**Project Scope:**

* **Project Summary:**

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.

The project tries to predict evaluate the life expectancy for different countries in years. The data offers a timeframe from 2000 to 2015. The link to the DataSet required is here: <https://www.kaggle.com/kumarajarshi/life-expectancy-who> . The output algorithms have to be test if they can maintain their accuracy in predicting the life expectancy for data they haven’t been trained.

* **Project Requirements:**

The technologies required for the project are Python, IBM studio and IBM Watson. This project is supposed to be done within 29days. All the supporting requirements are provided in the dashboard section, One can easily understand and use them. Candidate with prior knowledge of Machine Learning can start working on the project by his own. For candidates without prior knowledge of Machine Learning, supporting links, Slack channel connectivity and timely Zoom meeting sessions are being conducted. The dataset required is also provided by the link.

* **Functional Requirements:**

The function and the flow of the project are quite simple and clear

The project commences with the project scope. A precise scope for the project is to be made having explanations on the following topics.

Prepare Project Scope Document with following Headings

* Project Summary
* Project Requirements
* Functional Requirements
* Technical Requirements
* Software Requirements
* Project Deliverables
* Project Team
* Project Schedule

After this setting up the development environment is mandatory. Into this, the candidate has to connect digitally with the project mentor and to submit his/her work using the technologies such as GitHub, Slack and the Zoho writer.

The candidate has to create account on the IBM cloud and need to access node-red application.

* **Technical Requirements:**

We define the task to solve as follows  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.

We need to optimize the model architecture and feature set, and test the performance of several models. The following sections describe:

* the dataset
* the train-validation-test split
* our methods for creating the input data for the model
* our methods for determining the model architecture
* our methods for feature selection
* the evaluation protocol.

A high-level approach may also be used for health determination attribute classifications:

* Individual lifestyle behaviours e.g., spending patterns, exercise, diet;
* Physical and social environments e.g., living density, pollution levels;
* Socioeconomic factors e.g., education level, financial status;
* Health outcomes e.g., illnesses;
* Health systems e.g., health insurance status.
* GDP.
* Sex differences.
* Regional variations.
* **Software Requirements:**

The software required throughout the project are quite easy to use. The softwares required are:

* Python IDE
* IBM Watson
* IBM cloud
* Node-red application
* Zoho writer
* Slack channel
* Auto-Ai (non-technical candidate)

The links to learn and master the above softwares are provided in the dashboard.

The IBM Cloud platform is built to support your needs whether it's working only in the public cloud or taking advantage of a multicloud deployment model. With our open-source technologies, such as Kubernetes, Red Hat OpenShift, and a full range of compute options, including virtual machines, containers, bare metal, and serverless, you have as much control and flexibility as you need to support workloads in your hybrid environment. You can deploy cloud-native apps while also ensuring workload portability.

Node-RED is a programming tool for wiring together hardware devices, APIs and online services in new and interesting ways. It provides a browser-based editor that makes it easy to wire together flows using the wide range of nodes in the palette that can be deployed to its runtime in a single-click.

Infuse IBM Watson into your apps and workflows to tap into organizational data and put AI to work across multiple departments – from finance, to customer care, to supply chain. With Watson, you can create better, more personalized experiences for customers, scale the expertise of your best people across the organization, and make smarter decisions based on deep insights from data.

* **Project deliverables:**

Whilst there have been a few attempts to calculate the Life Expectancy for classified groups and diseases, little has been done to foresee and predict generic or personalized life expectancy, and could be attributed to several contributing factors.

1. The overwhelming number of variables that could be considered in predicting LE;
2. Lack of accumulated data in one storage location for data processing and analysis to generate meaningful data;
3. Difficulty of centralizing heterogeneous networks from different countries and regions;
4. Unpredictable and fast changing lifestyle of humans with the increase of sophisticated technologies;
5. Limited methods of health data collection such as data only from patients in a healthcare environment (hospitals, insurance companies), which may exclude the general population.

The project model will provide an individually customized life expectancy prediction. This will also review existing works and techniques in the prediction of human LE, and reach to a conclusion that it is feasible to predict a PLE for individuals using evolving technologies and devices such as big data, AI, machine learning techniques, and PHDs, wearables and mobile health monitoring devices. This model tends to make rather pessimistic predictions, while doctors tend to do the opposite. Pessimistic predictions could promote early recognition and anticipation of the palliative phase, and timely discussion of ACP strategies.