**Project Goals:**

1. This project aims to predict the expectancy of life given dataset using 21 features.
2. The task is to be done using Python.
3. The process involves the application of regression.
4. The project also aims to optimize the parameters of the model.

**Assumptions:**

1. Regression can be defined as a method or an algorithm in Machine Learning that models a target value based on independent predictors. It is essentially a statistical tool used in finding out the relationship between a dependent variable an independent variable.

In regression, we assume that there is a linear function of features. Therefore, it is aimed to find the coefficient of features.

1. We also assume that the given data is sufficient enough to predict coefficients.
2. We also assume that the test set will belong to some distribution as train set.

**Constraints:**

Features are linearly dependent on the dependent variable.

**Major Deliverables:**

1. WEEK 1:

* The project aims to deliver the goals with above mentioned assumptions. For this, the structure of model is to be crafted. The relevant details will be included in the Project Scope Document and will be submitted by the end of the first week.

1. WEEK 2:

* The second week's report will include the following:
* Exploration of IBM Cloud and Watson services.
* Creating a Node-Red starter application in IBM cloud.
* Overview of IBM Watson machine learning platform.

1. WEEK 3:

* This week’s report will include the details of the model and the details of the tools and modules to be used.

1. WEEK 4:

* The fourth week's report will include the following:

Collection of Life Expectancy dataset.

Creation of necessary IBM cloud services.

Creation of Machine Learning project using IBM Watson Studio by importing the dataset into its Jupyter Notebook.

At last creation of endpoints for Red-Node Integration to integrate Machine Learning services.

**Requirements:**

Life expectancy dataset. Download link:

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

IBM Cloud/IBM Watson Studio.

Sklearn Library: already contained in Jupyter Notebook.

Visualisation tools: matplotlib library is used.

**Key Milestones:**

|  |  |
| --- | --- |
| **Report** | **Date** |
| Report 1 | 26/05/2000 |
| Report 2 | 02/06/2000 |
| Report 3 | 09/06/2000 |
| Report 4 | 16/06/2000 |