# PROJECT REPORT

# **PROJECT STATEMENT:** PREDICTING LIFE EXPECTANC USING MACHINE LEARNING

**PROJECT BY:** Nakirekanti Tejaswini

B.Tech IIIrd year

Mahtama Gandhi Institute of technology,Hyderabad

**1 INTRODUCTION**

1.1 Overview

1.2 Purpose

**2 LITERATURE SURVEY**

2.1 Existing problem

2.2 Proposed solution

**3 THEORITICAL ANALYSIS**

3.1 Block diagram

3.2 Hardware / Software designing

**4 EXPERIMENTAL INVESTIGATIONS**

**5 FLOWCHART**

**6 RESULT**

**7 ADVANTAGES & DISADVANTAGES**

**8 APPLICATIONS**

**9 CONCLUSION**

**10 FUTURE SCOPE**

**11 BIBILOGRAPHY**

**APPENDIX**

A. Source code

1. **INTRODUCTION**

**1.1** **Overview**

Life expectancy is the key metric for assessing population health. The average time a human being is expected to live. Life expectancy depends on various factors like Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors. This project provides a way to 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.

For prediction purpose, the program uses Machine Learning Regression Algorithm. The user interface of the project was developed by using Node-Red application. The project also uses other IBM services like IBM Cloud, IBM Watson Studio, IBM Machine Learning Service and Cloudant.

**1.2 Purpose**

There are several approaches to estimate the life expectancy of a country by considering various factors. We build a regression model based on data provided by the World Health Organization (WHO) to evaluate the life expectancy for different countries in years. The data offers a timeframe from 2000 to 2015.

For better life expectancy rate we consider important immunization factors like Hepatitis B, Polio and Diphtheria and also regional variation, economic factors, education, expenditure on healthcare system and some specific disease related deaths that happened in the country, GDP, population. These factors helps to find the accurate expectancy rate.

1. **LITERATURE SURVEY**
   1. **Existing Problem**

Life expectancy is calculated in different ways. Generally Life expectancy is calculated by incorporating data on age-specific death rates for the population, which requires enumeration data for the number of people, and the number of deaths at each age for that population and from them the average life expectancy for each of the age groups within the population can be calculated.

But the results will be accurate if the collected data is complete. 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.

* 1. **Proposed Solution**

This project builds a regression model based on data provided by the World Health Organization (WHO) to evaluate the life expectancy. It has been observed that in the past 15 years, there has been a huge development in health sector resulting in improvement of human mortality rates especially in the developing nations in comparison to the past 30 years. Therefore, in this project we have considered data from year 2000-2015 for 193 countries for further analysis. This study will focus on immunization factors like Hepatitis B, Polio and Diphtheria, mortality factors, economic factors, social factors and other health related factors as well.

The output of this project helps the countries to draw the insights and develop the factors influencing the rate and build it to improve the expectancy rate.

1. **THEORITICAL ANALYSIS**
   1. **Block diagram**

Collecting the Dataset

Creating necessary IBM services

Creating Jupyter notebook in IBM Watson and importing the dataset

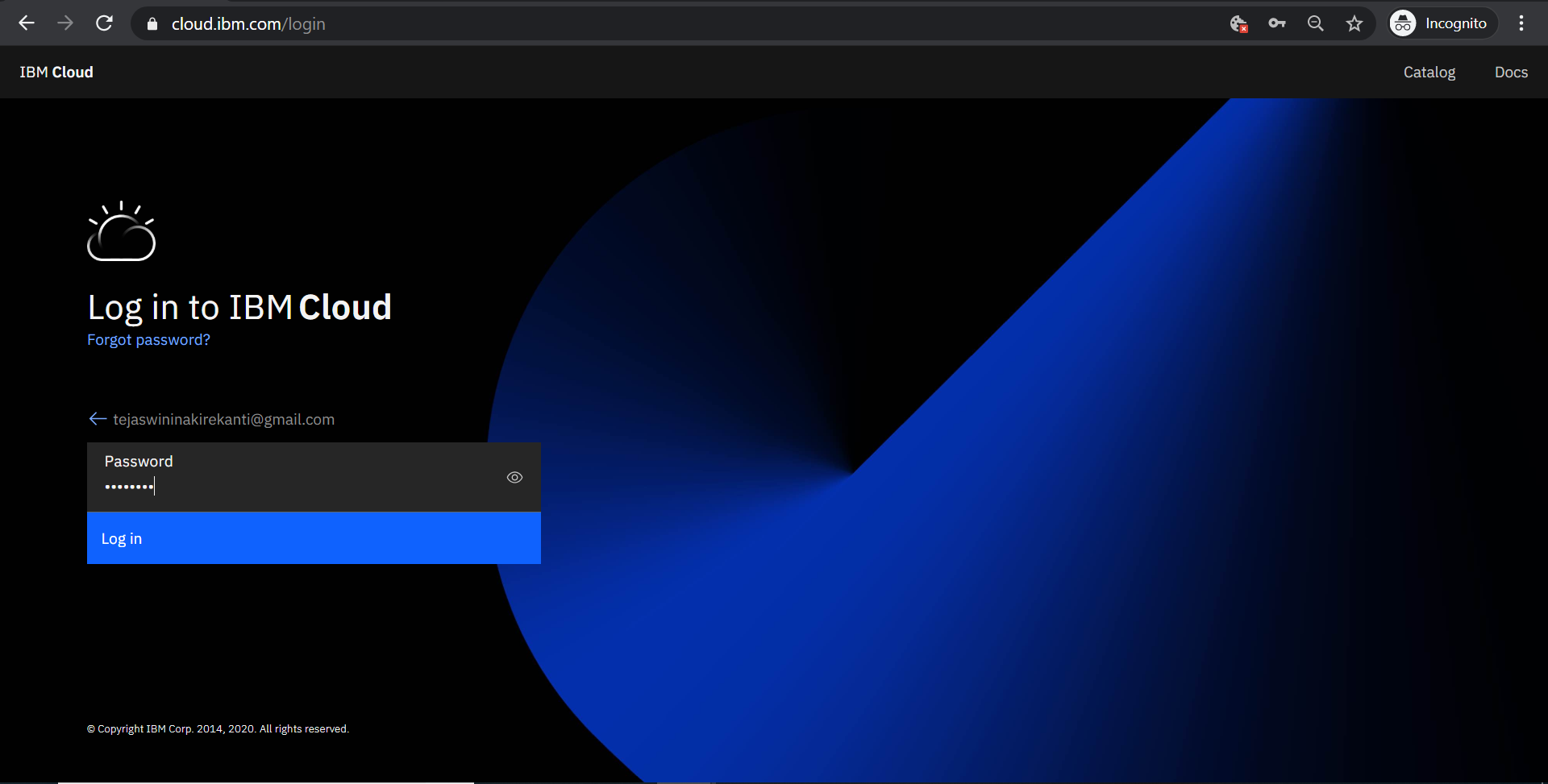
Building a ML model and creating the end points for Node-red integration.

