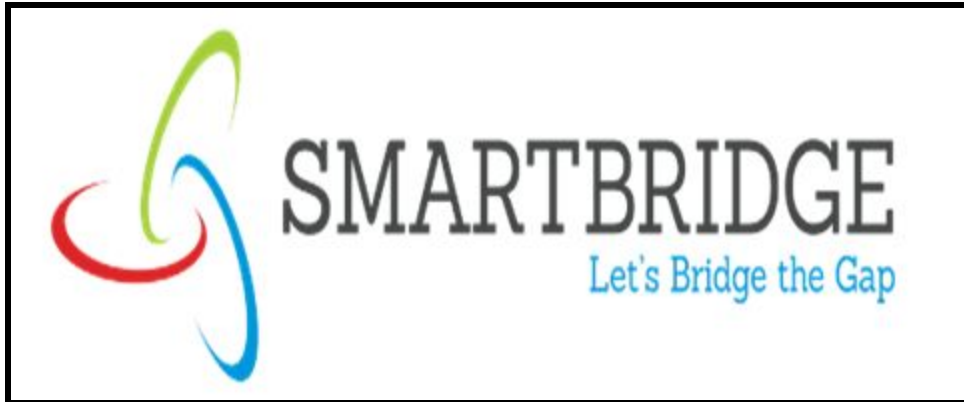


**Internship Title**  
**Predicting Life Expectancy using Machine Learning**



By:

T Likitha

3rd Year - B.Tech

Information Technology

Gayatri Vidya Parishad College of Engineering(A)

## Table of Contents:

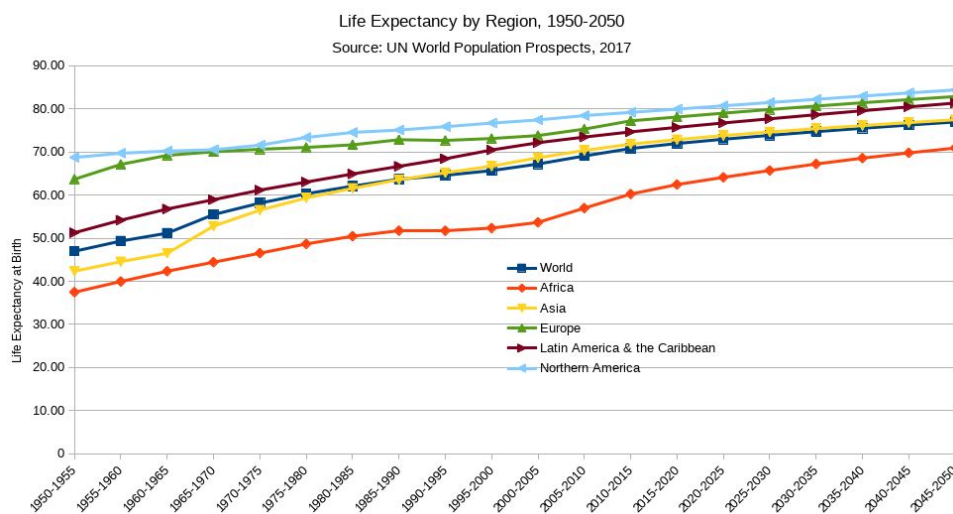
<b>1</b>	<b>INTRODUCTION</b>
	1.1 Overview
	1.2 Purpose
<b>2</b>	<b>LITERATURE SURVEY</b>
	2.1 Existing problem
	2.2 Proposed solution
<b>3</b>	<b>THEORETICAL ANALYSIS</b>
	3.1 Block diagram
	3.2 Hardware / Software designing
<b>4</b>	<b>EXPERIMENTAL INVESTIGATIONS</b>
<b>5</b>	<b>FLOWCHART</b>
<b>6</b>	<b>RESULT</b>
<b>7</b>	<b>ADVANTAGES &amp; DISADVANTAGES</b>
<b>8</b>	<b>APPLICATIONS</b>
<b>9</b>	<b>CONCLUSION</b>
<b>10</b>	<b>FUTURE SCOPE</b>
<b>11</b>	<b>BIBLIOGRAPHY</b>
<b>12.</b>	<b>APPENDIX</b>

# 1.Introduction

## 1.1 Overview

The project titled ‘Predicting Life Expectancy using Machine Learning’ aims to predict the life expectancy of several countries depending on various factors.

