

Predicting Life Expectancy Using Machine Learning

(Smart Internz).

Category : MachineLearning TimePeriod : 30days

Data : WHO Data sheet from Kaggl

Expectancy with 22 features From 193 countries and 2938

Records(2000-2015)years data

Tools : Python,IBM Watson studio,IBMcloud,Node red app

Project submitted by : **MOHIT SINGH MAHUR**

INTRODUCTION

OverView

This Project is done under the guidelines from SmartInterz Team. This helped me to start learning Machine Learning. I'm going to use a huge data and some IBM Cloud Tools to complete our project.

I think it one of My grate Opportunities

In this project we are going to work on Huge data set with python to handle that data

Purpose

Purpose of this project is to predict life expectancy of a country by using some dependent features .

This will help Government and other Non-Governmental Organizations to take most effective steps to increase the life expectancy of Humans

LITERATURE SURVEY

Existing Problems on Life Expectancy

The most important causes of death in Western industrialized countries are cardio and cerebrovascular diseases and malignancies. For instance, in Germany in 2008, 68.6% of all women and 65.9% of all men died from these diseases.

The third most frequent cause of death are respiratory diseases which cause less than 10% of deaths each year.

Not only because of diseases even climatic changes, natural disaster, human mistakes in huge industries where many people work.

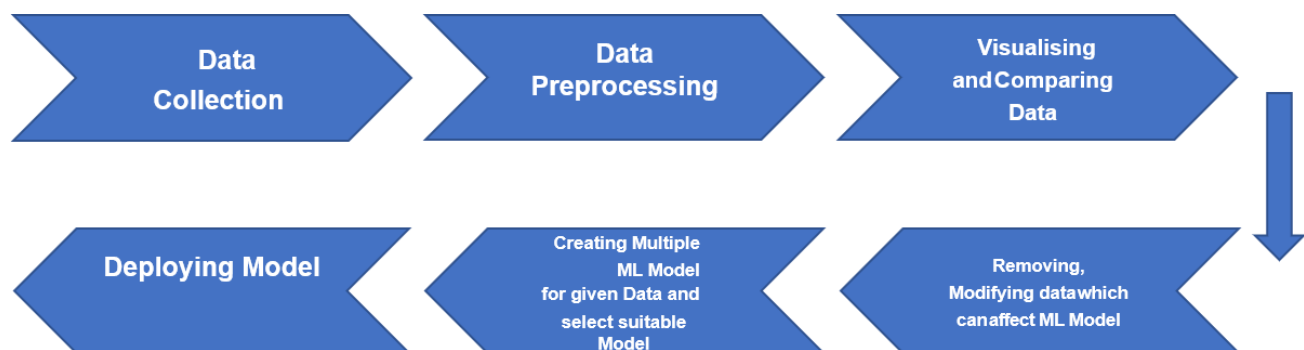
Proposed Solution

Regular physical activities reduces the risk and improves many diseases including arterial hypertension, diabetes , obesity, coronary heart disease, chronic heart failure, and chronic obstructive pulmonary disease .

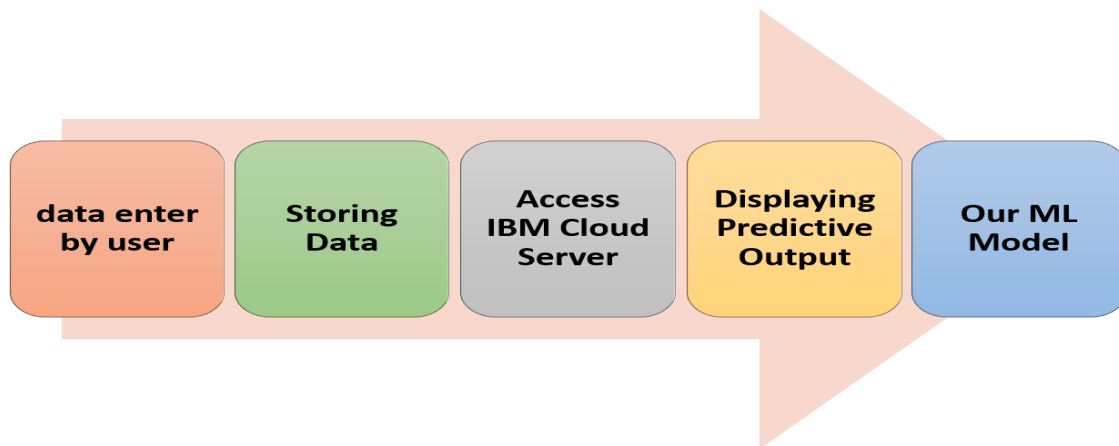
In addition, the risk of colon , breast ,lung diseases, and pancreatic cancer is reduced, Even there physical Fitness will help them to withstand the natural disaster and climatic changes.

