NAME : AADHITHYA V

COLLEGE : KPR INSTITUTE OF ENGINEERING AND TECHNOLOGY

INTERNSHIP TITLE : Predicting Life Expectancy using Machine Learning- SB51678

PROJECT ID : SPS_PRO_215

PROJECT TITLE : Predicting Life Expectancy using Machine Learning

Email id : 19ec001@kpriet.ac.in

Github Link :-

https://github.com/SmartPracticeschool/llSPS-INT-2007-Predicting-Life-Expectancy-using-Machine-Learning

SBID : SB20200051678

PROJECT REPORT : -

INTRODUCTION : -

OVERVIEW : -

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 difference , Mental illness, Physical illness, Education , Year of birth and other demographic factors. This problem statement provides a way to predict average life expectancy of people living in country when various factors such as year , GDP , education, alcohol intake of people in the country, expenditure, on healthcare system and some specific diseases related deaths that happened in the country are given.

PURPOSE : -

Life expectancy at birth is one of the most frequently used health status indicators. Gains in life expectancy at birth can be attributed to a number of factors, including rising living standards , improved lifestyle and better education , as well as greater access to quality health services. Life expectancy plays an important role when decisions about the final phase of the life need to be made. Good prognostication for example helps to determine the course of treatment and helps to anticipate the procurement of health care services and facilities or more broadly facilitates advance care planning. Thus predicting life expectancy of the human being becomes

signicant.

LITERATURE SURVEY : -

EXISTING PROBLEM :-

Life expectancy is a commonly used stastic to measure the health of a population . On average the current worldwide life expectancy is approximately 70 years, but this is influenced by many factors such as gender , geographical location , family history , alcohol consumption, diet and social habits. In this competitive and health concious society , we are making smarter foods and lifestyles to make our lives easier but how do we know if our habits are actually prolonging our lives ? what if youncould find out how many years or days are remaining in your future. The seeker desires a novel approach for combining information to accurately predict the longevity of an individual human being.

