SMARTBRIDGE

Let's Bridge the gap

Summer Internship report

Predicting Life Expectancy using Machine Learning

21/05/2020-18/06/2020

Submitted by:

Shubhangi Shreya

From:

SmartBridge Educational Services Pvt Ltd.

Plot No 132, Above DCB bank, 2nd floor,

Bapuji Nagar, Habsiguda,
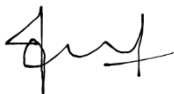
Nacharam Main Road, Hyderabad – 500076

Date: 21/05/2020.

Dear **Shubhangi Shreya**

SmartBridge Educational Services Pvt Ltd, is pleased to offer a training cum internship opportunity. During this period you would be associated with our mentors and The Smart Practice School Platform.

 For further details you can contact us on +91 8499004200.

Thanks and Regards,

Ch. Jaya Prakash

Program Manager – SIP2020,
Date: 16/05/2020.

Preface

This report documents is the work done during the summer internship at **SmartBridge Educational Services PVT Ltd**, for the prediction of lif- expectancy under the supervision of **mentors of SmartBridge and smart practice school platform**. The report will give an overview of the tasks completed during the period of internship with technical details. Then the results obtained are discussed and analyzed. I have tried my best to keep report simple yet technically correct. I hope I succeed in my attempt.

Shubhangi Shreya

Index

Introduction

Overview

The dataset for this project is taken from Kaggle. It is a platform for predictive modelling and Analytics competitions. Here organization and researchers post the data. Statisticians and data scientist from all over the world compete to produce the best models.

Problem Statement- Predicting Life Expectancy Using Machine Learning

Problem Description :

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.

Purpose

Built a machine learning model for the prediction of life expectancy.

LITERATURE SURVEY

Existing problem

Predicting the **lifespan** of people, or their "Personal **Life Expectancy**" (PLE) would greatly alter our lives. On one hand, it may have benefits for policy making, and help optimise an individual's health, or the services they receive.

Proposed Solution

Predicting life expectancy is not a new concept. Experts do this at a population level by classifying people into groups, often based on region or ethnicity.

Also, tools such as deep learning and artificial intelligence can be used to consider complex variables, such as biomedical data, to predict someone's biological age.

Biological age refers to how "old" their body is, rather than when they were born. A 30-year-old who smokes heavily may have a biological age closer to 40.

Calculating a life expectancy reliably would require a sophisticated system that considers a breadth of environmental, geographic, genetic and lifestyle factors – all of which have influence.

With machine learning and artificial intelligence, it's becoming feasible to analyse larger quantities of data. The use of deep learning and cognitive computing, such as with IBM Watson, helps doctors make more accurate diagnoses than using human judgement alone.

This, coupled with predictive analytics and increasing computational power, means we may soon have systems, or even apps, that can calculate life expectancy.

IBM cloud provide platform to make model using AUTO AI. I have made model without using a line of code then deployed it. Then using API key, instance key and URL a node red flow created. By sharing the URL of node-red flow it can be user friendly and can also use as fun.

