Life Expectancy using Machine Learning Model

**INTRODUCTION**

* 1. **Overview**:

We are in an unpredicted era where human are living longer. Life expectancy is one of the most important factors in end-of -life decision making. It helps to create advance health care policies. Prediction of individual life span represents a significant challenge for ageing research that is important for understanding factors influencing longevity, as well as identifying life-span–associated characteristics that can be studied as surrogates of longevity in laboratory experiments. Many algorithms can be used such as Linear Regression, Ridge Regression, Lasso Regression, Re, Linear Regression with Polynomic features, Decision Tree Regression, Random Forest Regression

**1.2 Purpose:**

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.

**LITERATURE SERVEY**

* 1. **Existing Problem:**

1. The government doesn't get to know the reasons for the increasing death rate.
2. People are not aware of the cause of deaths in their surroundings.
3. People doesn't know which disease is how much fatal and for which they should
   1. **Proposed Solution:**

We are aiming framework to study life expectancy and life span over the time. It would have on greater effect on economy. Prognosis of life is not only instrumental predicting living rate but also helps in deciding whether there is tendency of diseases in a continent.

**THEORETICAL ANALYSIS**

* 1. **Block Diagram:**

**USER**

**KK** INPUT REQUEST

**WEB INTERFACE**

**MACHINE LEARNING MODEL**

PREDIC PREDICT WRITE

**3.2 Software Design:**

Node-red App Design

**CREATE A FORM WITH REQUIRED FIELDS**

**SET REQUEST FOR TOKEN AND SET GLOBAL VARIABLE**

**ADD 1ST HTTP REQUEST FOR ACCESSING SERVICE**

**STORE RECEIVED VALUE FROM MODEL AND DISPLAY ON THE FORM**

**ADD 2ND HTTP REQUEST FOR ACCESSING MODEL**

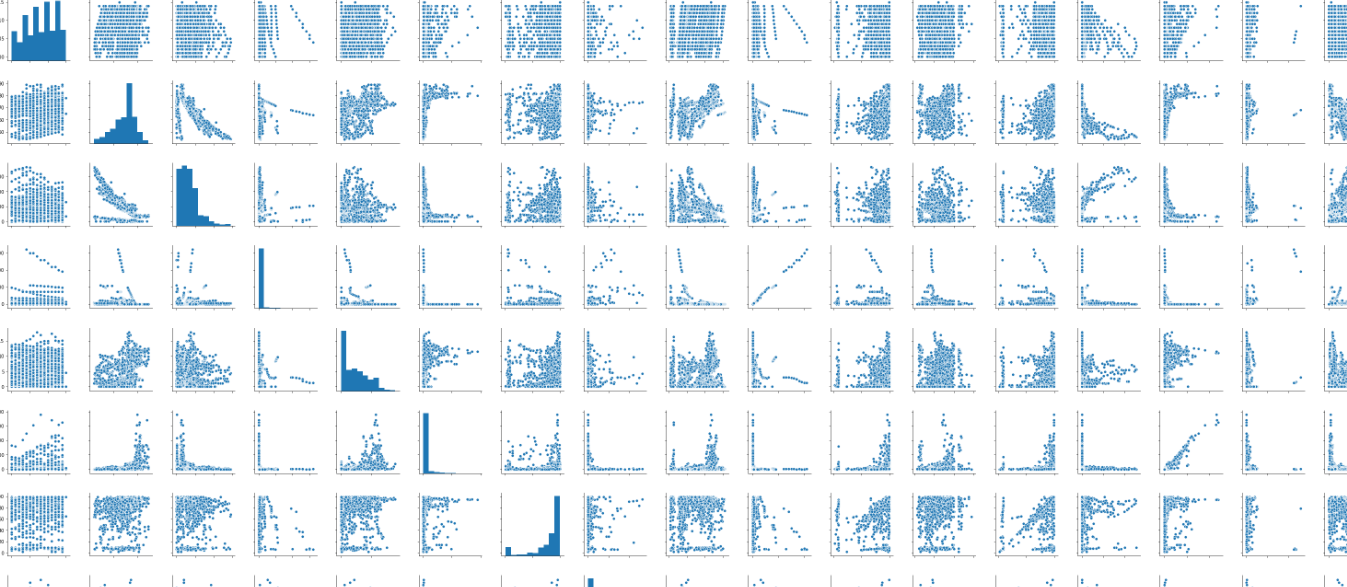
**GET THE VALUES FROM ENDPOINT**

**DISPLAY THE PREDICTION**

**EXPERIMENTAL INVESTIGATION**

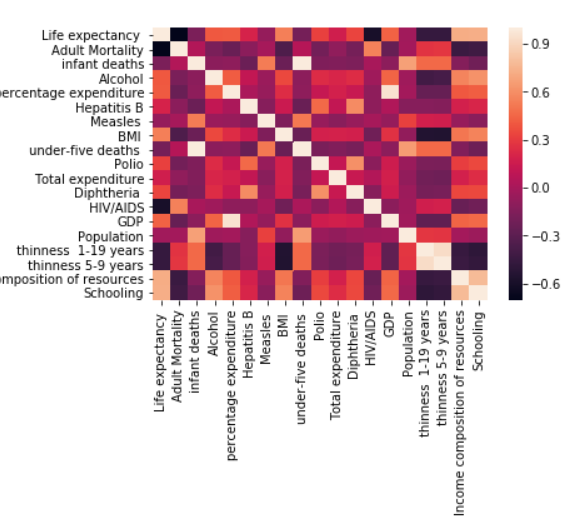
1. Pairplot:

A pairplot allows us to see distribution of both single variable and relationship two variables. Hence using pair plot relationships between two variables of the dataset can be understood

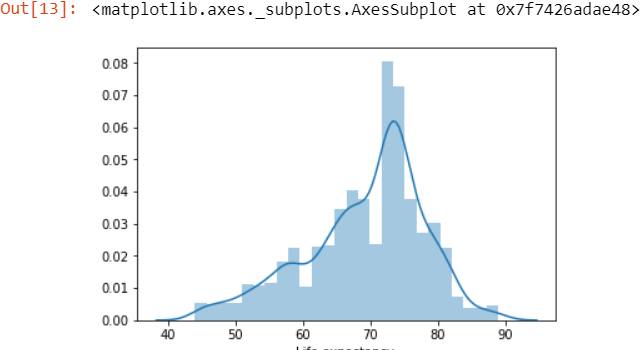


1. Heatmap:

