

1. Project Summary

The aim of this project is to create a new model using WHO data set and to predict life expectancy based on certain values.

I have implemented Random Forest Regression while considering data from period of 2000 to 2015. This study will focus on immunization factors, mortality factors, economic factors, social factors and other health related factors as well. Since the observations in this dataset are based on different countries, it will be easier for a country to determine the predicting factor which is contributing to lower value of life expectancy. This will help in suggesting to a country which area should be given importance in order to efficiently improve the life expectancy of its population.

2. Project Scope

Life expectancy is one of the most important factors in end-of-life decision making. Advance care planning improves the quality of the final phase of life. Sometimes physicians however tend to overestimate life expectancy and miss the window of opportunity. In this project by using machine learning and data processing we will be predicting life expectancy from Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors.

With the help of machine learning it will be a feasible and promising approach to predicting life expectancy. It will help in timely recognition of the right moment to start Advance Care Planning.

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 in the dataset.

3. Team

Individual Project

4. Functional Requirements

- Predict Life Expectancy

- Process the dataset and find dependent and independent variables relationship.

5. Technical & Software Requirements

- The project will be built and deployed on IBM Cloud using IBM's Watson machine learning platform.
- Node-RED flow will be used to Deploy the ML services.
- Code for the project will be written using Python 3.6
- Dependencies used - Pandas and Scikit-Learn etc.

6. Project Deliverables

Predicting Life Expectancy using Machine Learning to deliver the outcome for the given dataset of any county. It will predict the life expectancy based on the following:

- Country
- Year
- Status
- Adult Mortality
- infant deaths
- Alcohol
- percentage expenditure
- Hepatitis B
- Measles
- BMI
- under-five deaths
- Polio
- Total expenditure
- Diphtheria
- HIV/AIDS
- GDP
- Population
- thinness 10-19 years
- thinness 5-9 years
- Income composition of resources
- Schooling

7. Project Schedule

Predicting Life Expectancy using Machine Learning will be develop completely with maximum accuracy with in given dead line which is 30 days (15 May to 15 June).