Building Node-red flow to integrate Ml model

**3.2 Hardware / Software designing**

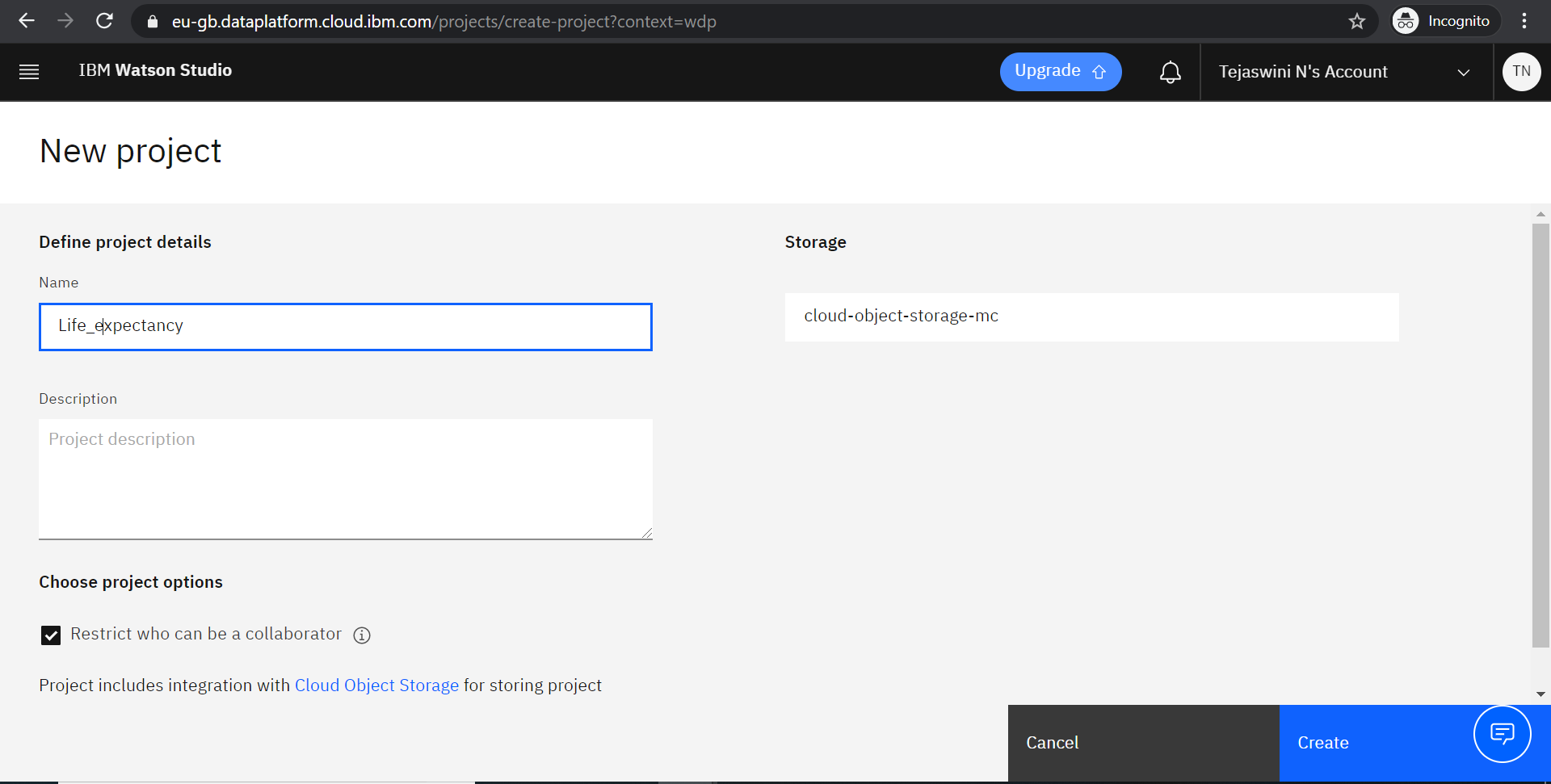
We create and deploy our model using IBM cloud services.