Theoretical Analysis

Machine Learning Model



Block Diagram



Software and Hardware

Software from IBM

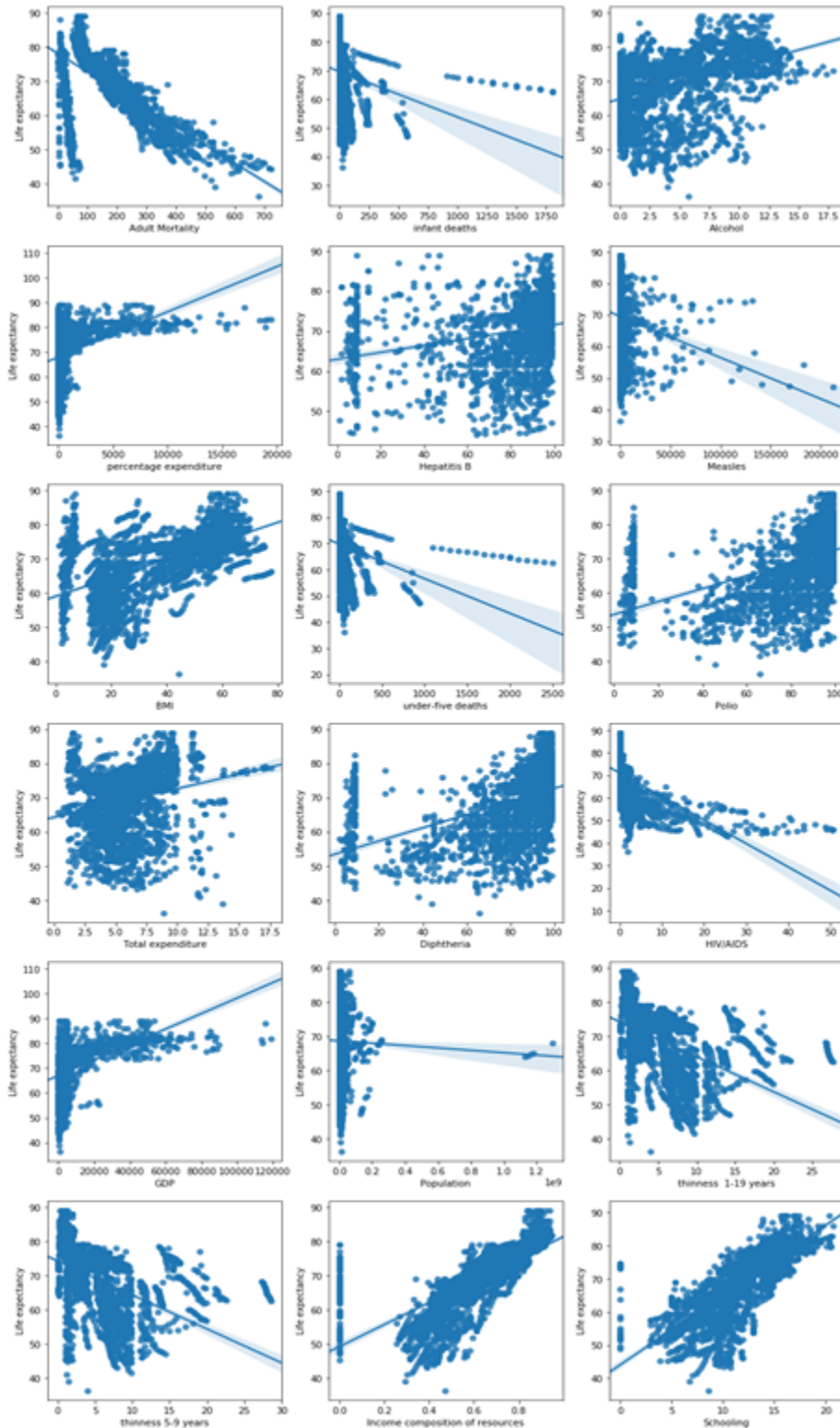
- a. IBM Cloudservice
- b. Watson Studio : To create our model using jupyternotebook
- c. Machine learning service : Deploying the createdmodel
- d. Node red app : Creating dashboard for our MLmodel