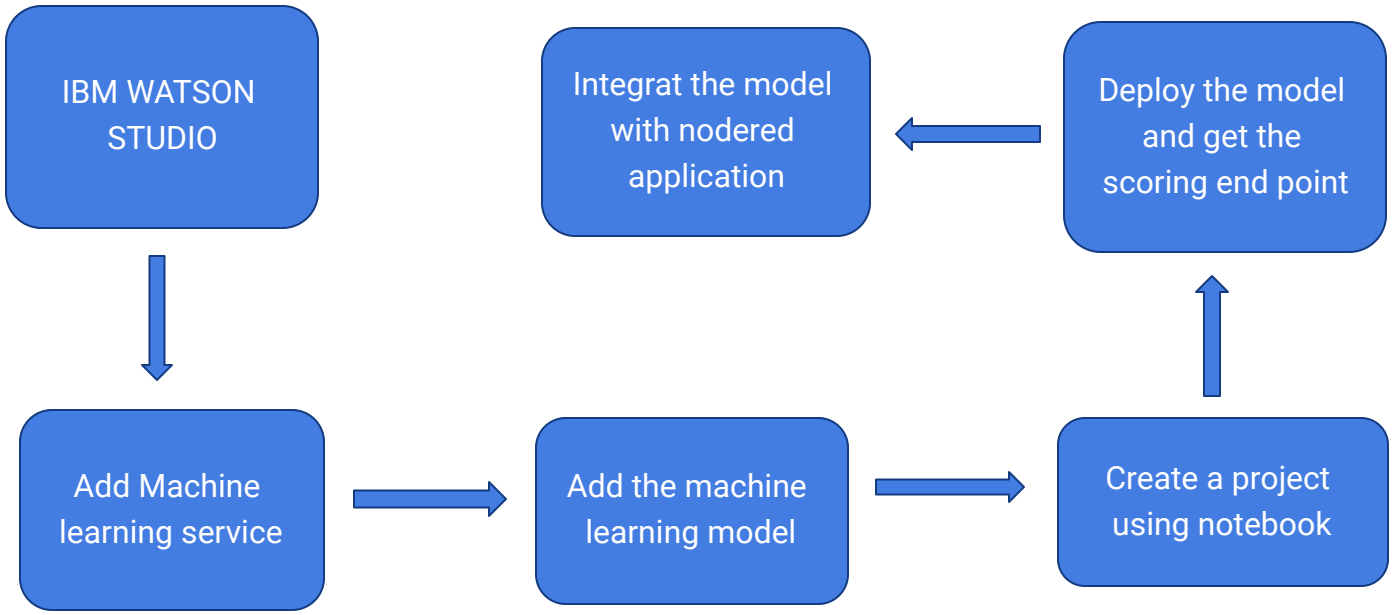
PROPOSED SOLUTION : -

There has been an explosion of break throughs in the field of Machine learning over the past few years. Machine learning algorithms are capable of lot and can do wonders for the healthcare sector. The proposed solution involves the use of Machine learning algorithms specifically regression models such as linear regression , ridge regression , etc. Life expectancy is correlated over time among countries and between males and females . Here we propose a method for forecasting life expectancy of an individual from a country taking inton certain factors as

Adult mortality rate , infant deaths ,alcohol , Hepatitis B , measles ,BMI , Polio , total expenditure , diptheria , HIV/AIDS , GDP of a country , population, income composition of resources , schooling and status of the country in terms of developing or developed .
This machine learning model will be made accesible to the users by integrating with nodered to create an interactive and user friendly and user interface .