THEORITICAL ANALYSIS

Block diagram
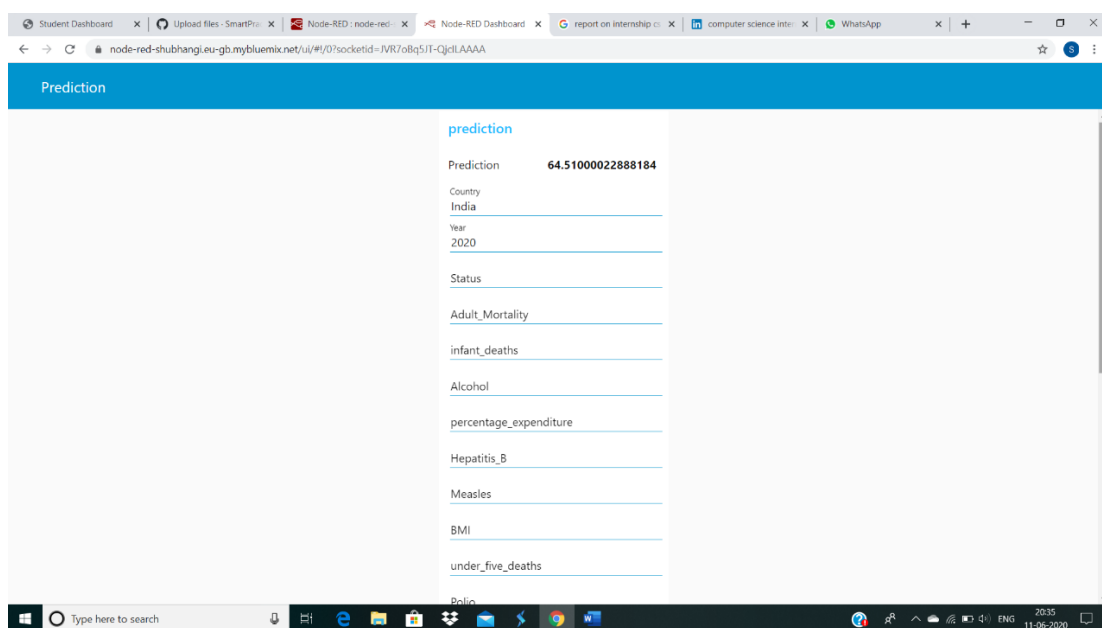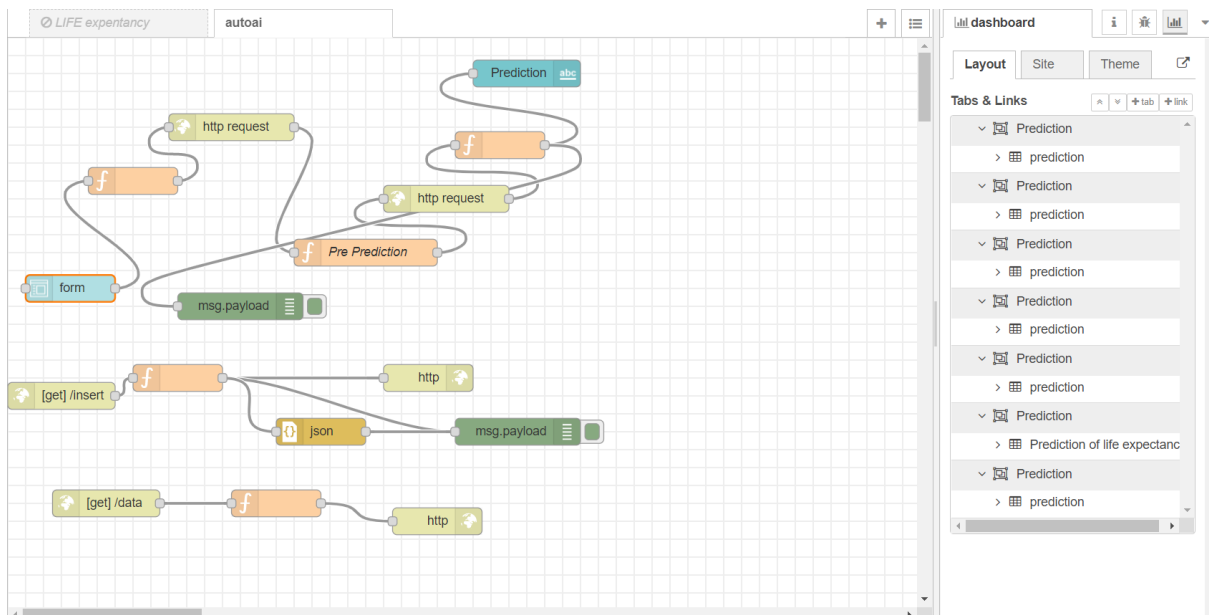
TRAINING
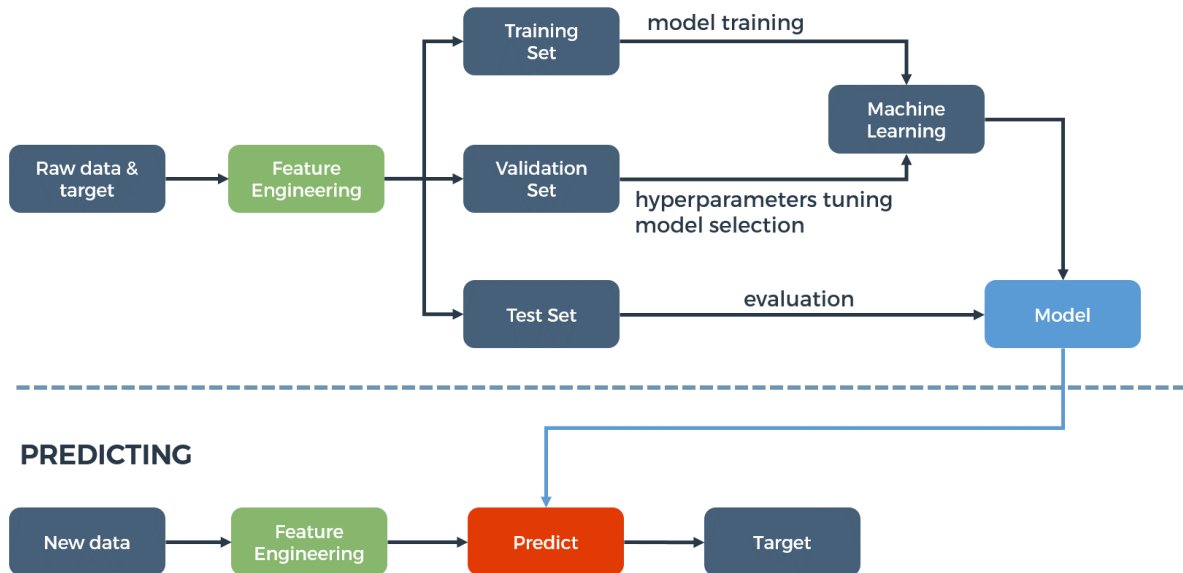
PREDICTING