The term “life expectancy” refers to the number of years a person can expect to live. By definition, life expectancy is based on an estimate of the average age that members of a particular population group will be when they die. In formulaic terms, life expectancy is denoted by  $e_x$ , where, “e” represents the expected number of years remaining and “x” represents the person’s present age. The life expectancy for a particular person or population group depends on several variables such as their lifestyle, access to healthcare, diet, economical status and the relevant mortality and morbidity data.



The above figure depicts the changes in Life Expectancy rates over the years

## **1.2 Purpose**

The project aims to determine the Life Expectancy of different countries based on various factors. Good prognostication for example helps to determine the course of treatment and helps to anticipate the procurement of health care services and facilities. This inturn will help in better planning and thus improved lifestyles.

## **2.Literature Survey**

### **2.1 Existing System**

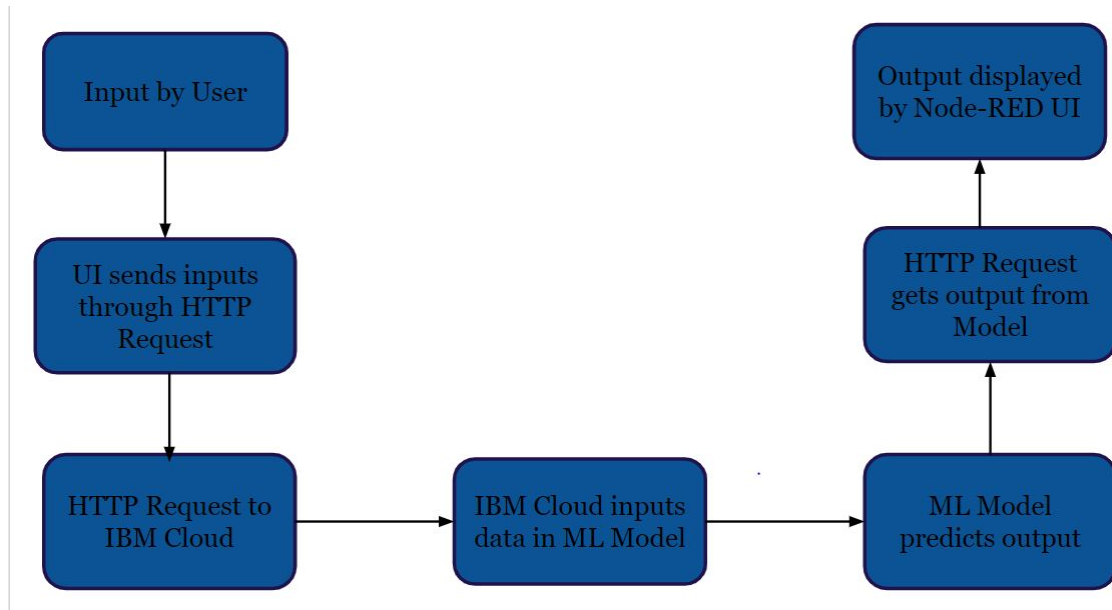
Although there have been lot of studies undertaken in the past on factors affecting life expectancy considering demographic variables, income composition and mortality rates. It was found that effect of immunization and human development index was not taken into account in the past. Thus, other factors have to be considered and also data related to this in the past years i.e. 2000-2015.

### **2.2 Proposed System**

The proposed system considers many factors like geographical condition, literacy, immunization, mortality factors and economic conditions etc. The data collected is by WHO for the years 2000-2015. Regression model of Random Forest Algorithm is used for better accuracy.

