1. INTRODUCTION

a. Overview

This project is to build a model while considering historical data from a period of 2000 to 2015 for all the countries. The model trained in this project will be able to predict the average lifetime of a human being given some input factors .With the help of this project any country is able to predict the expected lifetime of their countrymen and then accordingly take preventive measures to improve on their healthcare measures. This will also help countries in improving a particular field such as GDP ,alcohol intake,etc which have a high impact on a country's life expectancy.

b. Purpose

The average life Expectancy of a certain country says many things about that particular country. It ultimately helps in predicting the health conditions and the development of the health sector in that particular country. This ultimately helps the nation to find the area which needs attention in an urge to improve its contribution in average lifespan of a human being. The expectancy obviously depends upon the country's population, GDP, the economy of the country and many more factors. It is not enough to have a long life , Instead with having a long life one should have a fit life as well.

2. LITERATURE SURVEY

2.1 Existing problem

Past studies have revealed a lot of work in the field of predicting life expectancy of a human being. After reviewing existing works and techniques in the prediction of human Life Expectancy, and finally reached a conclusion that it is possible to predict a Average Life Expectancy for individuals using advancing technologies and devices such as big data, AI, machine learning techniques, and PHDs, wearables and mobile health monitoring devices, IOT. It is noticed that the collection of data is a huge challenge due to the privacy and government policy considerations, which will require collaboration of various bodies in the health industry. The interworking of a heterogeneous health network is also a challenge for data collection. Despite these challenges, a possibility of predicting Life by proposing an approach of data

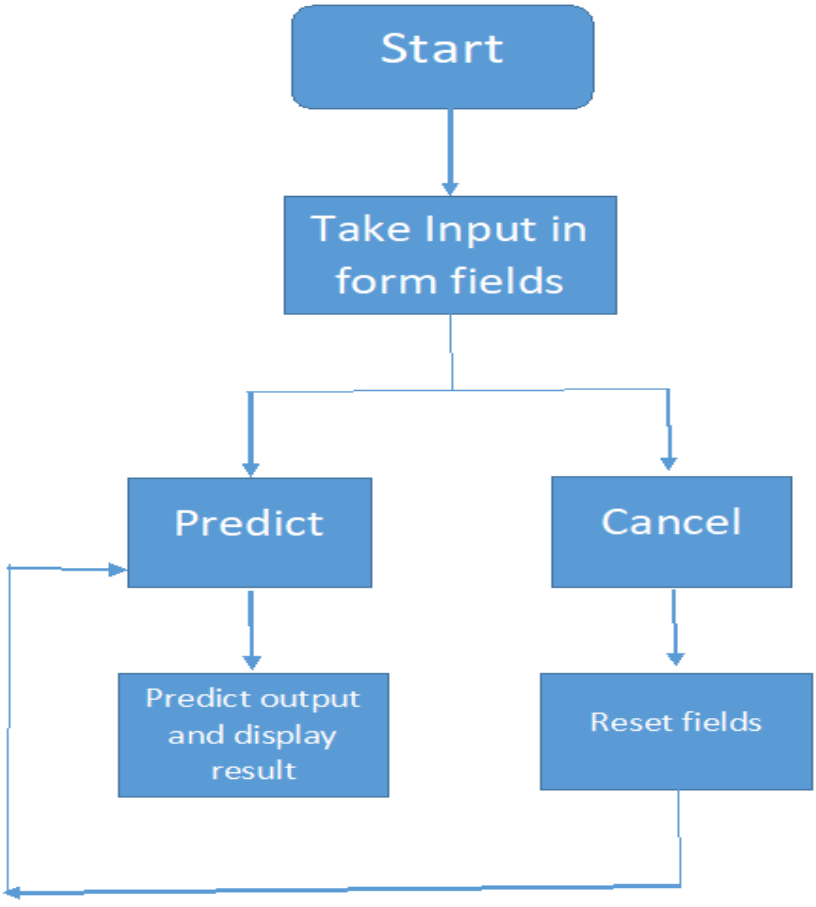
collection and application by smartphone,in which users can enter their information to access the cloud server to obtain their own predicted Lifespan based on the given inputs. To verify the accuracy of PLE prediction and validation of data quality, big data techniques and analysis algorithms need to be developed and tested in a real-life situation with several sample groups. As artificial intelligence technology is evolving and being applied rapidly, feasibility may be increasing to collect health data from the public as well as existing health agencies such as centralized health servers.

2.2 Proposed solution

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. Also, some of the past research was done considering multiple linear regression based on a data set of one year for all the countries. Hence, this gives motivation to resolve both the factors stated previously by formulating a regression model based on mixed effects model and multiple linear regression while considering data from a period of 2000 to 2015 for all the countries. Important immunization like Hepatitis B, Polio and Diphtheria will also be considered. In a nutshell, this study will focus on immunization factors, mortality factors, economic factors, social factors and other health related factors as well. Since the observations in 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. The model of" Predicting Life Expectancy using Machine Learning" uses IBM Cloud services, which helps to avoid any storage issues. The UI Presented to the users is a website url i.e. on users fingertips.