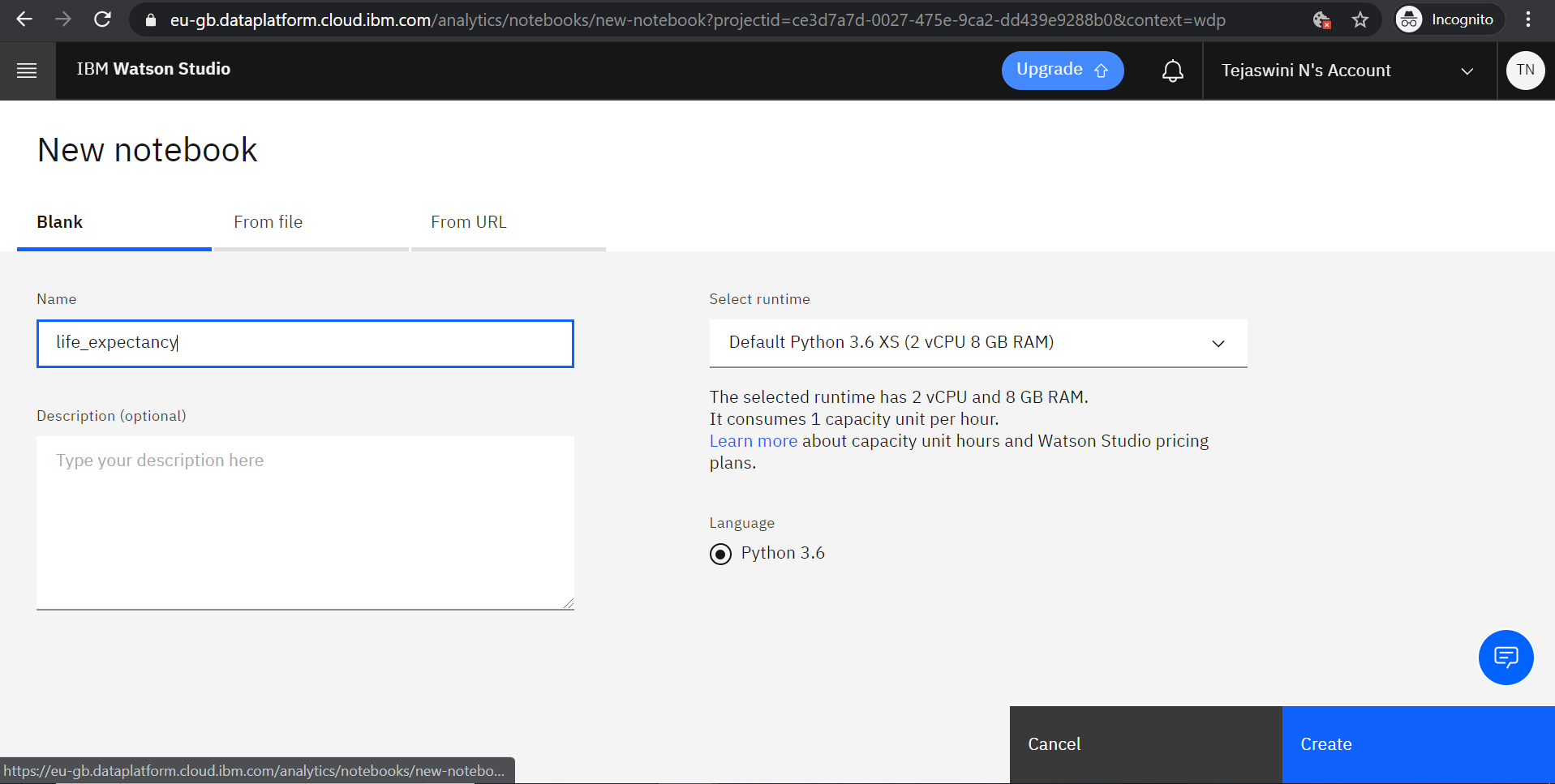
**Step 1:** Create IBM cloud account.



**Step 2:** Create Watson studio service

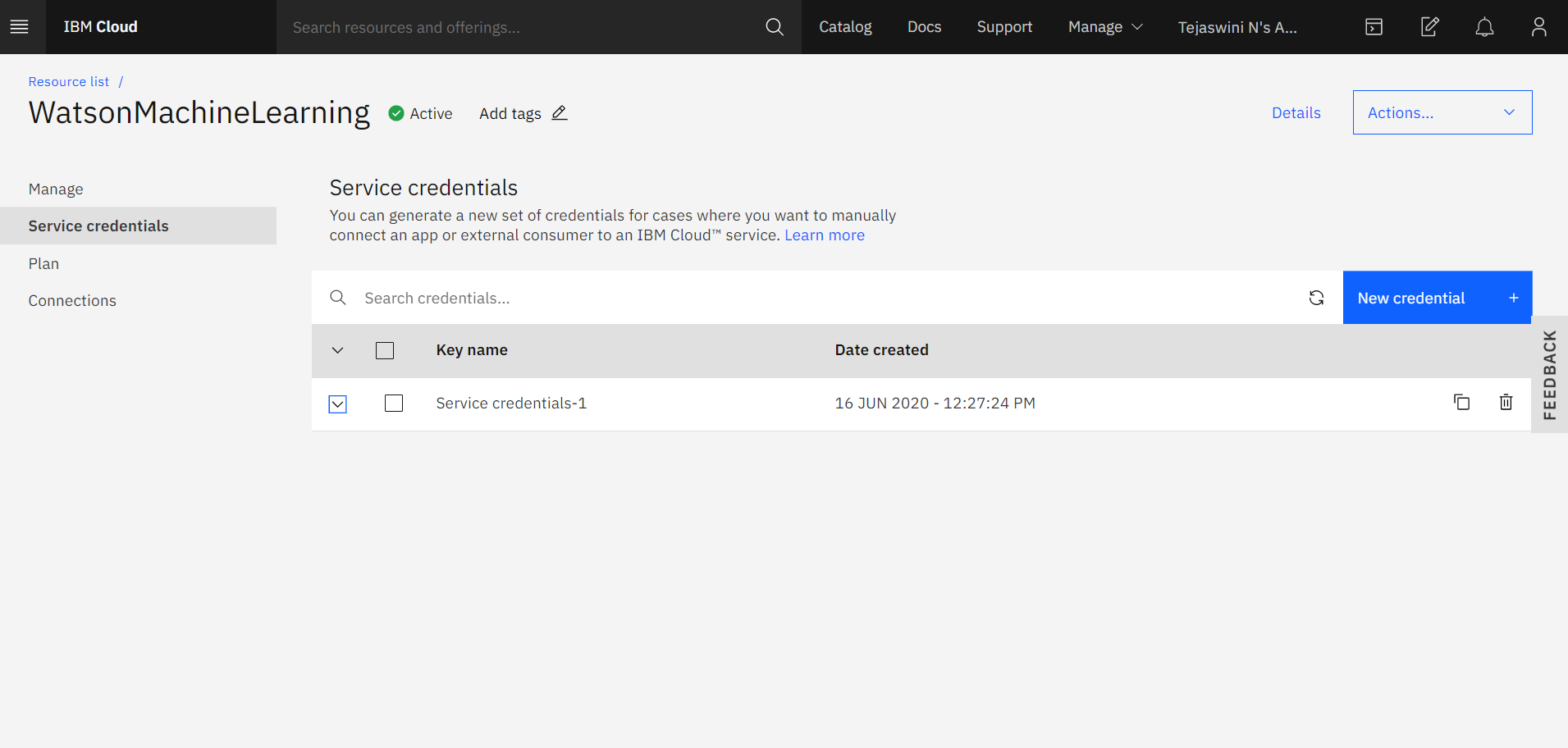
* Go to search bar and search for Watson studio.
* Create it using the plan and click on get started.
* Create new project and go to assets and create a notebook.