Hardware

1. IBM cloudstorage
2. IBM cloudserver

Experimental investigation

Comparing Life expectancy with other features

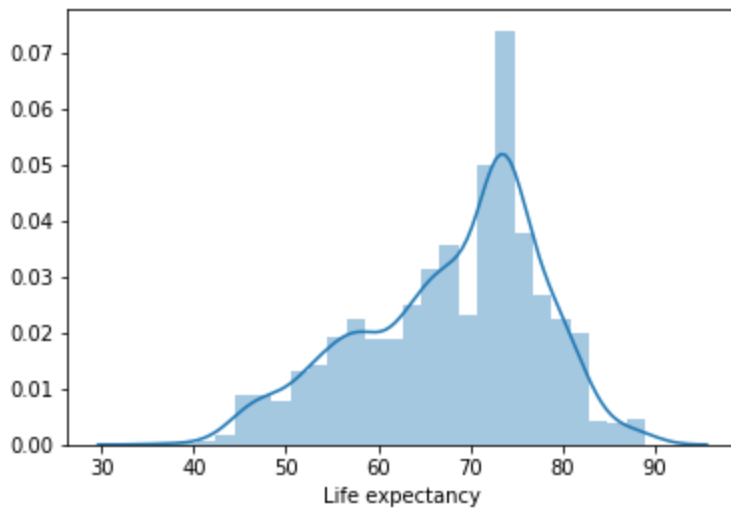


In above graphs we can see some regular pattern in between Life Expectancy and other features

Some of the important features based on the above graph are

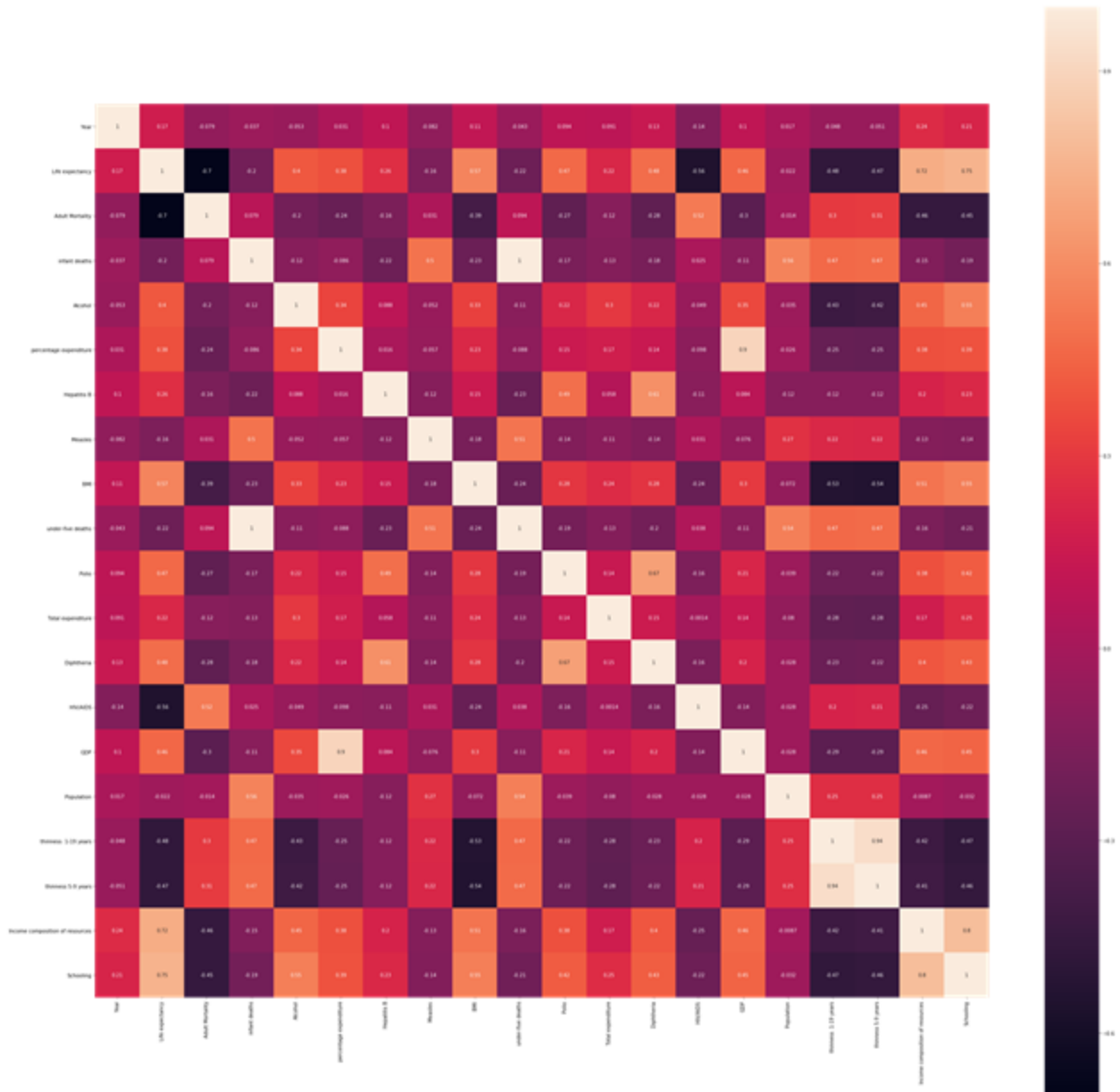
1. AdultMortality
2. Alcohol intake
3. Schooling
4. PercentageExpenditure
5. HIV/AIDS
6. Income composition of resources...