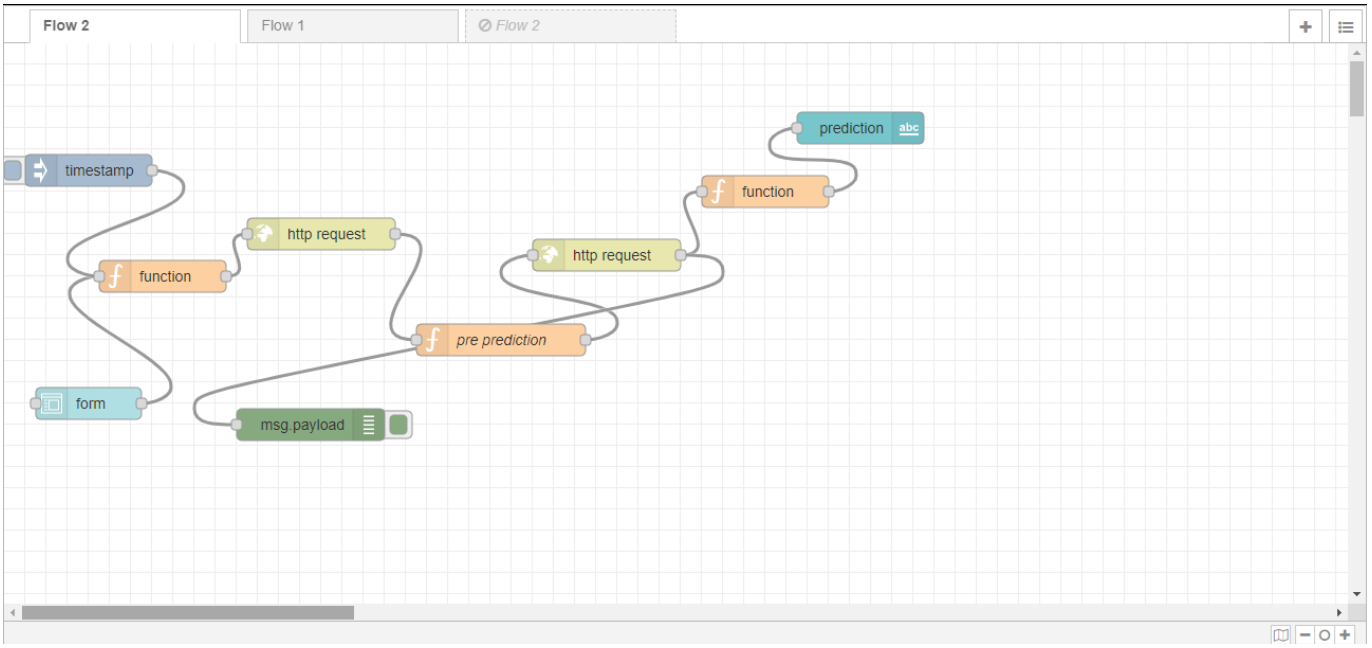
3. THEORITICAL ANALYSIS

3.1 Block diagram



3.2 Hardware / Software designing

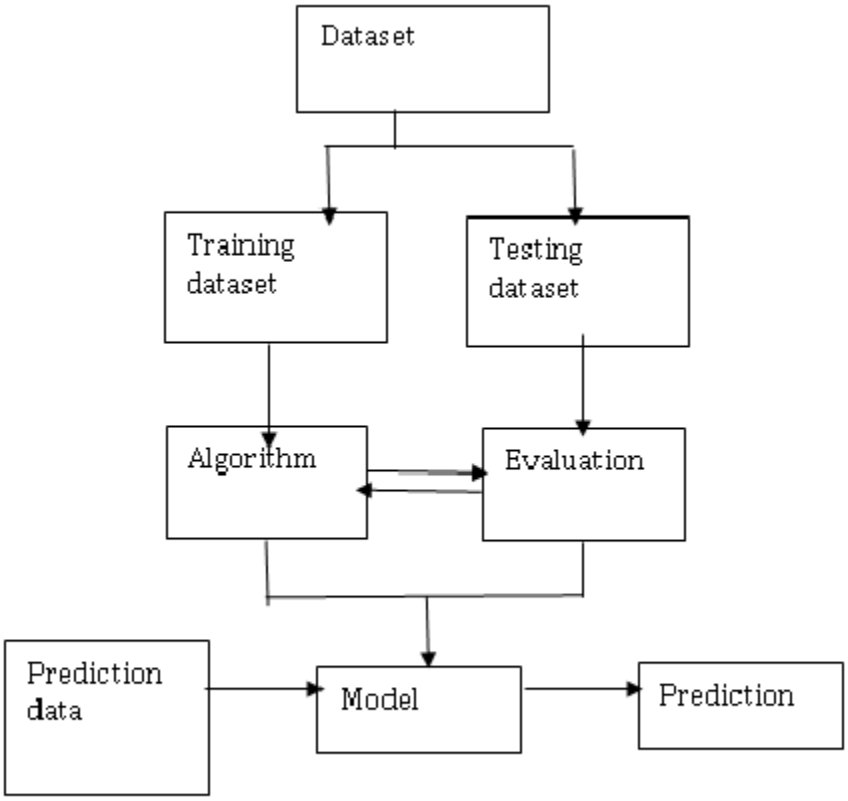
Gui of project with Node Red application.