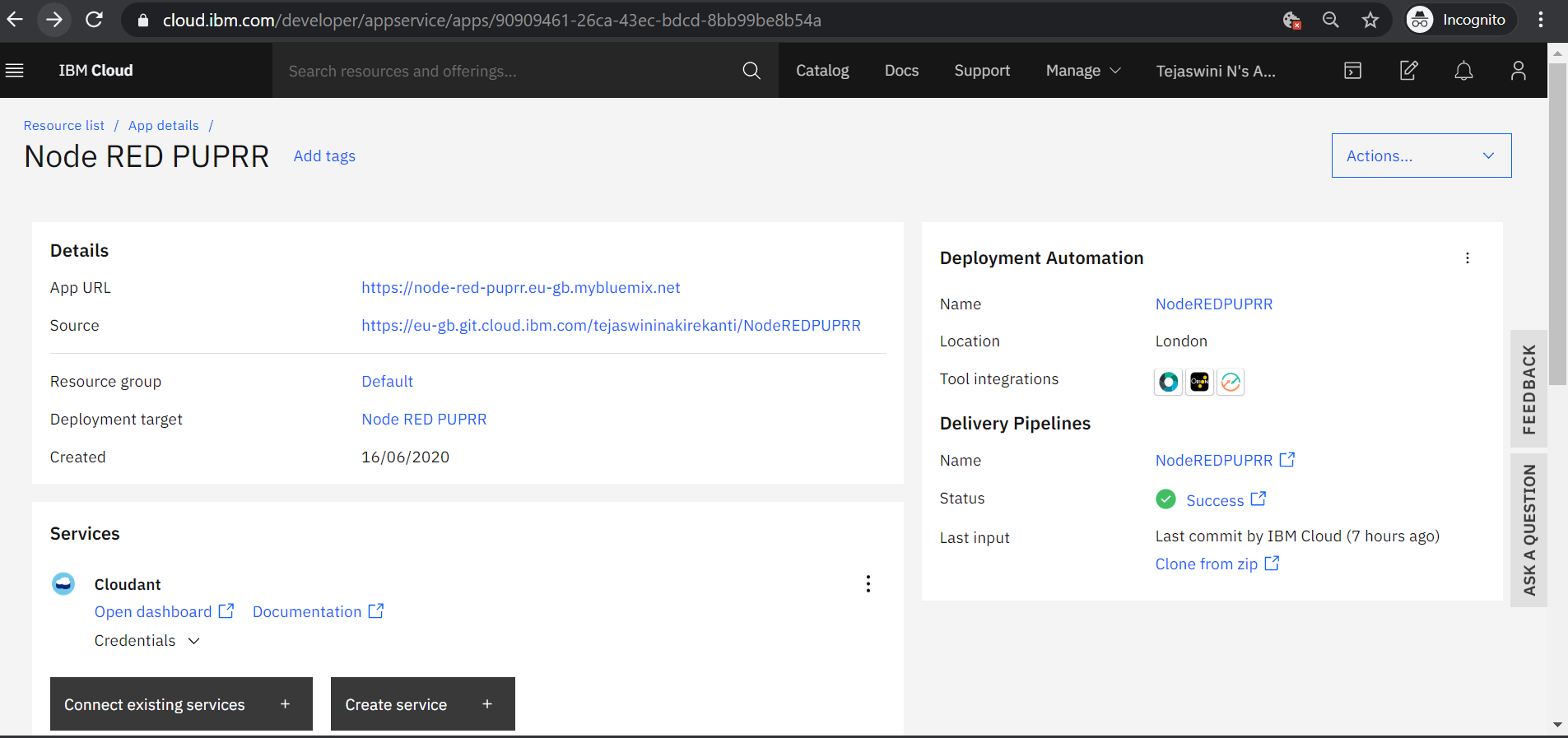
**Step 3:** Create Machine learning service

* Go to catalog and search for Watson Machine learning service and create it.
* Go to service credentials and add new credentials.