Life Expectancy Plot



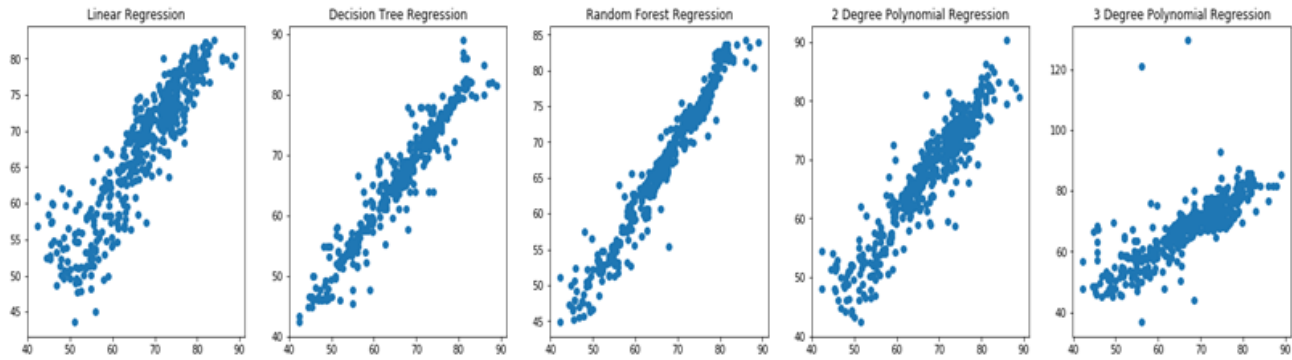
In above graph we can view that most of the countries having Life Expectancy in between 70 -80 years.

correlation between attribute visualizing usingHeat-map



1. Correlation to check how strong the attributes are dependent to each other
2. In the above graph , light color are show highly dependent attributes and dark are show less dependent attributes