4. EXPERIMENTAL INVESTIGATION

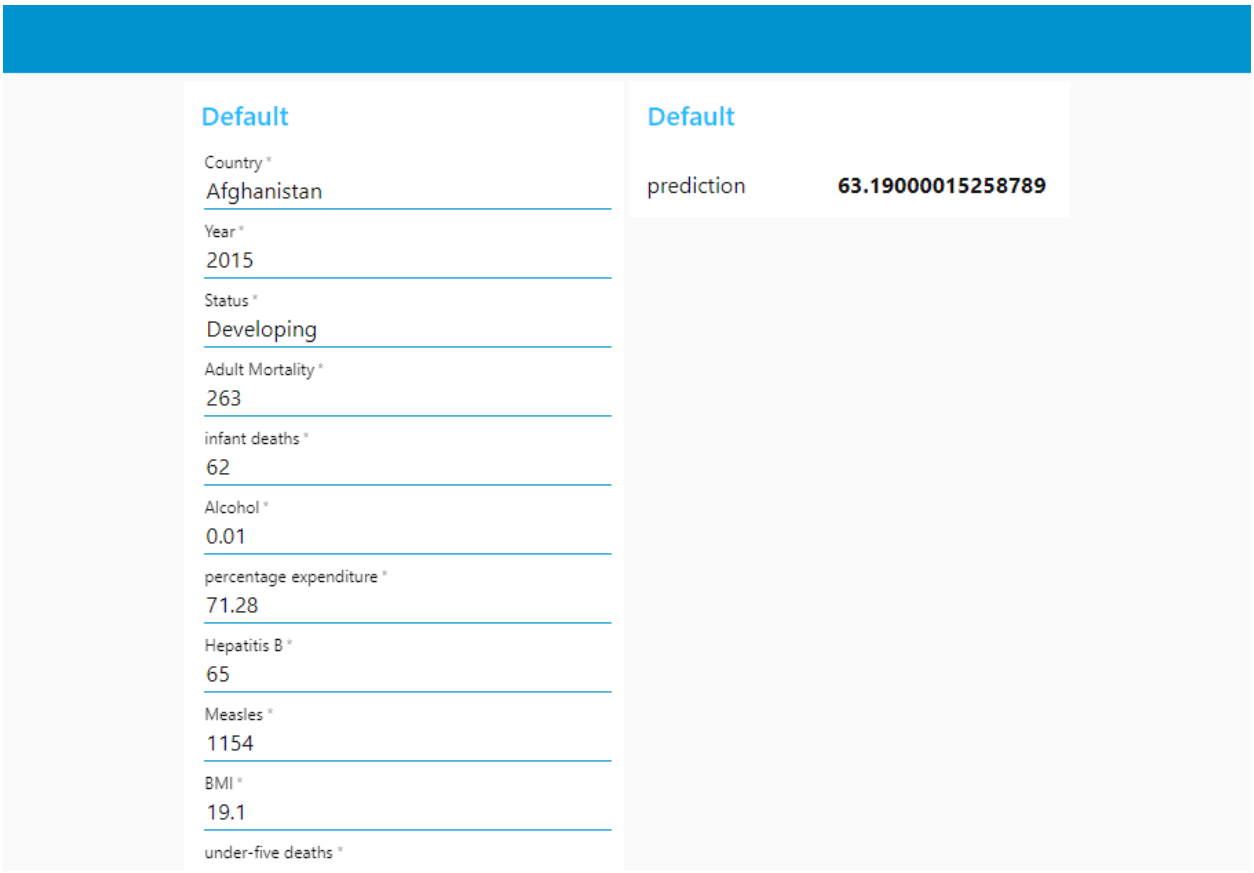
The problem statement falls into regression category where life expectancy is depends on medical status, adult mortality, Income composition of resources, consumption of alcohol, etc. After testing results from linear regression, SVR & random forest regressor accuracy got with random forest regressor was quite satisfying.

5. FLOWCHART



5. RESULT:

The user friendly Graphical User interface is shown in Figure 2. This GUI is connected to the trained machine learning model present in the backend(IBM watson notebook). The user has to fill in the inputs accordingly and click on the “Predict” button present at the end of the form. On clicking the “Predict” button, the user will be displayed the predicted life expectancy at the predict label, based on the inputs



6. ADVANTAGES & DISADVANTAGES

Advantages:

1. Life expectancy predictions have the potential to be beneficial to individuals, health service providers and governments.
2. For instance, they would make people more aware of their general health, and its improvement or deterioration over time. This may motivate them to make healthier lifestyle choices.
3. They could also be used by insurance companies to provide individualized services, such as how some car insurance companies use black-box technology to reduce premiums for more cautious drivers.
4. Governments may be able to use predictions to more efficiently allocate limited

resources, such as social welfare assistance and health care funding, to individuals and areas of greater need.

Disadvantages:

1. People may become distressed if their life expectancy is unexpectedly low, or at the thought of having one at all.
2. This raises concerns about how such predictions could impact those who experience or are at risk of mental health problems.
3. If so, issues related to data compliance, as well and collaboration with government and state agencies will need to be carefully managed.
4. Any system predicting life expectancy would handle highly sensitive data, raising ethical and privacy concerns.

7. APPLICATION:

1. Insurance
2. Tax
3 Pharmaceutical Industry

8. CONCLUSION

Prognostication of life expectancy is difficult for humans. Research shows that machine learning technique offers a feasible and promising

approach to predicting life expectancy. The research has potential for real-life applications, such as supporting timely recognition of the right moment to start Advance Care Planning.

9. FUTURE SCOPE

As future scope, we can connect the model to the database which can predict the life Expectancy of not only human beings but also of the plants and different animals present on the earth. This will help us analyze the trends in the life span. A model with country wise bifurcation can be made, which will help to segregate the data demographically.

10. BIBILOGRAPHY

1. Dataset: https://www.kaggle.com/kumarajarshi/life-expectancy-who
2. https://link.springer.com/chapter/10.1007/978-3-030-05075-7_15
3. https://ourworldindata.org/life-expectancy#:~:text=Life%20expectancy%20is%20the%20key,of%20death%20in%20a%20population.