Heatmap is useful in identifying which columns are more useful while predicting life expectancy particular column value. For example how much adult mortality will contribute towards life expectancy can be known from this heatmap.



3.Distplot:



Distplot is like a histogram which gives how a particular value falls in a particular column.

**FLOWCHART**

COLLECT DATASET

IMPORT REQUIRED LIBRARIES

REMOVE THE NULL VALUES

ANALYSE THE DATA AND PLOT GRAPH

TRAIN THE MODEL,TEST MODEL

PREDICT THE VALUE FOR DATASET

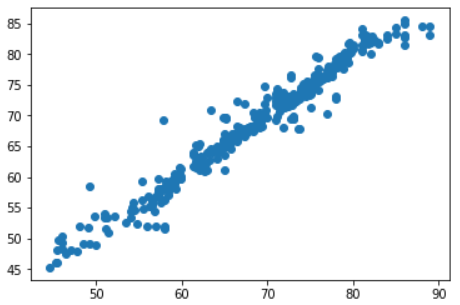
CALCULATE EVALUATION METRICS PARAMETERS

CONNECT THE MODEL AND UI APP

DISPLAY ON THE WEBPAGE

**RESULT:**

Scatter Plot using Random Forest Regression:



Metrics Parameters:

|  |  |
| --- | --- |
| MAE | 1.1583535108958865 |
| MSE | 3.4548060968523013 |
| RMSE | 1.8587108696223578 |

**Advantages:**

1. Government will be aware of the reasons behind the death rate.
2. Government will get to know which area has a higher death rate, so can increase medical and healthcare facilities in that area.
3. People will be aware about life expectancy and reasons behind increasing death rate and hence will take preventive measures accordingly.

**Disadvantages :**

1. It requires internet connection.

2. Application is unable to predict value for multiple sets of data at the same time.

**Applications :**

1. One of the applications of life expectancy is in the financial world, including life

insurance, pension planning.

2. Social benefits and healthcare facility management.

3. Life expectancy is used in pricing and underwriting life insurance .

**Conclusion:**

The value of life expectancy can be calculated using machine learning model and node red app. It reduces the complexity.

**Future scope :** Accuracy of the model can be improved further. And the system can be improved in a way such that the user will be able to predict life expectancy values for multiple sets of data at the same time.

**Bibliography:**

* https://www.who.int/whosis/whostat2006DefinitionsAndMetadata.pdf

● https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(15)60296-3/fullt ext

* https://medium.com/swlh/predicting-life-expectancy-w-regression-b794ca457cd4

**Appendix:**

* Source Code:

https://colab.research.google.com/drive/1WxQnppcybl6cUnGTL8nRnUqbTgjyyM7k?usp=sharing