Visualizing which model is most efficient

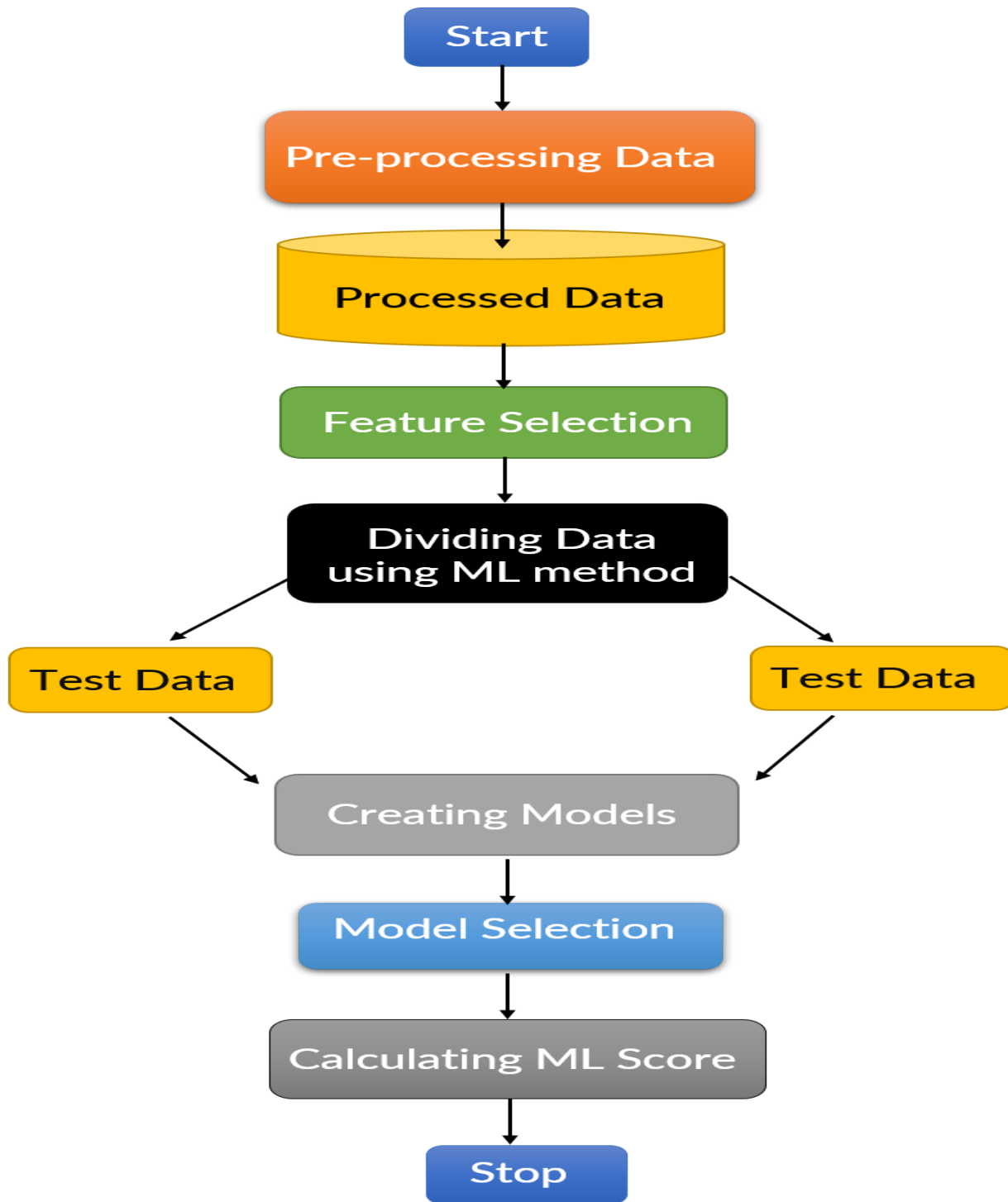


Here I used five models :

