooling.
* With the help of all these input values, the model will predict the life expectancy of such people.
* The accuracy level of prediction in my model is more than 95%.
* From the help of this model, the life expectancies of more than 190 countries can be detected.

REFERENCES:

1. [Predicting Life Expectancy Using Machine Learning](http://ls00012.mah.se/handle/2043/26877)
2. [Performance evaluation of machine learning algorithms in post-operative life expectancy in the lung cancer patients](https://arxiv.org/abs/1504.04646)
3. [Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients](https://www.sciencedirect.com/science/article/pii/S1568494613002627)
4. [Improved prediction of post-operative life expectancy after Thoracic Surgery](https://ijassa.ipu.ru/ojs/ijassa/article/view/351)
5. [Predicting life expectancy in heart failure](https://jamanetwork.com/journals/jama/article-abstract/181998)