THEORETICAL ANALYSIS : -

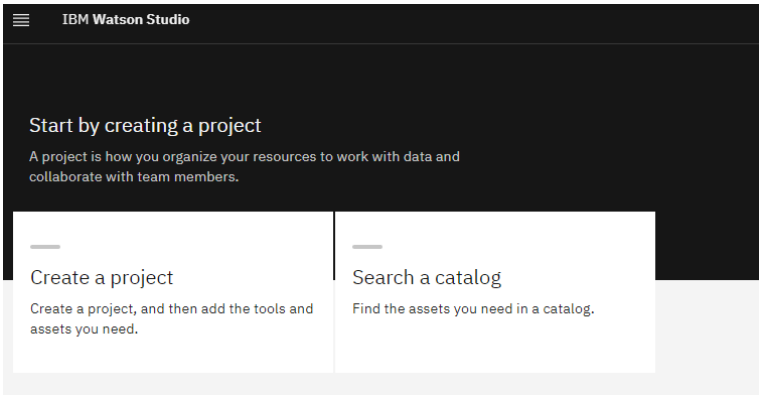
BLOCK DIAGRAM :-

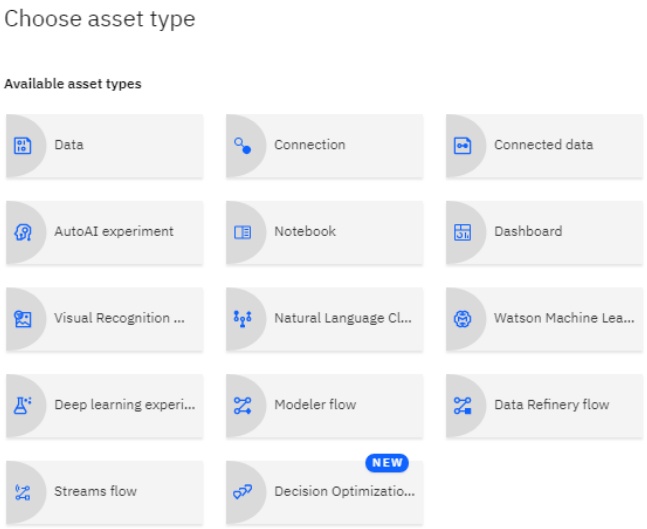


HARDWARE / SOFTWARE DESIGNING :-

IBM WATSON :-

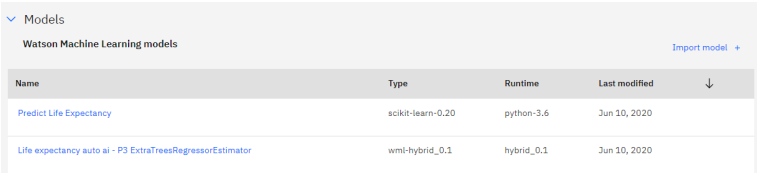
Log into your IBM Cloud ---> Get into IBM watson studio ---> Create an empty project ---> Add the dataset ---> Add notebook ---> Create a model --->Deploy the model ---> Get the scoring url .






```
deployment = client.deployments.create(published_model_uid, name="Predict Life Expectancy")

#######################################################################################

Synchronous deployment creation for uid: '0f880480-4619-4e80-a42d-4ea66fcc1cd7' started

#######################################################################################


INITIALIZING
DEPLOY_SUCCESS


------------------------------------------------------------------------------------------------
Successfully finished deployment creation, deployment_uid='120571dd-52dd-469b-b048-b9f3c4bca86a'
------------------------------------------------------------------------------------------------
```

```
scoring_endpoint = client.deployments.get_scoring_url(deployment)
scoring_endpoint

'https://us-south.ml.cloud.ibm.com/v3/wml_instances/25673d21-fefd-4fe2-8394-82dbc908c5b9/deployments/120571dd-52dd-469b-b048-b9f3c4bca86a/online'
```



NODERED :-

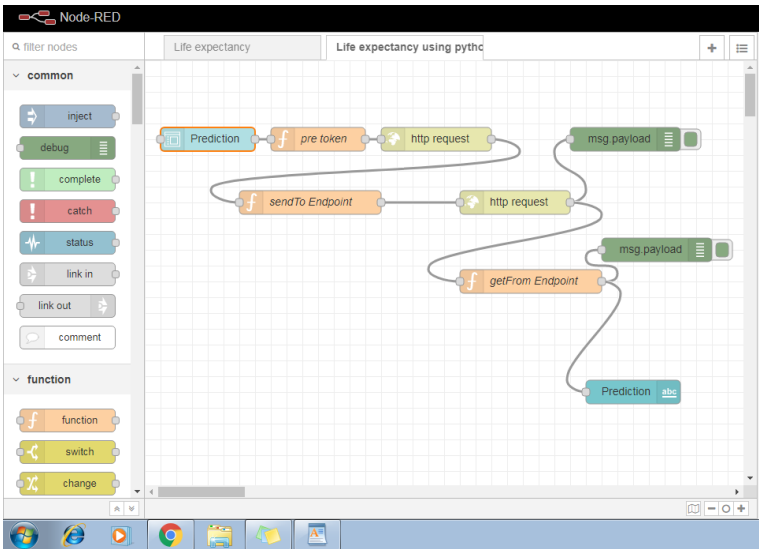
Create nodered application in IBM cloud ---> Create a flow to integrate ML with nodered.

