Predicting Life Expectancy using Machine Learning

Project Summary

Life expectancy is the statistical age that a person is expected to live, based on the year of its birth, its current age and many other demographic factors. It has many uses in the financial world, including life insurance and pension planning. Life expectancy is the single most influential factor that insurance companies use to determine life insurance premiums. It is also critical in retirement planning as many aging workers arrange their retirement plans' asset allocations based on a prediction of how long they expect to live.

The project will predict the Life Expectancy using **Machine Learning** on account of various factors such as **Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors.**

The Project will use **IBM Cloud services, IBM Watson Studio ad Python** to create a model, automate it and deploy it. **Node-Red** will also be used to integrate the endpoints created.

Project Requirements

1) Functional Requirements

The project is required to **predict the life expectancy** of people living in a country with maximum accuracy 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.

2) Technical Requirements

- Python programming language.
- IBM Watson studio for data processing and creating model.
- IBM Node-RED to develop front end of the project.

3) Software Requirements

Python, IBM Cloud Services, IBM Watson Studio, Internet Browser(preferably Google Chrome), Node-Red

Project Deliverables

- Project Documentation
- Notebook of ML Life Prediction model
- Watson Studio AutoAI ML model
- Node-Red Flow
- Working website of model

Project Team

1. Shrey Agarwal (Developer)
2. TheSmartBridge (Project Sponsor)

Project Steps

1) Collecting Dataset

For the project we are using the "Life Expectancy Data" by **WHO** which provides a statistical analysis on factors influencing Life Expectancy. The dataset contains different factors for 193 countries from a period of 2000-2015.

2) Setting up IBM Cloud Services

We setup an IBM Cloud account to access various services –

- IBM Watson Studio
- Node-RED
- IBM Machine Learning Services

These services are used for project development.

3) Creating a Watson Project & Machine Learning Services

We create a Project in Watson Studios in which we develop the project by creating assets and incorporating various services.

4) Creating Jupyter Notebook and importing Dataset

- We create a Jupyter notebook(the platform to construct the ML model) and import our dataset into the project.
- We choose the correct Regression model for our dataset (In this project we use a Linear Regression Model) and then train out model on the dataset after processing it.
- We then deploy the model in Watson studios. Endpoints are created using WatsonMachineLearningAPI.

5) Creating Node-RED Flow and integrating ML services

We create a flow using Node-RED to create the front end of the model and then integrate it with the machine learning services using API.

Conclusion

We created a project which successfully predicts the life expectancy of 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 submitted.