Predicting Life Expectancy using Machine Learning – SB51612
Project Scope Document

1. Project Summary

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.

Also, some of the past research was done considering multiple linear regression based on data set of **one year** for all the countries. But this dataset provided by data from WHO (https://www.kaggle.com/kumarajarshi/life-expectancy-who) considers a timeframe from 2000-2015 for 193 countries.

2. Project Requirements:
2.1 Functional Requirements:
Predicting the life expectancy rate of a country by considering the impact of GDP, education, alcohol intake, expenditure on healthcare etc.
2.2 Technical Requirements:
Python, IBM Cloud, IBM Watson Studio, Node-RED

Hardware Requirements:
Processor-i3 7th generation or higher
Speed: 2 GHz or more

3. Research of Previous:

In building this project we use the data from Kaggle. Which helps to build the model.

4. Algorithms:

Using Watson studio with Juyter notebook import python library and helper function.
from sklearn.linear_model import LinearRegression, Lasso, Ridge, etc.
Create a data frame.
Checking missing data then normalizing it.
Create Training and Test Sets.
Extract just the numeric values for the features and labels as the TensorFlow model will expect just numeric values as input.
Create a model then train it.
Prediction of life expectancy.

5. Deliverable

A machine learning model that will predict life expectancy(software).

6. Out of Scope

In the project, the user will not able to modify or not able to increase the accuracy of the ML model.

7. Project Team

- Shubham (Individual)

8. Project Schedule

Task	Days
1. Git 2. Zoho Writer 3. IBM Cloud Services 4. Watson Studio 5. Node-RED	0.5 day (20/05/2020) 0.5 day (20/05/2020) 1 day (21/05/2020) 3 days (22/05/2020) 2 days (25/05/2020)
1. Documentation 2. Data Preprocessing 3. ML Algorithms 4. Node-RED Integration	1 day (27/05/2020) 2 days (29/05/2020) 3 days (2/06/2020) 1 day (06/06/2020)