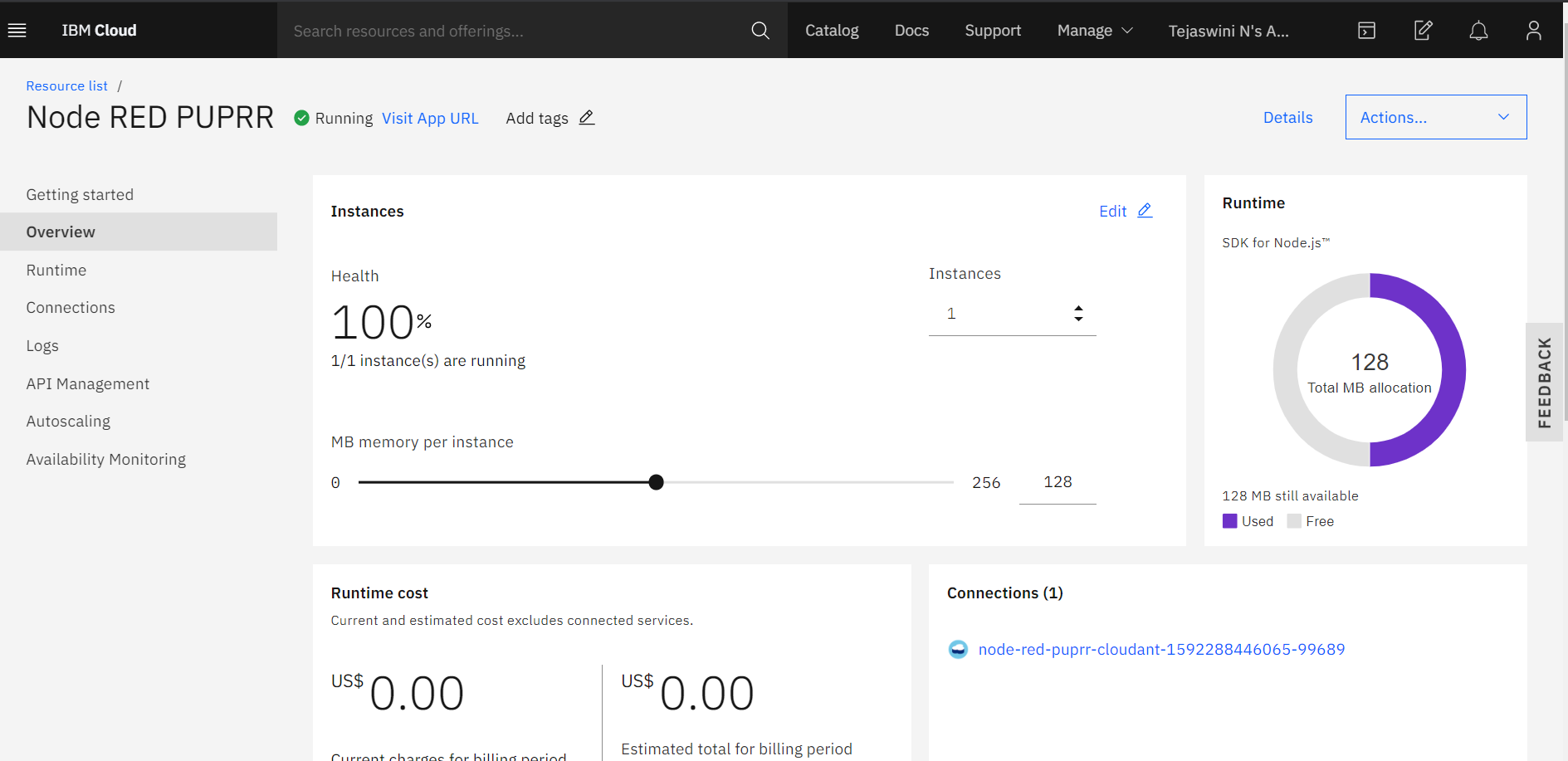


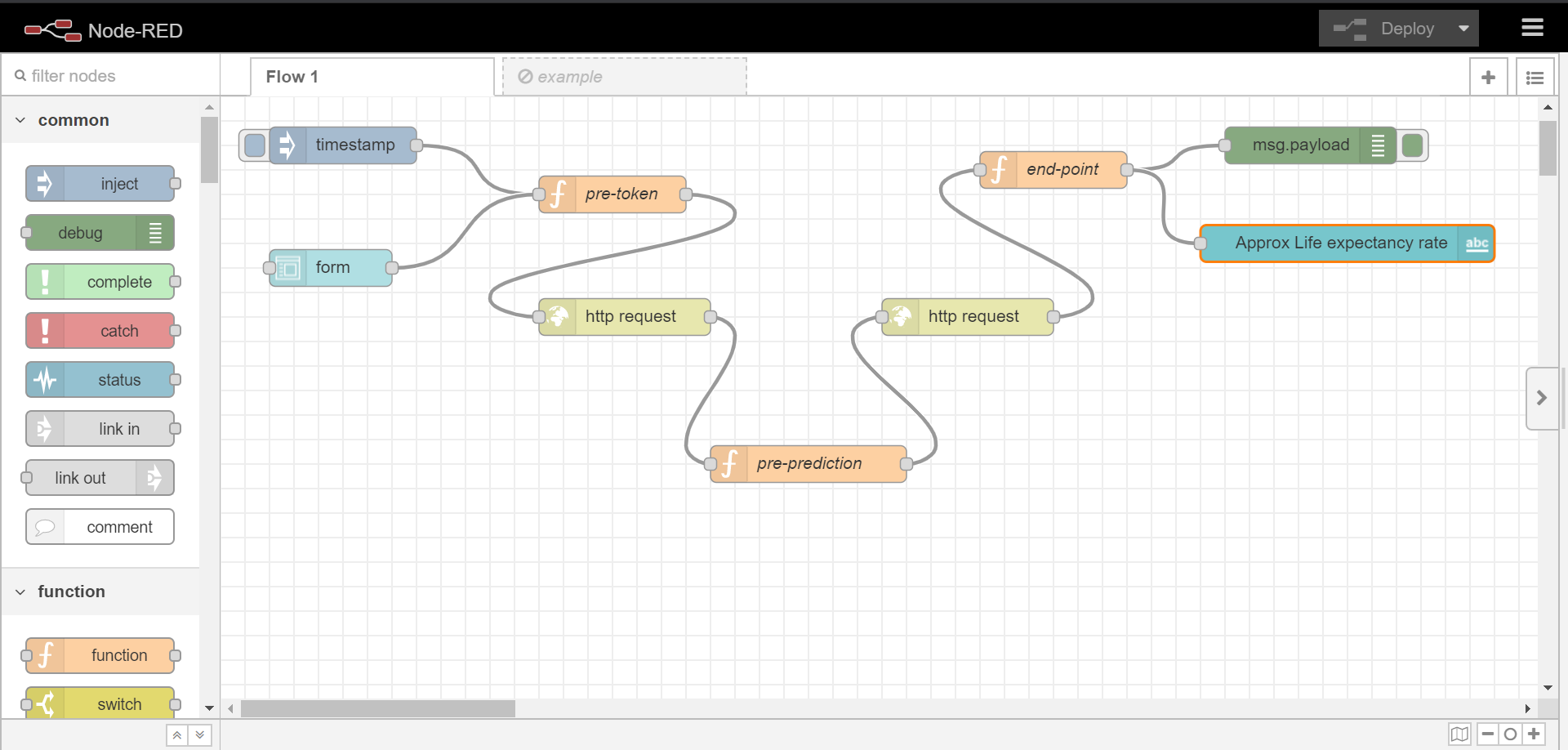
**Step 4:** Create Node-Red application.

