**Internship Title: Predicting Life Expectancy using Machine Learning - SB44471**

**Name: Samruddhi Shivaji Kshirsagar**

**Project\_id: SPS\_PRO\_215**

**Project Title: Predicting Life Expectancy using Machine Learning.**

**I. INTRODUCTION:**

There have been lot of studies and researches to notice the factors that are affecting life expectancy of a country considering dempographic variables,income composition and mortality rates. The WHO releases the data for var various countries and the reasons that are affecting the people with different factors like mortality,early deaths,polio etc.

Various Models of Machine Learning are used for the prediction, each prefer the different and the convinient model for the prediction of the life expectancy. Since the observations this dataset are based on different countries, it will be easier for a country to determine the predicting factor which is contributing to lower value of life expectancy. This will help in suggesting a country which area should be given importance in order to efficiently improve the life expectancy of its population.

**1.1 Overview:**

The project is based on on the Machine Learning concepts and the Python language. In this project the "Life Expectancy" dataset is used that is provided by WHO. This dataset contains the data of about 193 countries and various factors related to it. This dataset is described and then tested for the prediction of life expectancy. In this project I have been using the Forest Regressor Model of Machine Learning for the prediction and getting the accuracy of the project.

**1.2 Purpose:**

The purpose behind the project is to study the data of the Life Expectancy and analyze it. With the help of dataset and the elements of it like the various diseases worldwide like Polio,BMI,HIV, etc. which affect the life of the people we will predict the life expectancy of the various countries.

**2. LITERATURE SURVEY:**

Machine learning (ML) is the study of computer algorithms that improve automatically through experience.ML helps in analysis of the data. Various research work and the projects have been done for the prediction of life expectancy. Each one has its own solution which they prefered. To reach to a solution one can follow different ways. Every one working on this project propose solution by following various regression pattern like linear regression,decision tree regressor,forest regressor etc base on the accuracy they get or the way they prefer.

**2.1 Existing Problem:**

The death rate of the people across the world has been increasing . There are various different factors that are affecting it. Thus it is difficult to to understand the actual reason behind the increasing death rate. Thus we must predict the the reason which is causing the high death.

**2.2 Proposed Solution:**

We will be building a Machine learning (ML) model which will help in predicting the life expactancy of the the people.

This will we done with the help of given dataset of life expectancy that holds the data for 193 countries.

We will be using the Forest Regressor Model for the prediction of Life Expectancy.

**3.THEORITICAL ANALYSIS:**

Therotical Analysis is all about the actual implementation of the the project. It includes the stages of the development phase. As we are building the Machine Learning Model using the IBM cloud services and the Watson Studio It is primarily must that we must know about the following things:

1. Explore the IBM Cloud Services.

2. Learn about the Watson Studio and services asscociated with it.

3. Explore Use cases for the Machine Learning.

4. Use of Jupyter Notebook.

5. Automate your model in the Jupyter nb.

6. Deploy your model

**3.1 Block Diagram:**

For any project development we need too follow various stages that will help in building a project in the proper and sequential form. Thyus we will be describing the whole project development process with the help of the block diagram.

Train and Test the data using Forest regressor model

Import the dataset(Life Expectancy) in Jupyter Notebook

Enable the IBM Cloud ML Services

Deploy the model in Ibm Cloud using ML Watson credentials

Create the node red service and integrate the ML model

Create a UI using Node red flow and test the model

**3.2 Software/Hardware Design:**

It includes all the hardware or the software components used in the project development process.Thus following are the required softwares included for our project development.

1. Watson Studio
2. Jupyter Notebook
3. Machine Learning Services
4. IBM CLoud Foundary
5. Node-Red

No hardware component is involved in the project as all the project is the software developed.Thus we need only software design for the project. The project is developed with the hep of IBM Cloud Service.

**4. RESULT:**

After the model is developed and deployed we are able to calculate the life expectancy according the factors in the the dataset.

The prediction we get are dependent on the the Machine learning model we choose. The predection gives the accuracy based on the percentage of error we get by training and testing the data in the model.

**Accuracy: R^2: 0.89**

**Mean squared error: 6.24**

**Mean absolute error: 1.97**

**5. Advantages:**

- With help of this model we can get the predection of life expectancy across different countries.

- We can find the cause of death rate due to which the life has been reduced.

- The data is predefined thus the testing of the model has been easy.

**6.Disadvantages**:

- As dataset is predefine the model will not be able to calculate the data beyond it.

- It may give error if we try to predict for the cases outside dataset.

**7.CONCLUSION:**

We are sucessfully able to create a Machine Learning model using Forest Regressor to get the prediction of the model.

We get the minimum absolute error using Forest Regressor Model which leads to the high accuracy of the prediction.

**8. FUTURE SCOPE:**

In future one can build a Machine Learning Forest Regressor model by using the real time data for calculating the predictions. This will help in solving the real-time problems.

**9. BIBLOGRAPHY :**

**1.Developer IBM Building Models using Watson Studio**

<https://developer.ibm.com/technologies/artificial-intelligence/tutorials/watson-studio-using-jupyter-notebook>

**2. Kaggle Life Expectancy using WHO**

<https://www.kaggle.com/harshini564/ab-testing-life-expectancy-who>

**3. Random Forest Regressor**

<https://towardsdatascience.com/random-forest-in-python-24d0893d51c0>