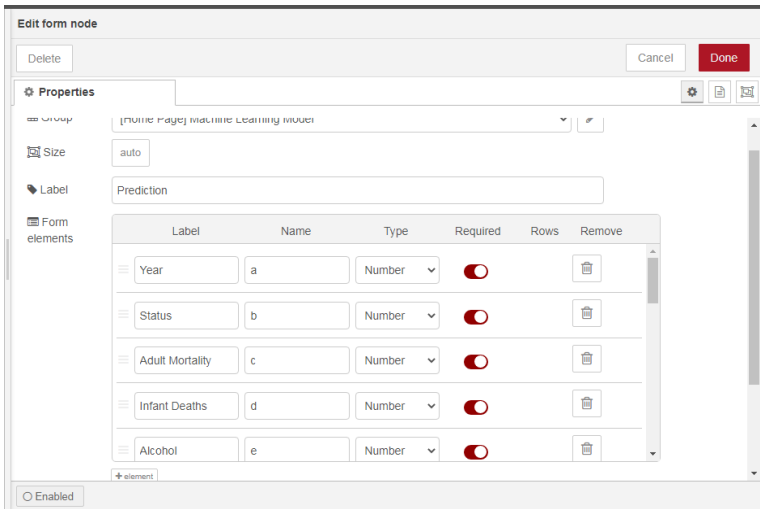
NODES :-

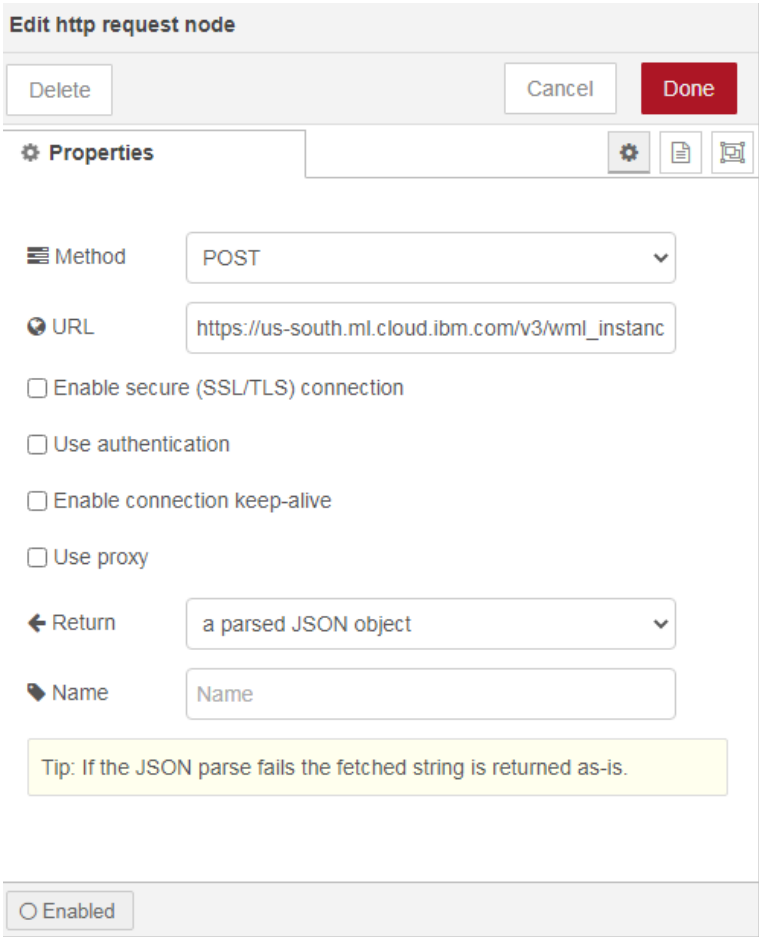
Form node = Add your required inputs

Function node = Enter the fields and values inserted in your model

http request = Add your wml credentials and the scoring url which connects with your model.







EXPERIMENTAL INVESTIGATIONS :-

Analysing the relationship between various factors on Life expectancy can help us improve the performance of the model and predict the target with more accuracy .

```
prediction = regression.predict(x_test)
```

```
plt.scatter(y_test, prediction)
plt.xlabel('Test Result', fontsize = 20)
plt.ylabel('Prediction', fontsize = 20)
plt.show()
```


```
df_data_0.head()
```

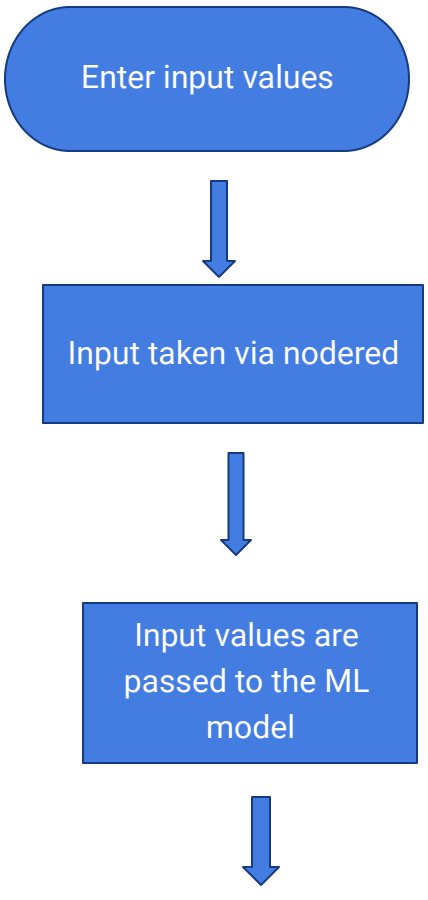
	Country	Year	Status	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	...	Polio	Total expenditure	Diphtheria	HIV/AI
0	Afghanistan	2015	Developing	65.0	263.0	62	0.01	71.279624	65.0	1154	...	6.0	8.16	65.0	0.1
1	Afghanistan	2014	Developing	59.9	271.0	64	0.01	73.523582	62.0	492	...	58.0	8.18	62.0	0.1
2	Afghanistan	2013	Developing	59.9	268.0	66	0.01	73.219243	64.0	430	...	62.0	8.13	64.0	0.1
3	Afghanistan	2012	Developing	59.5	272.0	69	0.01	78.184215	67.0	2787	...	67.0	8.52	67.0	0.1
4	Afghanistan	2011	Developing	59.2	275.0	71	0.01	7.097109	68.0	3013	...	68.0	7.87	68.0	0.1