* Go to search bar and search for Node-red application.
* Create it by selecting plan and cloud storage.
* Deploy the application and wait till it shows success.



* Go to dashboard and select cloud foundry apps and go to Node-red application.
* Click on Visit app URL and setup the Node-red editor.
* Go to editor and create the flow using nodes.

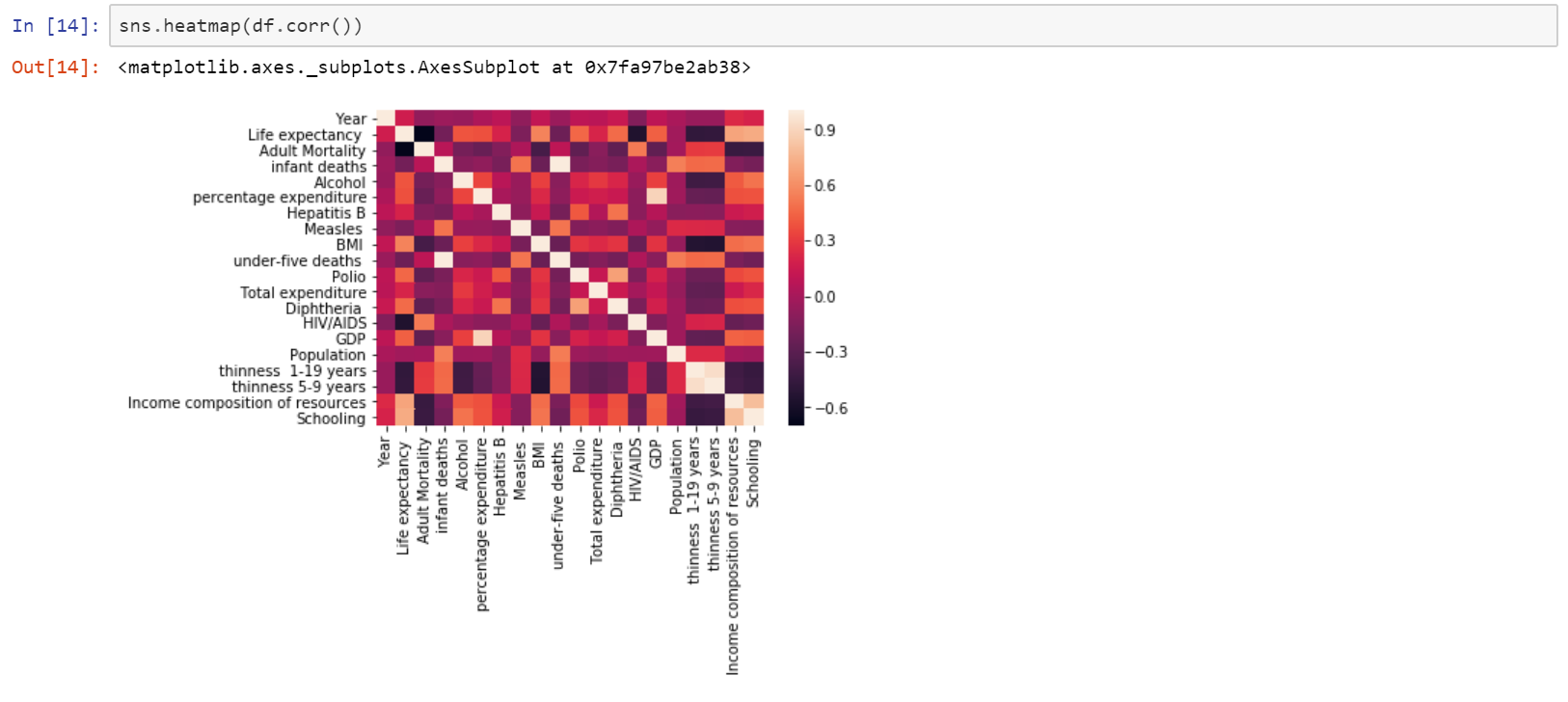




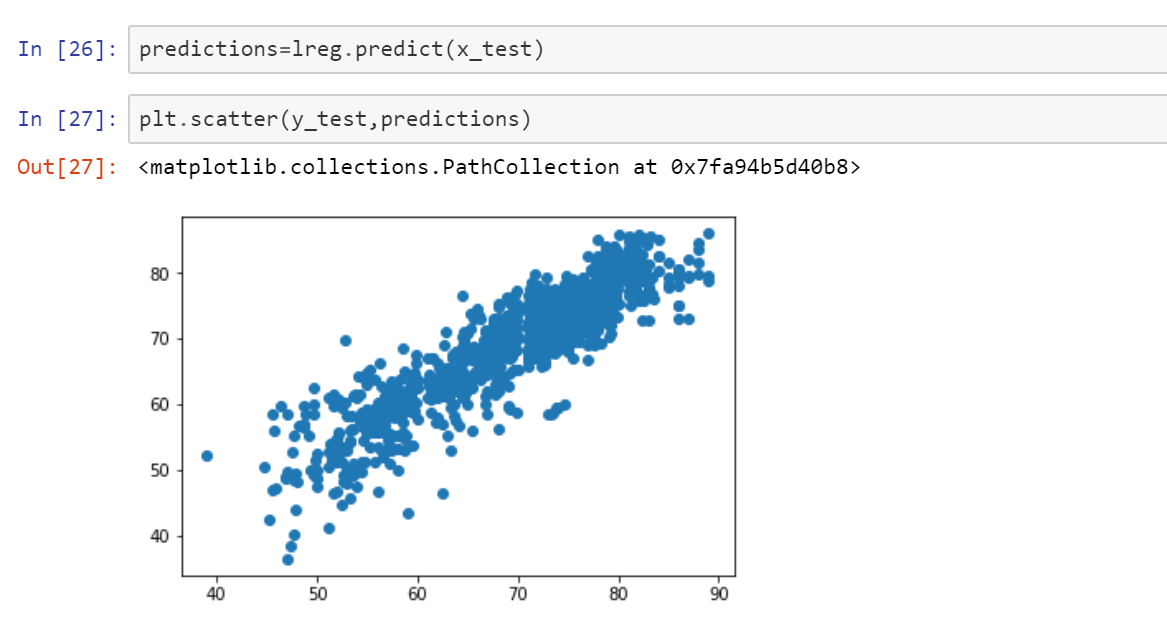
* Link to the Node-red flow <https://node-red-puprr.eu-gb.mybluemix.net/red/#flow/fb47ce2.20f533>

1. **EXPERIMENTAL INVESTIGATIONS**

* Correlation heat map drawn from the dataset.