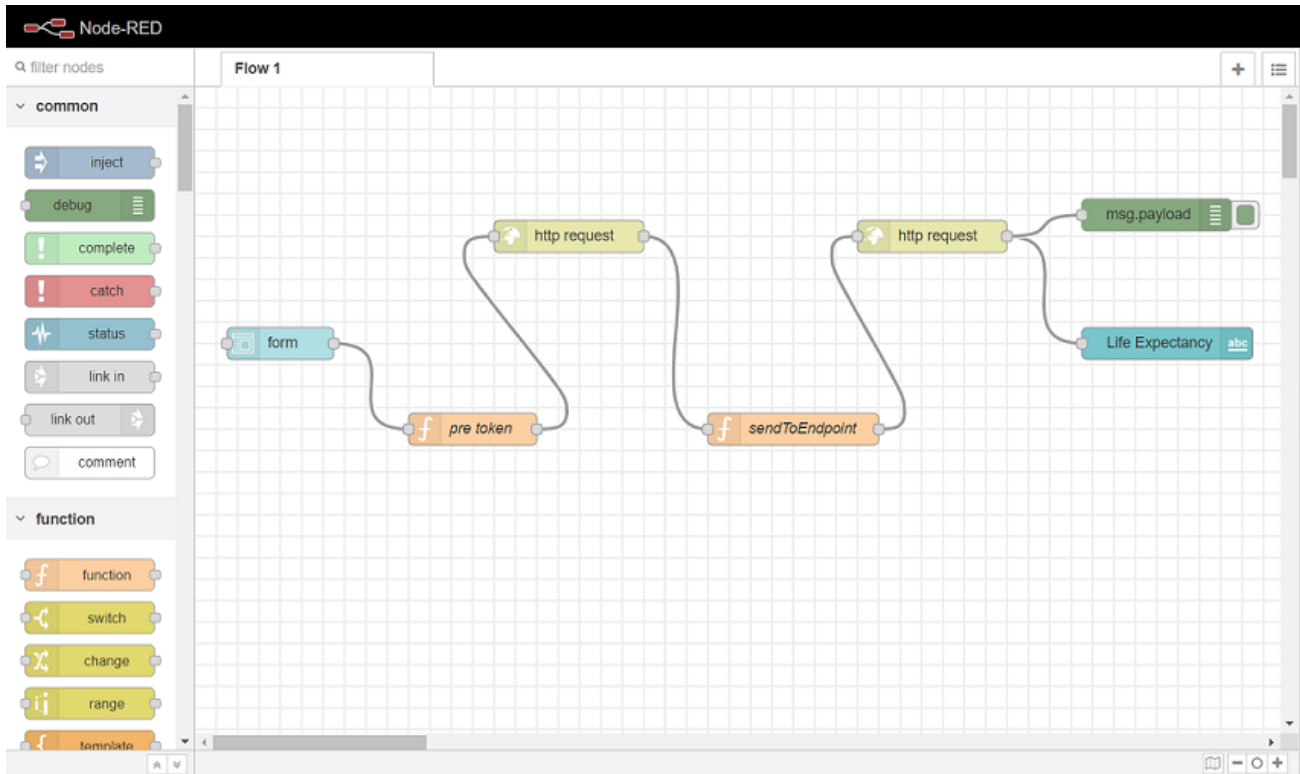
1. Linear regressionmodel
2. Decision treemodel
3. Random forest regressionmodel
4. 2-degree polynomialregression
5. 3-degree polynomialregression

From above graph we can see that Random Forest Regression is most effective model for our data.

FLOW CHART



Node-Red flow

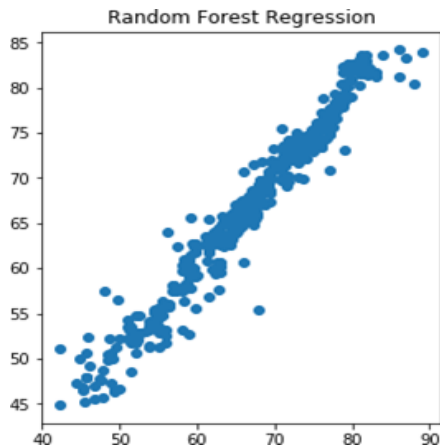


Explanation:

1. We are using form Node for taking input from user in the formFormat
2. PRE TOKEN is a function which is used convert local variables in to global and creating a MSG for HTTP request to get access from IBMcloud
3. Token function is used to create msg.header from token, instance id and data in the format of jsonobject
4. In Second HTTP request we are sending the json object to the model and getting predicted Life Expectancy from it
5. Edit text node is used to print the predicted values on dashboard

