

# **PROJECT REPORT DOCUMENT**

# **PREDICTING LIFE EXPECTANCY USING MACHINE LEARNING**

Submitted in fulfillment for the completion of  
the project at Smartbridge



Submitted by

VALAVALA VENKATA LAKSHMI GAYATRI

# 1.INTRODUCTION

## 1.1.Overview

Life expectancy is a statistical measure of the average time a human being is expected to live. Life expectancy depends on various factors including: financial status, regional variations, economic circumstances, gender differences, mental illnesses, physical illnesses, education, year of their birth, demographic factors and many more.

This project is aimed at predicting Life Expectancy rate of a country, given various factors, using Machine Learning algorithms.

In this project, we have taken the dataset related to life expectancy. Health factors for 193 countries has been collected from the WHO data repository website and its corresponding economic data was collected from the United Nations website.

## 1.2.Purpose

Since the observations 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.

# 2.LITERATURE SURVEY

## 2.1.Existing Problem:

Life expectancy is a statistical measure of the average time a human being is expected to live. This is based on various health, education and economic factors.

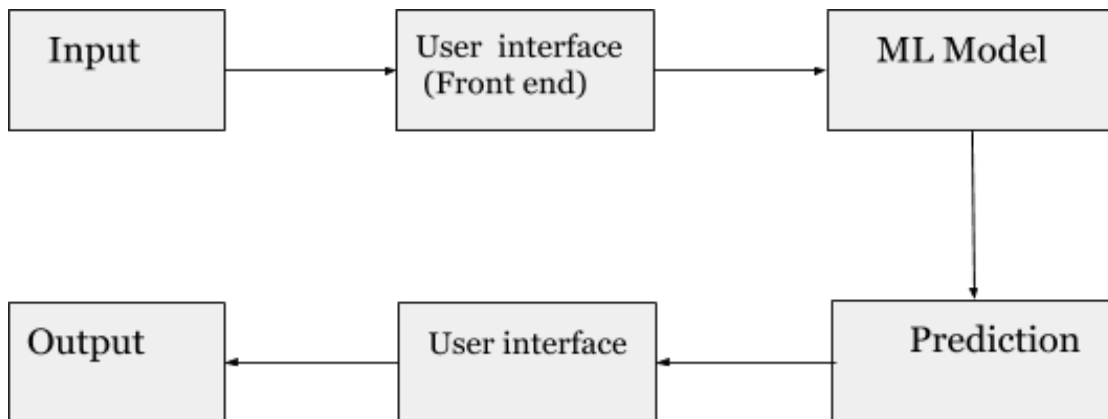
Although there have been a lot of studies undertaken in the past on factors affecting life expectancy considering demographic variables, income composition and mortality rates. It was found that the effect of immunization and human development index was not taken into account in the past.

## 2.2.Proposed Solution:

Design a regression model that takes into consideration all the available factors to predict accurate average life expectancy

### 3.THEORETICAL ANALYSIS

#### 3.1.Block Diagram



#### 3.2 Software /Hardware designing:

SOFTWARE REQUIREMENTS:-

- IBM Cloud
- IBM Watson Studio
- Node-red App

This problem statement is aimed at predicting average life expectancy of people living in a country given various features.

Firstly, the required data set is collected from <https://www.kaggle.com/kumarajarshi/life-expectancy-who>. Secondly, an IBM Watson studio project is created in the Watson Studio service, provided by IBM Cloud. Then, the data set is imported into this project and a machine learning model is formulated. This model can be opened in the Watson Studio project and tested giving the inputs.

A User Interface (UI) is designed for the project using Node-RED. In this User Interface the users can type in data regarding their country's gdp, alcohol intake, diseased count, deaths count, income and expenditure, education level, etc. and get the output which is the predicted life expectancy of the country.

### 4.EXPERIMENTAL INVESTIGATIONS:

1. Choose a Project Idea:

Predicting Life Expectancy of a person.

2. Conduct Background Research:

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

<https://ourworldindata.org/life-expectancy-how-is-it-calculated-and-how-should-it-be-interpreted>

### 3. Compose a Hypothesis:

Based on our study and information gathered we can predict the average age of a person.