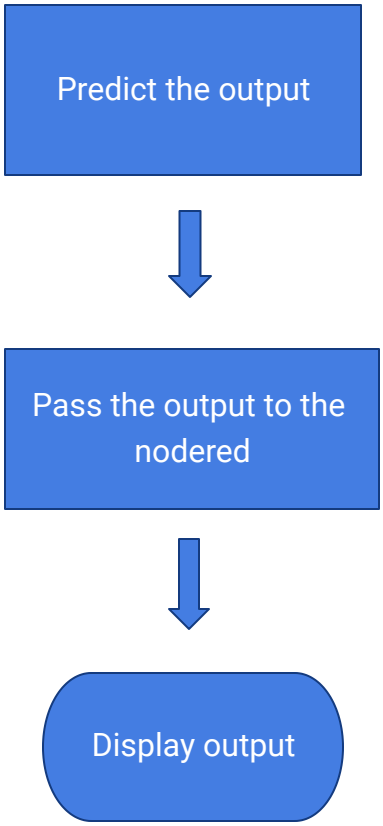
5 rows × 22 columns

```
df_data_0.columns.values
```

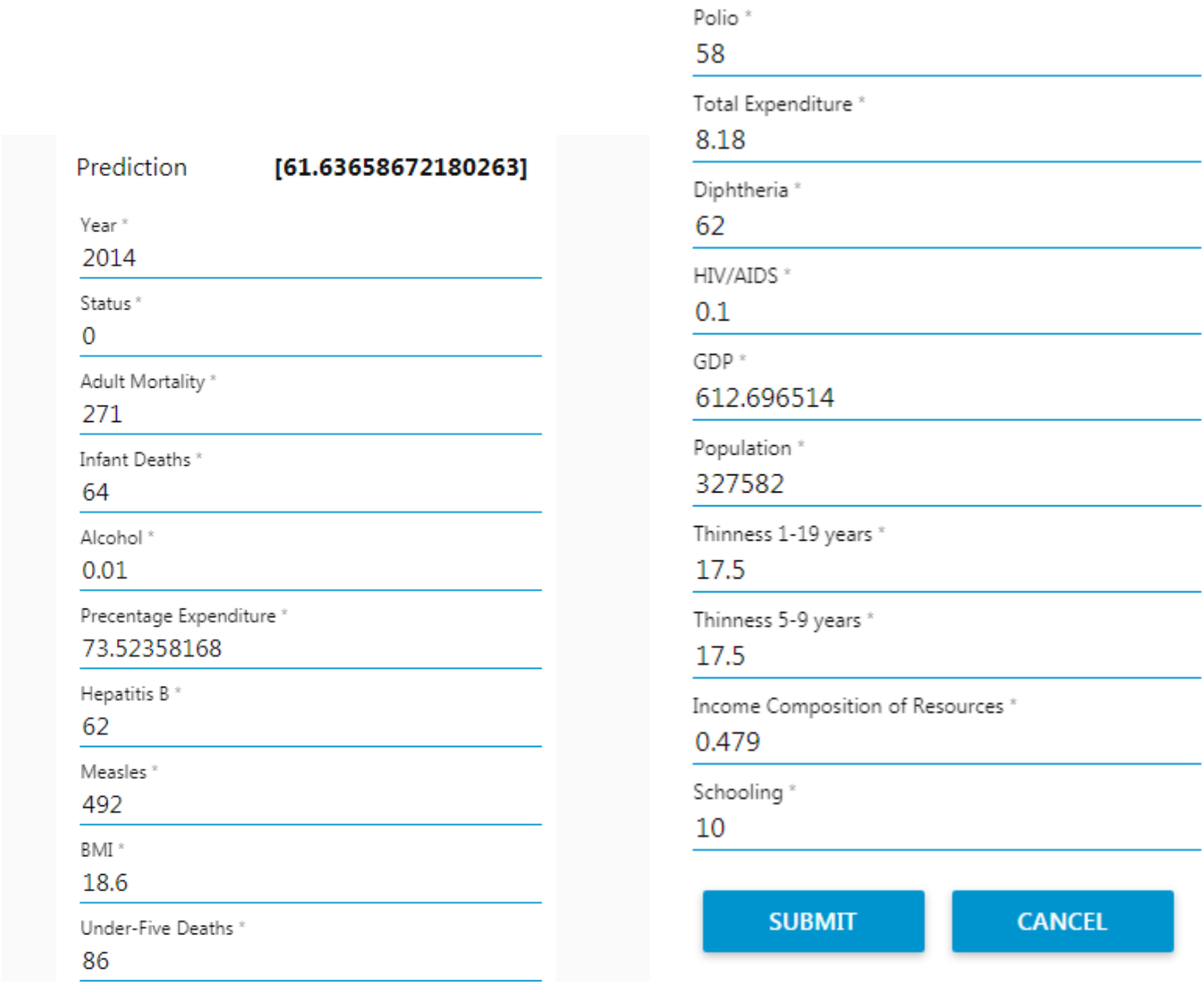
```
array(['Country', 'Year', 'Status', 'Life expectancy', 'Adult Mortality',
       'infant deaths', 'Alcohol', 'percentage expenditure',
       'Hepatitis B', 'Measles', 'BMI', 'under-five deaths', 'Polio',
       'Total expenditure', 'Diphtheria', 'HIV/AIDS', 'GDP', 'Population',
       'thinness  1-19 years', 'thinness 5-9 years',
       'Income composition of resources', 'Schooling'], dtype=object)
```

FLOW CHART : -




Predict the output
Pass the output to the nodered
Display output

RESULT :-



Thus the result obtained using the ML service and Nodered.

ADVANTAGES : -

One of the biggest advantages of embedding machine learning algorithms is their ability to improve over time . Machine learning typically improves efficiency and accuracy .

Using nodered also simplifies the effort put into creating the front end . The

programmer doesn't need extensive knowledge on HTML and javascript. It also makes the integration between Machine learning model and UI easier.

DISADVANTAGES : -

Using machine learning interface comes with its own problems . since the whole point of it is minimise human involvement, it also makes error detection and fixing much more problamatic. It takes a lot of time to identify the root cause of a problem. It is time consuming.
At the same time , Nodered does not give many features to customize our UI.

APPLICATIONS :-

1 . Individuals can predict their own life expectancy by inputing values in the corresponding fields . This could make people aware of their mental health .
2 . Based on the factors used to calculate life expectancy of an individual and the outcome , healthcare will be able to fund and provide better services to those with greater need
3 . Insurance sector will be able to provide individualized services to people based on the life expectancy outcomes and factors .

CONCLUSION : -

Predicting life expectancy of an individual can greatly alter our

lives. Human behaviour and activities are so unpredictable , it may almost be impossible to predict the lifespan with accuracy . However with the help of Machine learning algorithms such as regression models, we can get close to predicting a nearest roundabout value . With the help of machine learning algorithms one can ease the process of automating the application and predicting the expectancy with great accuracy . It also reduces the time and effort .

