Predicting life expectancy using machine learning

● Project Summary

In this project, we have to create a new model based on the data provided to evaluate the life expectancy

The data offers a timeframe from 2015 to 2022. The output algorithms have been used to test if they can maintain their accuracy in predicting the life expectancy for data they haven't been trained. Following algorithms have been used:

1. Linear Regression
2. Ridge Regression
3. Lasso Regression
4. Elastic Net Regression
5. Linear Regression with Polynomic features
6. Decision Tree Regression
7. Random Forest Regression

● Project Requirement

The scope of this project is to predict the life expectancy by creating and training the model based on the provided dataset. The project relies on accuracy of data. Prediction will be based on the factors like Immunization related factors, Mortality factors, Economical factors and Social factors. The Global Health Observatory (GHO) data repository under World Health Organization (WHO) keeps track of the health status as well as many other related factors for all countries. The data-sets are made available to public for the purpose of health data analysis. The data-set related to life expectancy, health factors for 193 countries has been collected from the same WHO data repository website and its corresponding economic data was collected from United Nation website. Among all categories of health-related factors only those critical factors were chosen which are more representative. It has been observed that in the past 15 years, there has been a huge development in health sector resulting in improvement of human mortality rates especially in the developing nations in comparison to the past 30 years. Therefore, in this project we have considered data from year 2000-2015 for 193 countries.

● Functional Requirements:

1. A data model based on dataset.
2. A GUI (graphical user interface) or a dashboard to predict the life expectancy easily.

● **Technical Requirements:**

1.Data model created must have at least 90% accuracy.

2. We can use datasets with the help of machine learning and data science with the help of python.

3. The prediction of life expectancy must be in years.

● **Software Requirements:**

Python IDE, Excel, IBM Cloud, IBM Watson

● **Project Deliverables:**

1. A Machine learning model to predict life expectancy.
2. A user interface to interact with the model.

● **Project Team:**

This Project contain a team of one member.

Name: Rohit Tiwari

● **Project Schedule:**

Sr.no	Topic	Sub-Topic	Week	Date
1.	Project Planning & Kick-off	i. Project Scope, Schedule, Team & Deliverables. ii. Setup the Development Environment	WEEK-1	15/05/2020 To 22/05/2020
2.	Explore IBM Cloud Platform	i. Create IBM Cloud Account. ii. Create A Node-RED Starter Application	WEEK-2	22/05/2020 To 29/05/2020

3.	Explore IBM Watson Services	i. Explore IBM Watson Use cases ii. Explore IBM Watson Machine Learning	WEEK-3	29/05/2020 To 05/06/2020
4.	Introduction to Watson Studio	i. Build Your Own ML Model in IBM Watson Studio ii. Automate Your ML Model	WEEK-3	29/05/2020 To 05/06/2020
5.	Predicting Life Expectancy with Python	i. Collect the Dataset for The Project ii. Create Necessary IBM Cloud Services iii. Create A Watson Studio Project iv. Create Machine Learning Service v. Create A Jupyter Notebook In IBM Watson And Import Data vi. Build A Machine Learning Model and Create Endpoints for Node-RED Integration vii. Build Node-RED Flow to Integrate ML Services	WEEK-4	05/06/2020 To 14/06/2020