Hardware / Software designing

A node red flow is made for the prediction of life expectancy. Which is given below.

https://node-red-shubhangi.eu-gb.mybluemix.net/ui/#!/0?socketid=JVR7oBq5JT-QjcILAAAA

Demonstration of node red flow

Auto AI model

EXPERIMENTAL INVESTIGATIONS

- Collection of data set from Kaggle.
  https://www.kaggle.com/c/predicting-life-expectancy/data
- On IBM Watson studio machine learning using auto ai build a model to predict life expectancy.
  - ➢ To do so first create account on IBM Watson studio.
  - ➢ Using Add to project choose auto ai.
  - ➢ Then upload data set
  - ➢ Choose best way to predict.
  - ➢ Save as a model which is on the top
  - ➢ Deploy the model.
  - ➢ Test the model.
  - ➢ Create service credential

- Create cloud foundry app
  https://node-red-shubhangi.eu-gb.mybluemix.net/red/#flow/c8ce04dc.84f0c8
- Make node red flow
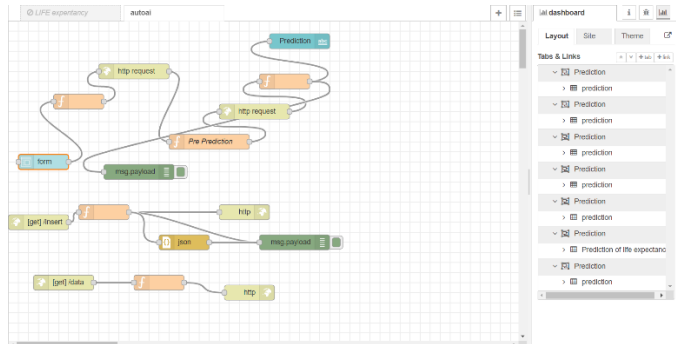
- Then add API key, instance Id and url.

- After deploying the model from dashboard UI can be seen.