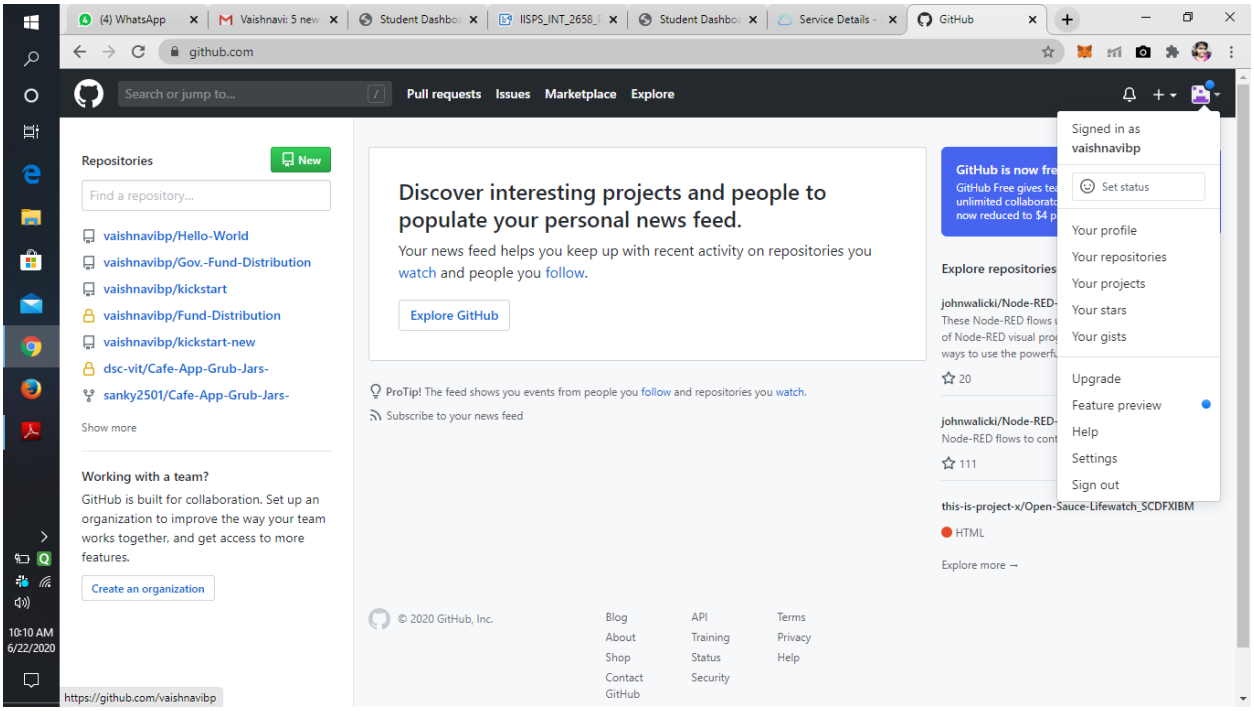
APPENDIX

A. Source code

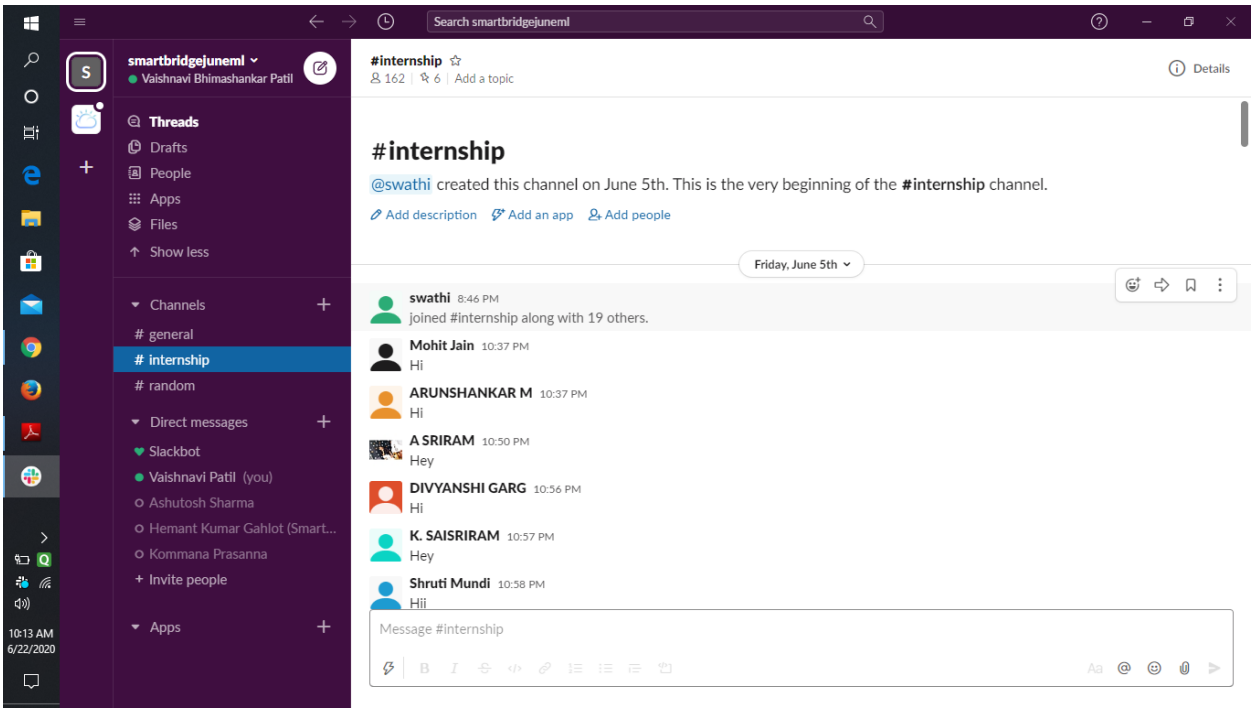
https://github.com/vaishnavibp/Life-Expectancy

