Predicting Life Expectancy Using Machine Learning

PROJECT MANAGER : Hardik Vagrecha **Date :** 12/06/2020

PROBLEM STATEMENT:

In Supervised Machine Learning, Regression algorithms helps us to build a model by which we can predict the values of a dependent variable from the values of one or more independent variables.

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:

- This project is an end-to end project which will formulate a model while considering data from a period of 2000 to 2015 for all the countries.
- This project will predict the average time a human being is expected to live based on some factors.
- A country can predict the expected life of their citizens.According to that , the country can take necessary preventive measures to improve the healthcare system.
- This will serve as an example for countries to assess to improve life expectancy for their citizens.
- This will help in suggesting a country which area should be given importance in order to efficiently improve the life expectancy of its population.

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

FUNCTIONAL REQUIREMENTS : The project flow will be as follows:

1. Preprocessing of the dataset- 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, R2 score, 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 Service
- Github
- Python v3
- 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. The country can improve factors like healthcare system, immunizations, etc to increase the life expectancy of people living in the country 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 the other for UI.