FLOWCHART

TRAINING

RESULT

Prediction of life expectancy based on country, year, status, adult molarity, GDP and population etc.
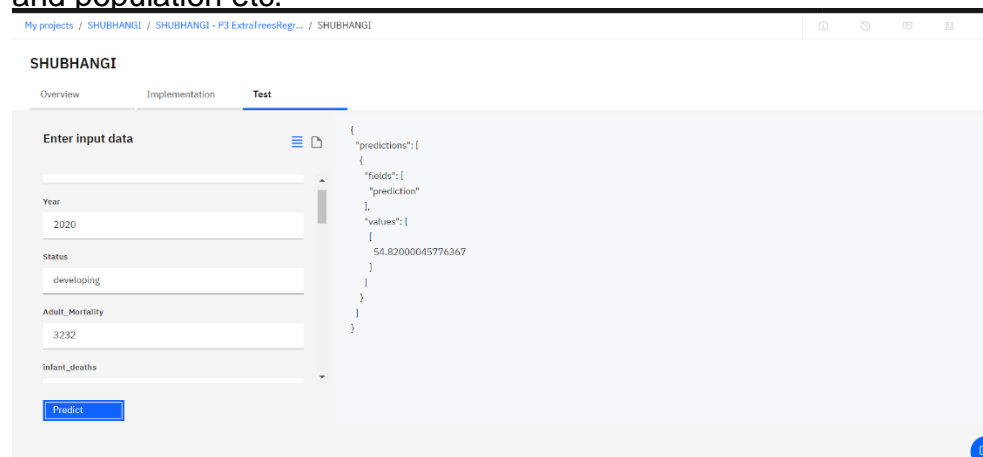
Application

Some interesting correlations here:

- There is a strong positive correlation between 'Schooling' and 'Life Expectancy'. This may be because education is more established and prevalent in wealthier countries. This means countries with less corruption, infrastructure, healthcare, welfare, and so forth.

- Similarly to the point above, there is a moderate positive correlation between 'GDP' and 'Life Expectancy', most likely due to the same reason.

- Surprisingly there's a moderate positive correlation between 'Alcohol' and 'Life Expectancy'. I'm guessing that this is due to the fact that only wealthier countries can afford alcohol or the consumption of alcohol is more prevalent among wealthier populations.

Similarly more result can be abstract.

Conclusion

I can honestly say that my time spent interning with Smartbridge resulted in one of the best summers of my life. I have gained new knowledge, skills and met many new people. I achieved several of my learning goals, however for some the conditions did not permit. I got insight into professional practice. I experienced that financing, as in many organisations, is an important factor for the progress of projects. Related to my study I learned more about how to make chat bot, database, and many more feature of IBM cloud. There is still a lot to discover and to improve. I gained a lot's of insight regarding almost every aspect. Bootcamps during this period was very helpful to complete project. All mentors were very supportive.

The internship was also good to find out what my strengths and weaknesses are. This helped me to define what skills and knowledge I have to improve in the coming time. It would be better that the knowledge level of the language is sufficient to contribute fully to projects.

At last this internship has given me new insights and motivation to pursue a career in machine learning.

BIBLIOGRAPHY

How to create a Kick Off Meeting Agenda

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

GitHub account creation

https://github.com/

https://www.youtube.com/watch?v=7YUTc4Cigc8&feature=youtu.be

Learn how to add, edit, delete text using Writer.

https://www.zoho.com/writer/help/working-with-text.html

IBM Academic Initiative account

https://my15.digitalexperience.ibm.com/b73a5759-c6a6-4033-ab6b-d9d4f9a6d65b/dxsites/151914d1-03d2-48fe-97d9-d21166848e65/

create Node-red application

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

https://www.youtube.com/watch?v=s7wmiS2mSXY&feature=youtu.be

https://www.w3schools.com/howto/howto_make_a_website.asp

https://www.youtube.com/

APPENDIX

https://node-red-shubhangi.eu-gb.mybluemix.net/ui/#!/0?socketid=JVR7oBq5JT-QjcILAAAA

https://eu-gb.dataplatform.cloud.ibm.com/ml/deployments/80718439-26b1-4c0f-8e5a-4df15688e25e/test?projectid=550379ee-8e2a-4bcf-9567-8a62b9419f46&mlInstanceGuid=98168165-ebfd-4004-82ce-936211a17656&context=wdp&flush=true&wmlv4=true