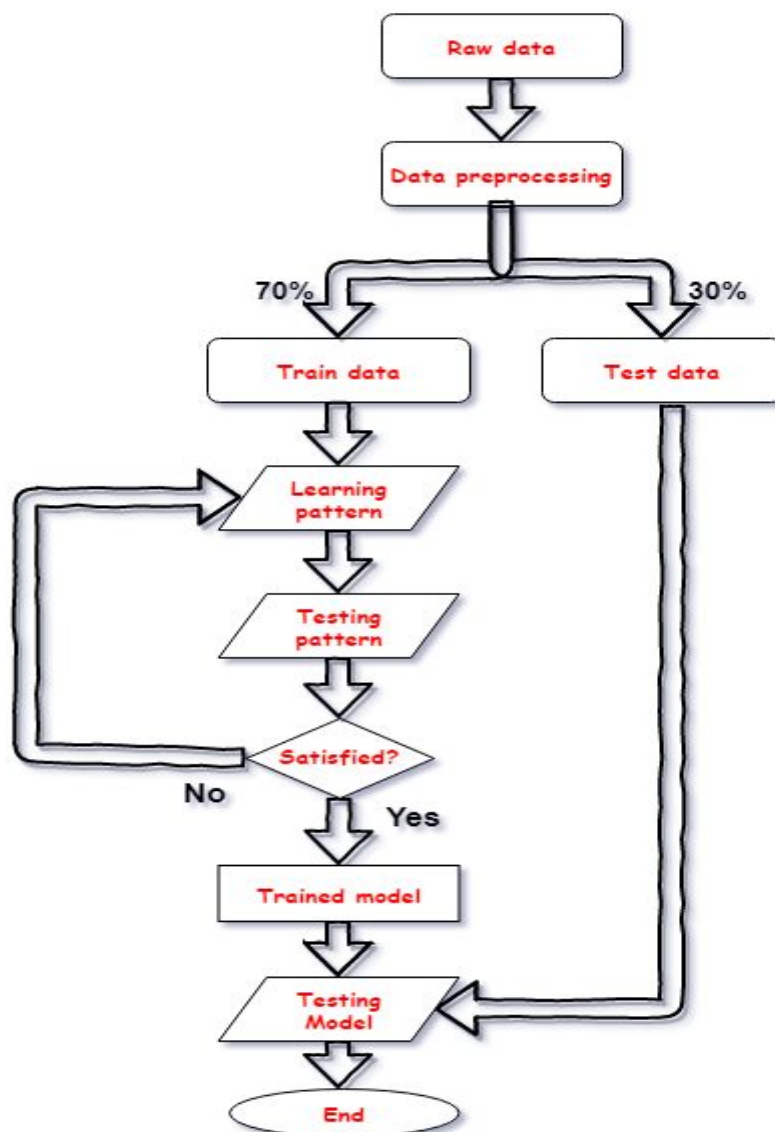
### 4. Design your Experiment:

First we need to collect the suitable data for our problem statement. Next we need to construct the regression model for this problem.

### 5. Draw Conclusions:

After construction of our model, we can predict the average life expectancy of a person.

## 5.FLOW CHART:



# 6.Ouput:

Machine Learning Model

Prediction

Adult Mortality

infant deaths

Alcohol

percentage expenditure

Hepatitis B

Measles

BMI

under 5 deaths

Polio

Total expenditure

Diphtheria

HIV/AIDS

GDP

Population

thinness 1-19 years

thinness 5-9 years

Income composition of resources

Schooling

☐ Developing Country

SUBMIT

CANCEL

Student Dashboard x

Predicting Life Exp x

Node-RED : node- x

Node-RED Dashbo x

Life Expec

net/ui/#!/0?socketid=QzJd2UVz\_XRH2vM1AAAF

Machine Learning Model

Prediction43.8

Adult Mortality275

infant deaths71

Alcohol0.01

percentage expenditure7.097109

Hepatitis B68

Measles3013

BMI17.3

under 5 deaths100

Polio68

Total expenditure

# CONCLUSION:

The end product is an web page created and deployed on node-red app of IBM cloud. The backend of web page is a Decision Tree Regressor Model with 92.4 R2 score created and deployed on watson studio using machine learning service.

The web-page has input fields similar to dataset columns such as Country, BMI, percentage expenditure, Alcohol etc and an output field named as prediction i.e similar to dataset column Life expectancy which gives the life expectancy prediction based on the given values.