## **3.Theoretical Analysis**

### **3.1 Block Diagram**



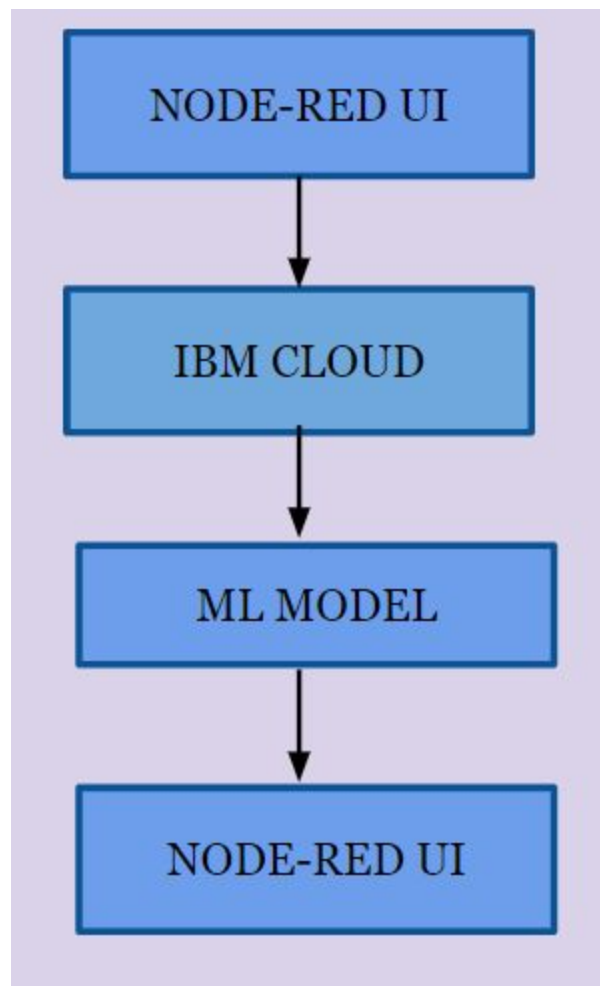
### **3.2 Software Designing**

Input is taken from the user using a “Form” element in Node-Red. Then, an HTTP request is made to the IBM cloud that further makes an HTTP request to the deployed model using model’s instance id. After verification of id, the model sends an HTTP response which is finally parsed by the Node-Red application and the result is displayed on the user screen.

## **4.Experimental Investigation**

Experimental investigation for the dataset was done. Later, the dataset was thoroughly researched to find the factors affecting the life expectancy for the region. Also software like Node-RED was also experimented with.

## **5.Flowchart**



## 6. Result

Home

Predicting Life Expectancy

Prediction 56.38999999999999

Status 0

Adult Mortality 222

Alcohol 0.2

percentage expenditure 73.89

Hepatitis B 76

BMI 18.2

under-five deaths 93

Polio 67

Total expenditure 8.3

Diphtheria 65

HIV/AIDS 0.3

GDP 63.54

thinness 1-19 years 17.5

thinness 5-9 years 17.2

Income composition of resources 0.454

Schooling 10

SUBMIT CANCEL

## 7. ADVANTAGES AND DISADVANTAGES

### Advantages

- Helpful for predicting life expectancy
- Helpful in medical fields
- Helpful in analysing factors affecting life expectancy
- Random forest regressor is stable and also has less noise.

## Disadvantages

- Data can be wrong.
- Accuracy is 93%.
- The UI requires internet connection.

## **8.Applications**

- The project can be used as a basis to develop personalized health applications.
- As the model uses a wide range of features for prediction, 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
- The government can plan and develop their health infrastructures by keeping the most correlated factors in mind.
- The project can help governments to keep track of their countries' health status so they can plan for the future accordingly.



## **9. Conclusion**

The project thus predicts the life expectancy rate given various features. The model has an accuracy of 93% life expectancy rate of different countries depending on various factors. The model has an accuracy of 92.5%. This was then implemented on the NODE-RED using IBM Cloud service. The project makes a good use of machine learning in predicting life expectancy of a country that can help respective government in making policies that will serve for the benefit of the nation and entire humankind.

## **10. Future Scope**

Some future improvements can be made on current project. They are as follows:

1. As more data comes, that can be fed to the model for more accurate predictions.
2. Currently, the project is just a web application. It can be developed to support other platforms like Android, IOS and Windows Mobile.
3. Other regression models can also be used for prediction and later the best among them should be chosen.

4. User interface can be modified for various countries according to their regional languages.

## **11. Bibliography**

- <https://www.youtube.com/>
- Dataset: <https://www.kaggle.com/kumarajarshi/life-expectancy-who>
- IBM Tutorials: <https://developer.ibm.com/tutorials/>
- Node-RED Tutorials: <https://nodered.org/docs/tutorials/>

# Appendix

- Code

[https://github.com/SmartPracticeschool/llSPS-INT-1629-Predicting-Life-Expectancy-using-Machine-Learning/blob/master/Life%20Expectancy%20%20\(2\).ipynb](https://github.com/SmartPracticeschool/llSPS-INT-1629-Predicting-Life-Expectancy-using-Machine-Learning/blob/master/Life%20Expectancy%20%20(2).ipynb)

- Node-RED Flow: <https://node-red-tmoyt.mybluemix.net/ui/>