PROJECT SCOPE DOCUMENTATION

Project Title: PREDICTING LIFE EXPECTANCY USING MACHINE LEARNING
kick off date: 20.05.2020

Project summary:

      Life expectancy is a statistical prediction for how long a person will live.Based on actuarial science, life expectancy takes into account several individual-level as well as population-level factors to arrive at a figure.

      This project involves building a machine learning model to predict the life expectancy rate of a country given various features like year, GDP, education, alcohol intake of people in the country, expenditure on healthcare system and some specific disease related deaths that have happened.A typical Regression Machine Learning project leverages historical data to predict insights into the future.

      It works on the data set provided by the World Health Organization(WHO) to evaluate the life expectancy in different years of a country. The time frame offered in the data set is from the year 2000 to 2015.

      Some of the Regression based algorithms that have been used to predicting the life expectancy for the data that hasn't been used in training the model are:
- Linear Regression
- Linear Regression with Polynomic Feature
- Random Forest Regression
- Decision Tree Regression
- Ridge Regression
- Lasso Regression
- Elastic Net Regression.

      The model that involves Python to code these regression techniques uses the Jupyter Notebook in IBM Watson Studio to import data and automate the ML model otherwise the IBM Watson Studio services are used to auto AI the experiment.A Node - RED flow is built to integrate the ML services or Auto AI.

Project Requirements:

The following are needed to successfully develop this project :

1. A dataset comprising of the countries and factors influencing the life expectancy there.
2. A regression based machine learning model coded in Python.
3. A Git Hub account.
4. An IBM Cloud account.
5. A jupyter notebook in IBM Watson Studio to develop the Python codes.
6. A Node-RED flow application to integrate the model.

Functional Requirements:

The following are the functionl requirements of this project:

1. A supervised Machine learning model based on Regression written in Python to process the dataset for the desired output.
2. An IBM Cloud service is to store the required data.
3. An IBM Watson Studio service to create an ML model and automate it.
4. A Node-RED flow to integrate ML services.

Software Requirements:

The following are the software requirements of this project:

1. Jupyter Notebook for developing python codes in
2. Git Tool
3. IBM cloud account
4. IBM Watson Studio service
5. Node-RED flow application

Project Deliverables:

A machine learning model that will be able to predict the life expectancy of a country given an input data comprising of various factors influencing it.

Project Team:

This is an individual project developed by Kishore R.

Project Schedule:

Tasks	Duration
1. Setup the Development Environment	1 Day
2. Create IBM Cloud Account	0.5 Day
3. Create a Node-RED starter application	1 Day
4. Explore IBM uses cases and IBM Watson Machine Learning	3 Days
5. Build an ML model in IBM Watson Studio	2 Days
6. Automate the ML Model	1 Day
7. Collect dataset for the project	0.5 Day
8. Create the IBM Cloud service	1 Day
9. Create a Watson Studio project	1 Day
10. Configure Watson studio and create Machine Learning service	1 Day
11. Create A Jupyter Notebook In IBM Watson And Import Data	0.5 Day
12. Build A Machine Learning Model and Create Endpoints For Node-RED Integration.	2 Days
13. Build Node-RED Flow to Integrate ML Services	2 Days
14. Create AutoAI experiment	1 Day
15. Build Node-RED Flow to Integrate AutoAI	1 Day