PREDICTING LIFE EXPECTANCY

PROJECT MANAGER- Gyanesh Sharma **DATE -** 12/06/20

PROBLEM STATEMENT :

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 SUMMARY :

The aim of this project is to predict life expectancy of people in a country. The model is based on supervised learning approach which uses
several factors related to the person and the country they reside.
Some factors are: illness, age, alcohol intake, financial status of the person also the healthcare system and facilities provided by the country, immunization related factors, economical factors, history of disease in the country and many other factors. Also to increase accuracy, countries are one hot encoded as it also plays a major role in predicting life expectancy. For training model Random Forest Regression is used as after evaluating various regression algorithm it was giving best result with minimum rse error.

PROJECT REQUIREMENTS : An appropriate dataset is needed to build the model.

FUNCTIONAL REQUIREMENTS : The project flow will be as follows:

1. Data Preprocessing- Cleaning of the dataset, Eliminating noise
2. Exploratory Data Analysis
3. Spliting the dataset into training data and testing data
4. Training the model using machine learning algorithms
5. Prediction of the model by user-input ‘
6. Checking accuracy, RMSE of the model
7. Optimizing if needed
8. Deploying the model
9. Creating UI using Node-Red

TECHNICAL REQUIREMENTS :

- Computer/Laptop
- Minimum 2.2GHz Processor
- Any Operating System.

SOFTWARE REQUIREMENTS:

- IBM Watson
- IBM Machine Learning Services
- Github
- Python v3
- Node-Red Application
- IBM Notebook / Jupyter Notebook.

PROJECT DELIVERABLES:

The project is about predicting the life expectency of people in a country using health, social, and economic variables. Countries can improve avg Life Excpectancy by tweaking and experimenting on various factors like healthcare system, immunizations, etc to increase the life expectancy of people living there from the predictions made by the model. User has to enter various fields like the adult mortality, GDP, disesases, etc of the country he resides in and can find his life expectancy.

PROJECT TEAM : Individual project.

PROJECT SCHEDULE:

The project is to be completed in 1 month and work for atleast 5 days a week. The project can be divided into two phases, one to build the model and th e other for UI.