FUTURE SCOPE : -

For future use one can integrate the life expectancy prediction with providing suggestions and medications to the individual using the application . This will help predict as well as increase the individuals life expectancy

Mental health can also be taken into account while predicting life expectancy with the help of sentiment analysis systems as well .

BIBLIOGRAPHY : -

https://www.allbusinesstemplates.com/download/?filecode=2KBA4&lang=en&iuid=9f9faa69-9fab-40ee-8457-ea0e5df8c8de

https://www.ibm.com/cloud/get-started

https://developer.ibm.com/tutorials/how-to-create-a-node-red-starter-application/

https://developer.ibm.com/technologies/machine-learning/series/learning-path-machine-learning-for-developers/

https://www.kaggle.com/kumarajarshi/life-expectancy-who

https://bookdown.org/caoying4work/watsonstudio-workshop/jn.html

APPENDIX : -

SOURCE CODE :

```
import types
import pandas as pd
from botocore.client import Config
import ibm_boto3

def __iter__(self): return 0
```

```
data3 = data2.copy()
features = data3.iloc[:,:-1]
features.columns.values
```

```
array(['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 1-19 years',
       'Thinness 5-9 years', 'Income Composition of Resources',
       'Schooling'], dtype=object)
```

```
targets = data3.iloc[:,-1]
targets.shape
```

```
(2928,)
```

```
features.shape
```

```
(2928, 20)
```

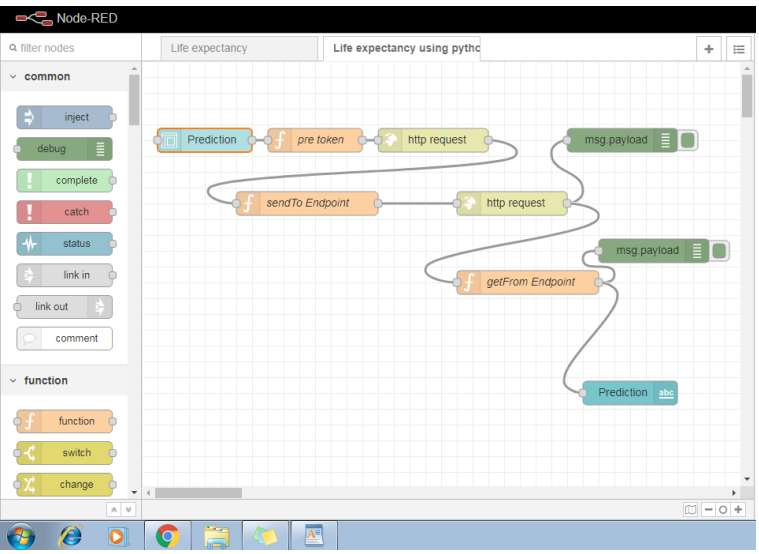
```
from sklearn.model_selection import train_test_split
```

```
train_test_split(features, targets)
```

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
import statsmodels.api as sm
```

```
df_data_0.head()
```

	Country	Year	Status	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	...	Polio	Total expenditure	Diphtheria	HIV/AI
0	Afghanistan	2015	Developing	65.0	263.0	62	0.01	71.279624	65.0	1154	...	6.0	8.16	65.0	0.1
1	Afghanistan	2014	Developing	59.9	271.0	64	0.01	73.523582	62.0	492	...	58.0	8.18	62.0	0.1
2	Afghanistan	2013	Developing	59.9	268.0	66	0.01	73.219243	64.0	430	...	62.0	8.13	64.0	0.1
3	Afghanistan	2012	Developing	59.5	272.0	69	0.01	78.184215	67.0	2787	...	67.0	8.52	67.0	0.1
4	Afghanistan	2011	Developing	59.2	275.0	71	0.01	7.097109	68.0	3013	...	68.0	7.87	68.0	0.1


Node-RED
filter nodes
Life expectancy
Life expectancy using pythc
common
inject
debug
complete
catch
status
link in
link out
comment
function
function
switch
change
Prediction
pre token
http request
msg payload
sendTo Endpoint
http request
msg.payload
getFrom Endpoint
Prediction
abc

THANKYOU