Result

Model Result



graph between predicting and actual value

Mean absolute error : 1.2867203219315895

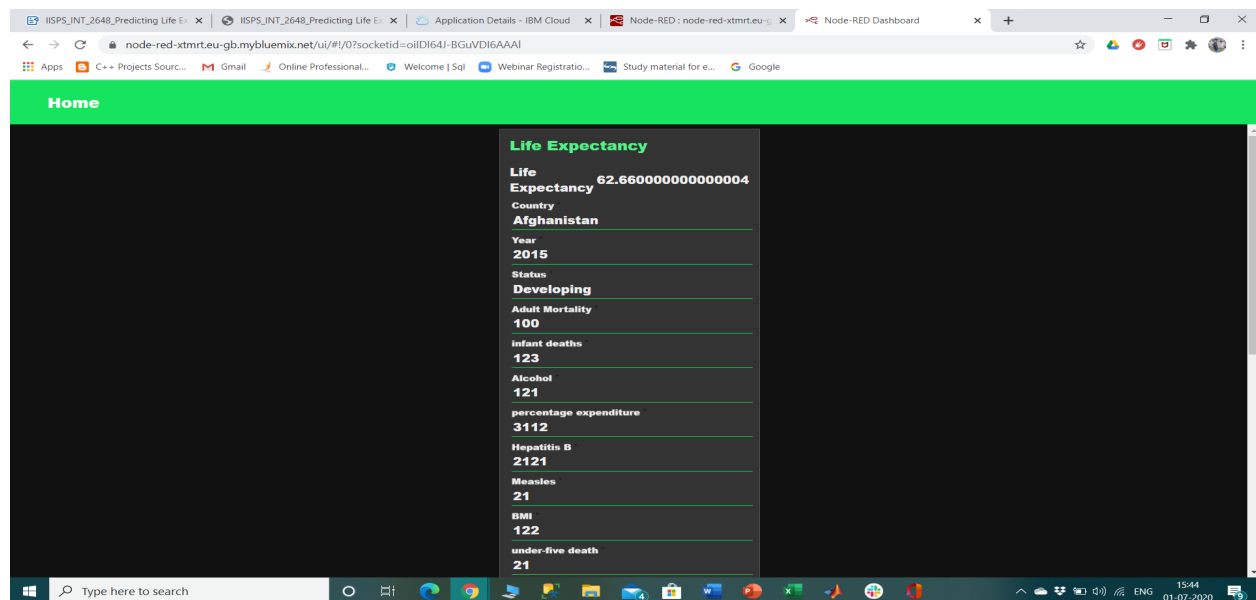
Mean Square Error : 3.8028237424547293

Root Mean Square Error : 1.9500830091190295

For This Record = {Afghanistan, 2015, Developing, **65** , 263, 62 ,0.01,
71.2796236, 65, 1154 , 19.1 ,83 ,6 ,8.16 ,65 ,0.1 ,584.25921 ,33736494
,17.2,17.3,0.479 ,10.1}

We got 67.19 as output but actual is 65, it is near to the result

Dashboard



In the above example is the same input but change adult mortality value, by decreasing it the Life Expectancy value is increasing.

Advantage and Disadvantage

Advantages

- i. Predicting life Expectancy will help us to monitor the health of the people
- ii. By predicting it, we can improve our health condition and public health care center
- iii. We can compare things and make decision for the future to increase life span of humans

Disadvantages

- i. If our predictions are wrong that may lead to wrong judgement for future.
- ii. People may get upset by looking at the result, If the results are published.
- iii. Some people make decisions by looking at the age of the person, If the age of the person exceeds over Life Expectancy then there is a chance of under caring of him/her .

Applications

- i. By using this model we can create App
- ii. Life-tables and Demographic applications
- iii. This type of thing may increase awareness on their lives
- iv. We can create applications like sharing which type of activities will increase the life expectancy

Conclusion

- i. Predicting life expectancy will help to improve our living conditions and lifestyle
- ii. These type of applications will increase the awareness among people.
- iii. It is clear that some of our habits will affect our life.

iv. Education is also required to increase life span of humans.

BIBLIOGRAPHY

Dashboard link

<https://node-red-xtmrt.eu-gb.mybluemix.net/ui/#!/0?socketid=oiDI64J-BGuVDI6AAA>

Mentors of SmartInterz really helped me a lot in deploying my ML model, Creating Dashboard for my project

Sincere Thank all the Mentor of SmartInterz