* Predictions from the model.



1. **Flow chart**

Http request to the IBM cloud for accessing Machine learning services.

Enter the fields in the UI form

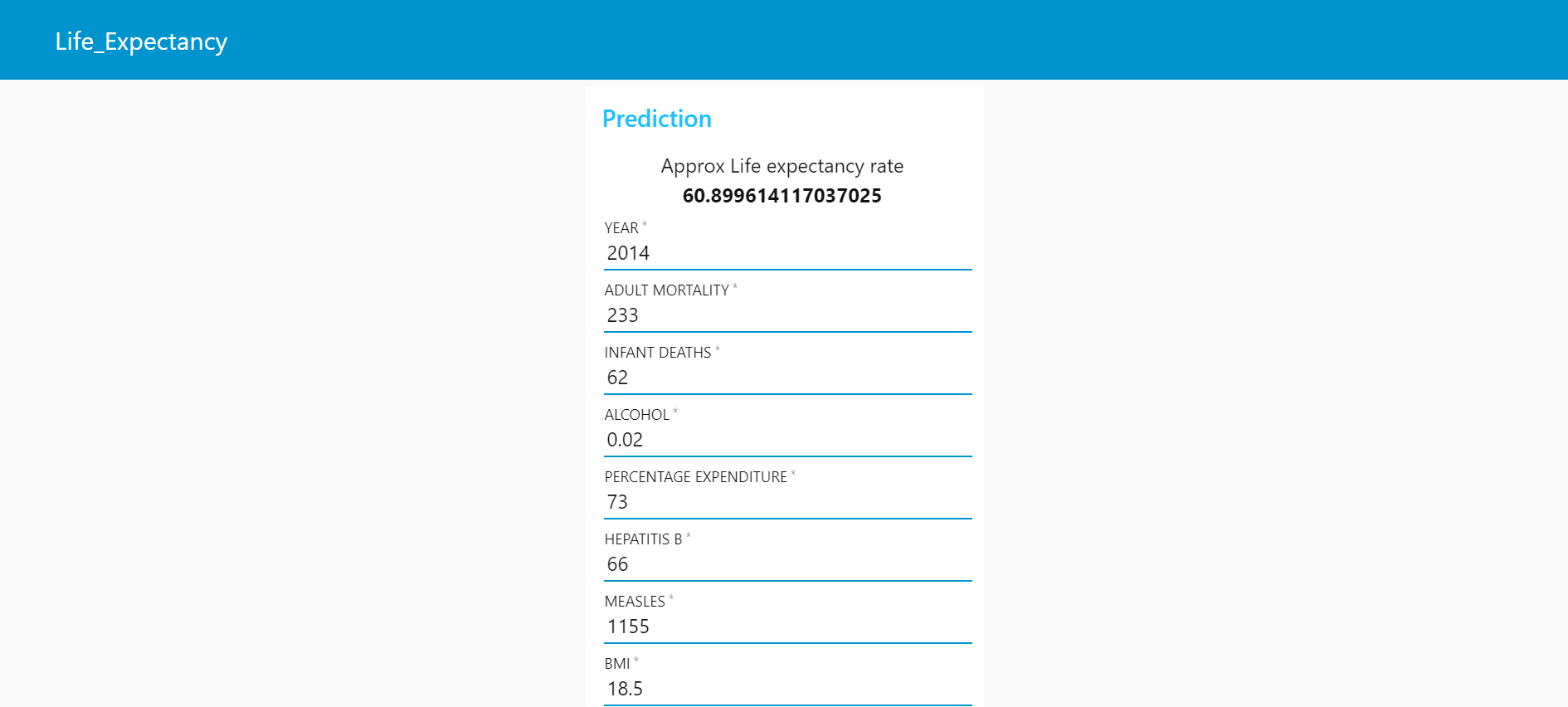
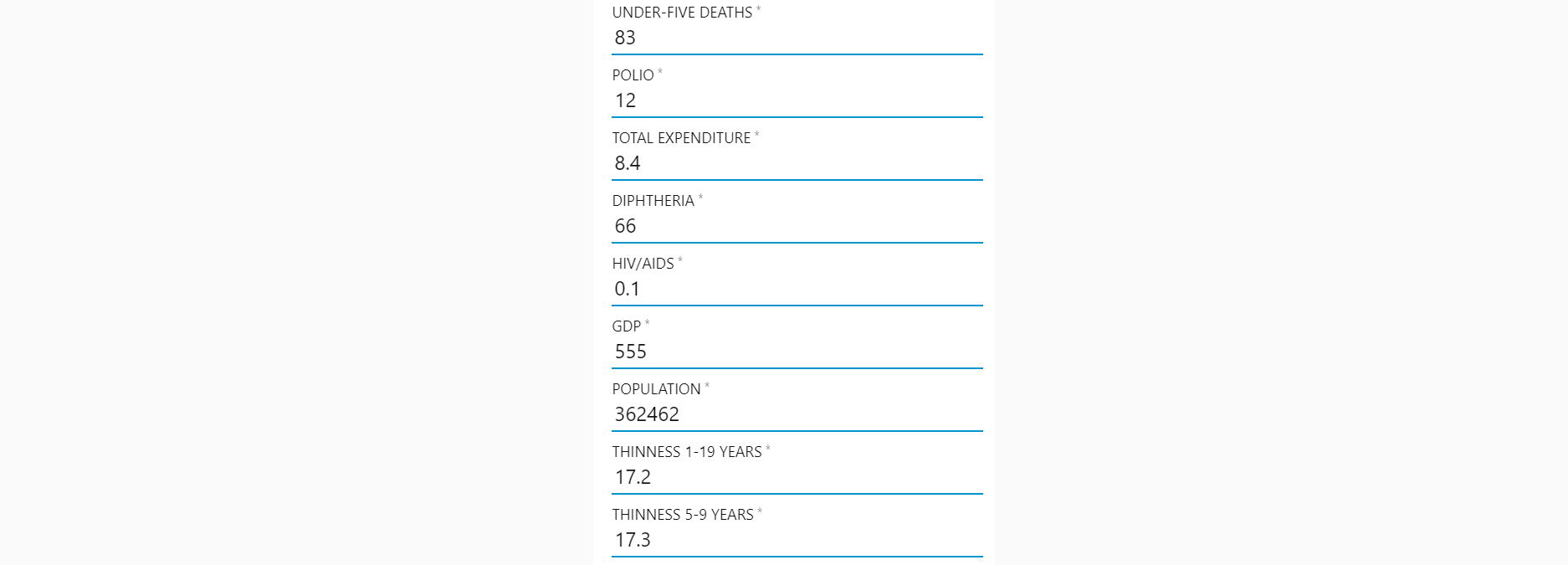
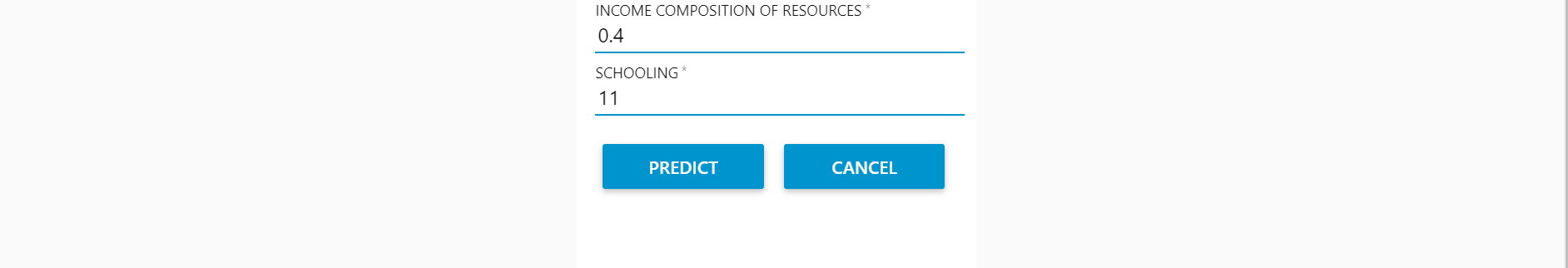
Http request to Machine learning model.

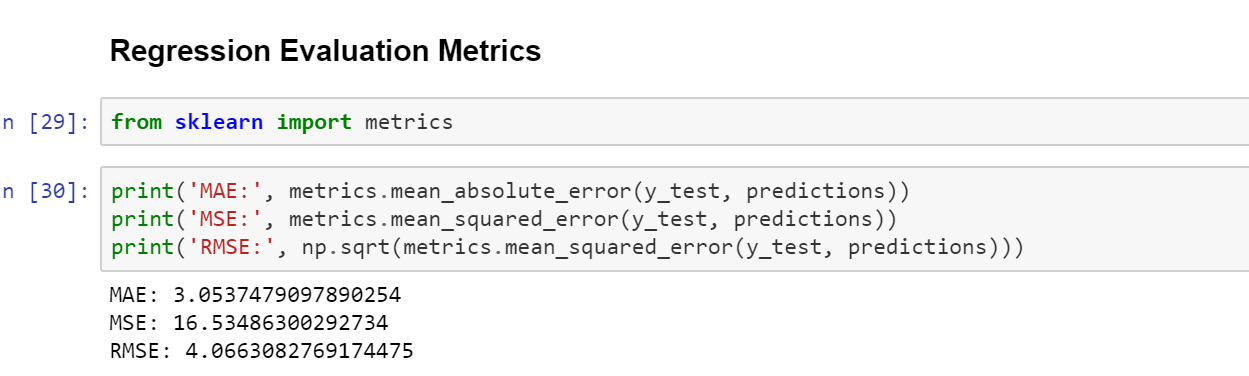
Predicted values by the model

User Interface

1. **Result**

* The user interface to predict the life expectancy.

* These metrics tells us how better our model is. 

1. **ADVANTAGES & DISADVANTAGES**

* Machine learning can review large amount of data and discover specific trends and patterns which helps in increasing the accuracy of the prediction.
* If the data is too large then it takes some time to train the model in machine learning.
* More the data more is the accuracy of the prediction.
* Using IMB cloud services we can generate and deploy our project so easily.
* We can easily create a user interface using Node-Red flow editor.

1. **APPLICATIONS**

* This application is a perfect use case for regression, which determines the relationship between one dependent variable (life expectancy) and a number of independent variables (development indicators).
* Lifeexpectancy is one of the most important factors in end-of-life decision making. Good prognostication for example helps to determine the course of treatment and helps to anticipate the procurement of health care services and facilities, or more broadly: facilitates Advance Care Planning.

1. **CONCLUSION**

* The features which plays a crucial role in calculating the expectancy rate accurately in our model are Adult mortality, GDP, Education, BMI, schooling, Income composition of resources. Therefore people, government and countries can evaluate the expectancy rate and can draw the insights so that they can improve those factors which are influencing to increase the rate.

1. **FUTURE SCOPE**

* We can predict more accurate rate if the dataset increases.
* Also we can suggest some parameters to increase the life expectancy rate in their country.
* Can suggest the changes in the life style if required.
* This application also helps medicine manufacture companies so that they can increase or decrease the production of medicine for particular disease.

1. **BIBILOGRAPHY**

* <https://developer.ibm.com/articles/introduction-to-machine-learning/>
* <https://developer.ibm.com/tutorials/build-and-test-your-first-machine-learning-model-using-python-and-scikit-learn/>
* <https://developer.ibm.com/tutorials/learn-regression-algorithms-using-python-and-scikit-learn/>
* <https://www.kaggle.com/kumarajarshi/life-expectancy-who>

**APPENDIX**

A. Source code