Task 1:Project planning and kickoff

subtask: Setup the environment
Github page

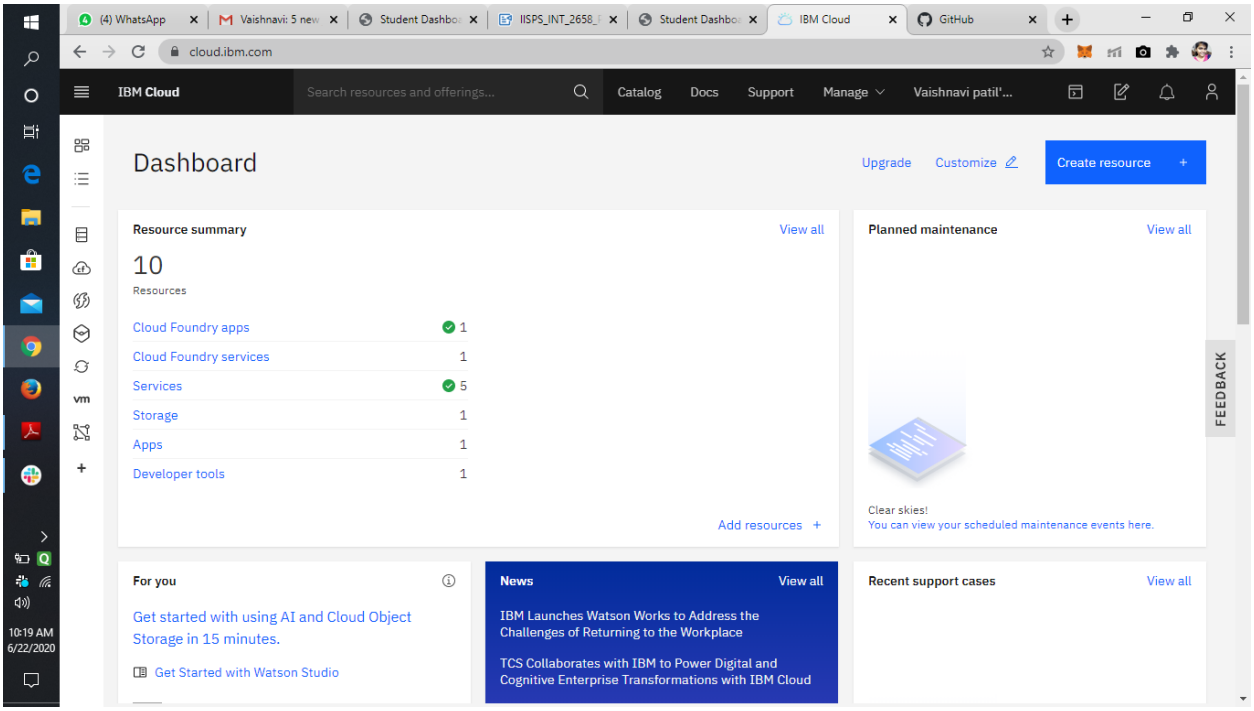


slack:

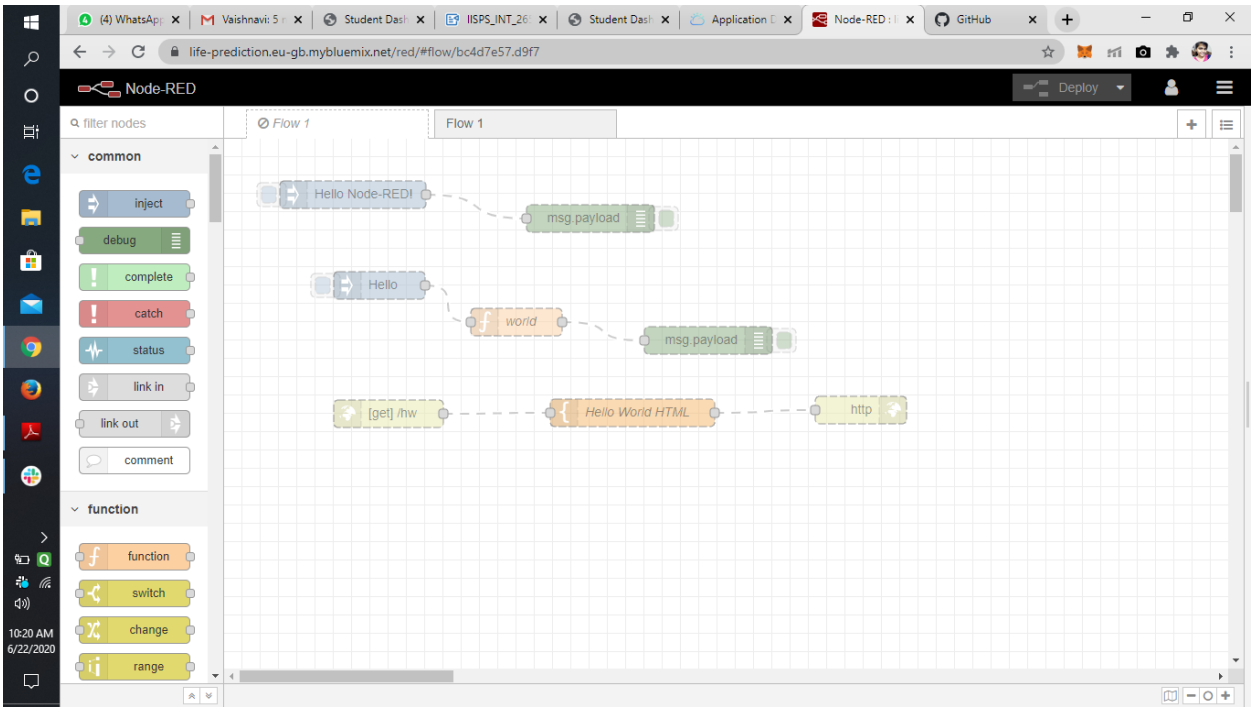


Task 2:Explore IBM cloud platform

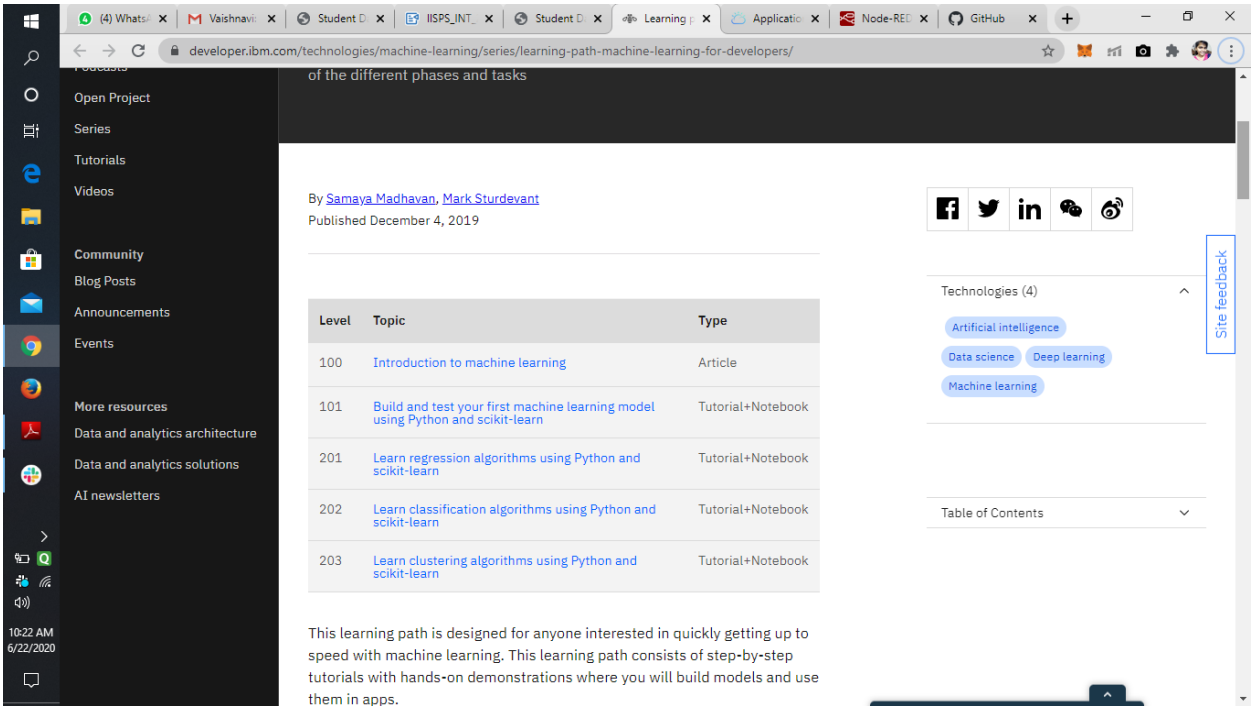
IBM cloud account:

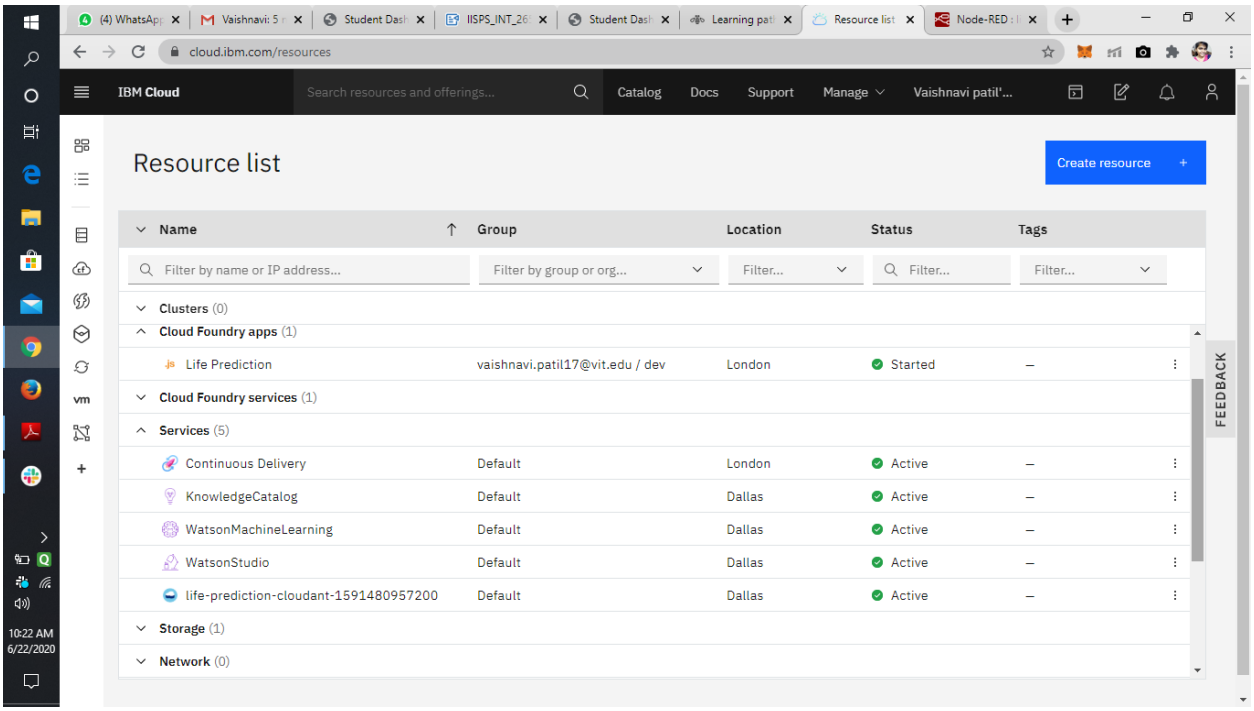


Node-red starter application:

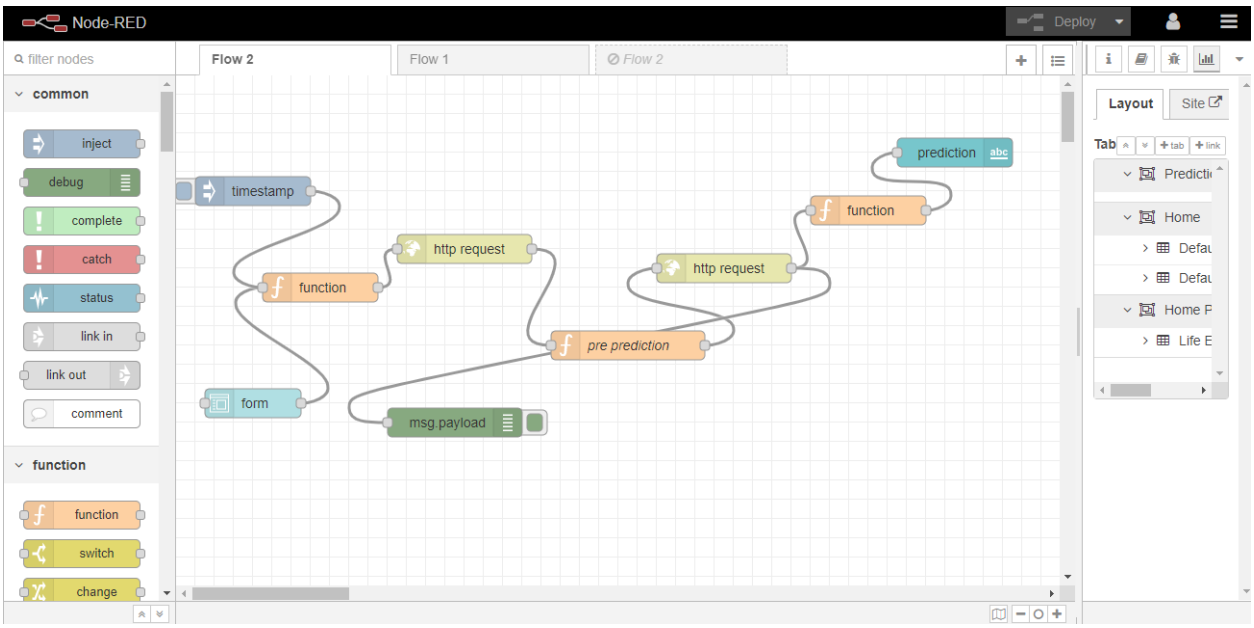


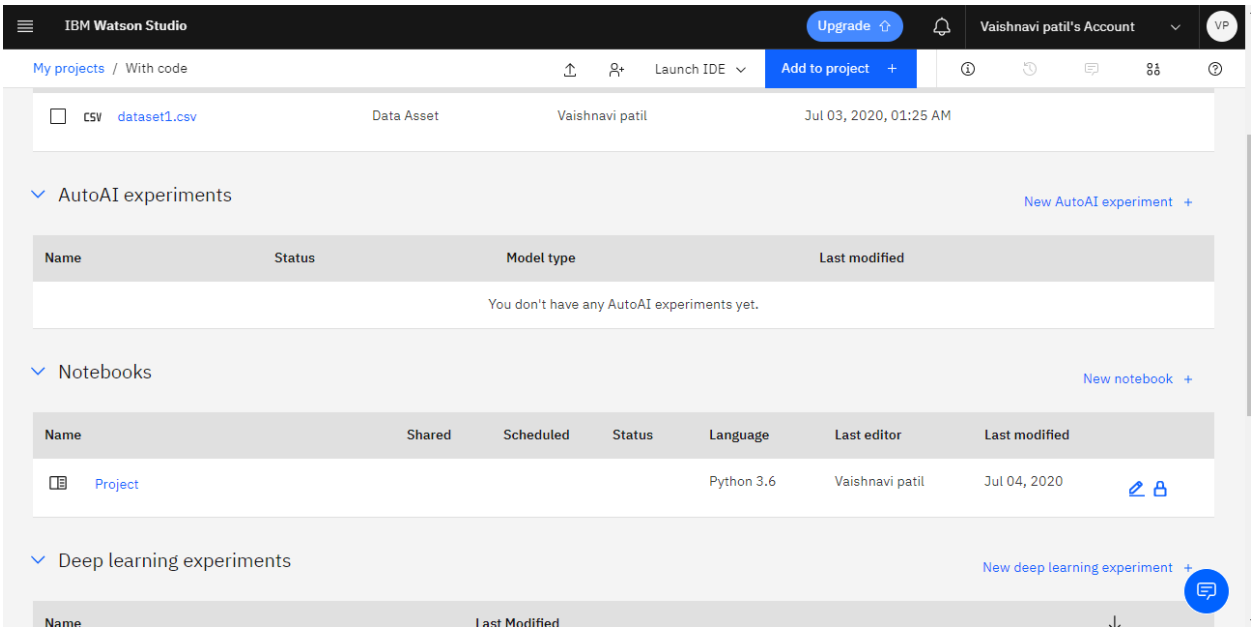
Task 3: IBM watson machine Learning services:

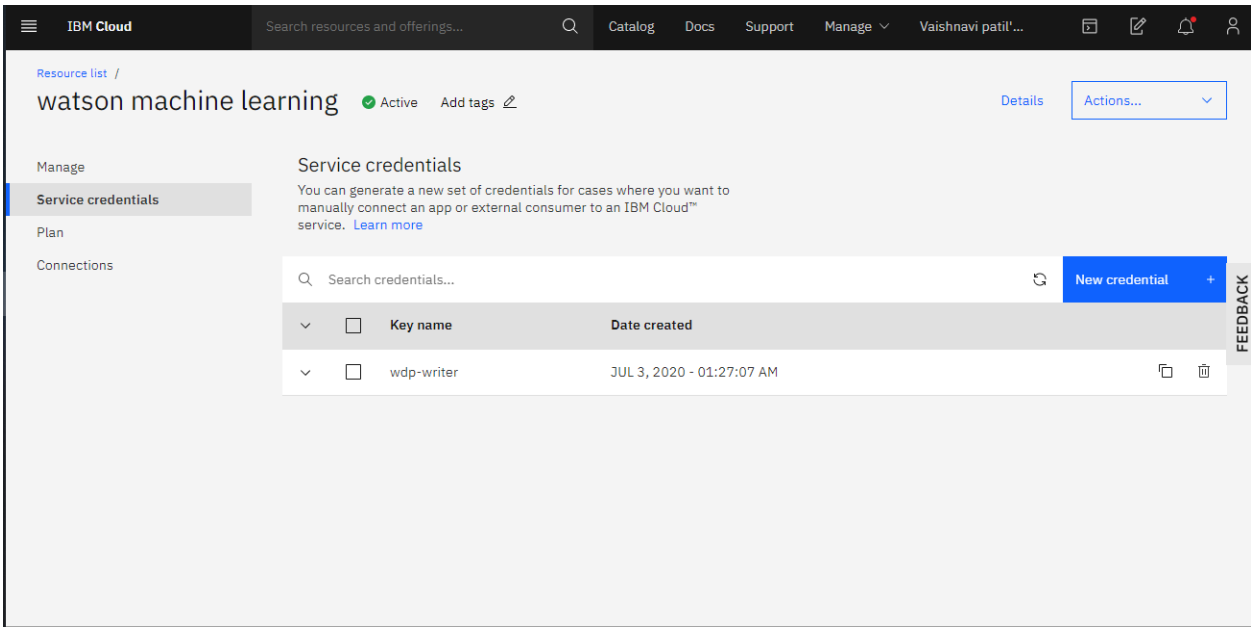


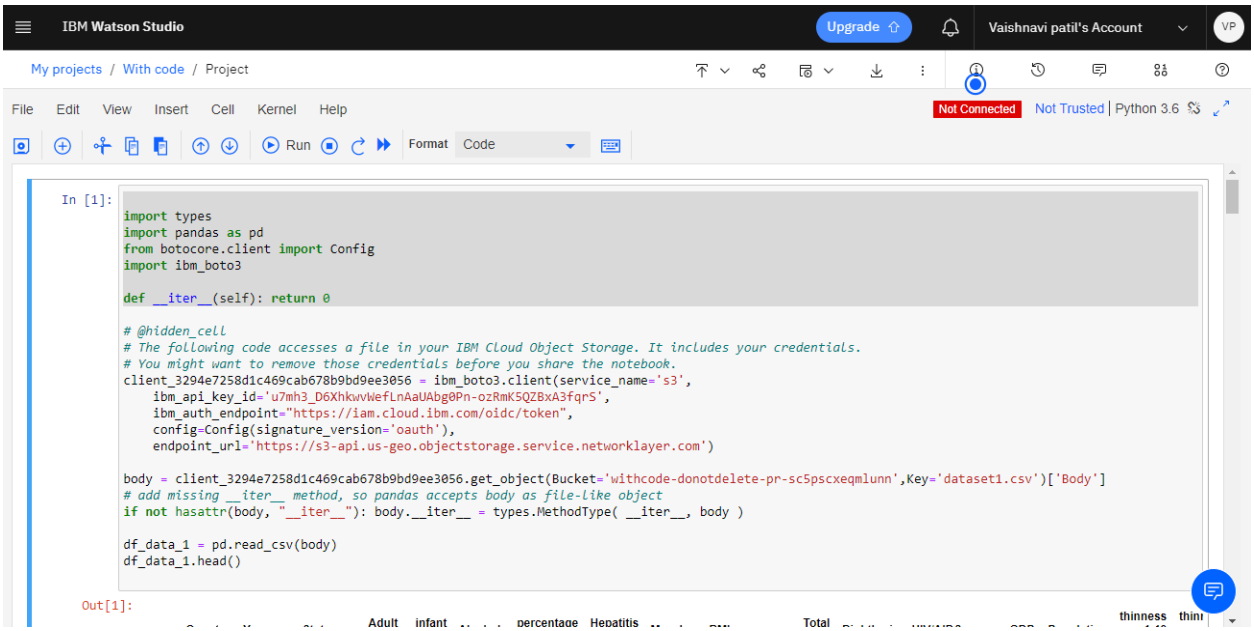


Task 4: Introduction to watson studio:




IBM Watson Studio
Upgrade
Vaishnavi patil's Account
VP
My projects / With code
Launch IDE
Add to project +
CSV dataset1.csv
Data Asset
Vaishnavi patil
Jul 03, 2020, 01:25 AM
AutoAI experiments
New AutoAI experiment +
Name
Status
Model type
Last modified
You don't have any AutoAI experiments yet.
Notebooks
New notebook +
Name
Shared
Scheduled
Status
Language
Last editor
Last modified
Project
Python 3.6
Vaishnavi patil
Jul 04, 2020
Deep learning experiments
New deep learning experiment +
Name
Last Modified


IBM Cloud
Search resources and offerings...
Catalog
Docs
Support
Manage
Vaishnavi patil'...
Resource list /
watson machine learning
Active
Add tags
Details
Actions...
Manage
Service credentials
Plan
Connections
Service credentials
You can generate a new set of credentials for cases where you want to manually connect an app or external consumer to an IBM Cloud™ service. Learn more
Search credentials...
New credential +
Key name
Date created
wdp-writer
JUL 3, 2020 - 01:27:07 AM
FEEDBACK


IBM Watson Studio
Upgrade
Vaishnavi patil's Account
VP
My projects / With code / Project
File Edit View Insert Cell Kernel Help
Not Connected
Not Trusted | Python 3.6
Run
Format Code
In [1]:
import types
import pandas as pd
from botocore.client import Config
import ibm_boto3
def __iter__(self): return 0
# @hidden_cell
# The following code accesses a file in your IBM Cloud Object Storage. It includes your credentials.
# You might want to remove those credentials before you share the notebook.
client_3294e7258d1c469cab678b9bd9ee3056 = ibm_boto3.client(service_name='s3',
ibm_api_key_id='u7mh3_D6XhkwvWefLnAaUAbg0Pn-ozRmK5QZBxA3fqrS',
ibm_auth_endpoint="https://iam.cloud.ibm.com/oidc/token",
config=Config(signature_version='oauth'),
endpoint_url='https://s3-api.us-geo.objectstorage.service.networklayer.com')
body = client_3294e7258d1c469cab678b9bd9ee3056.get_object(Bucket='withcode-donotdelete-pr-sc5pscxeqmlunn',Key='dataset1.csv')['Body']
# add missing __iter__ method, so pandas accepts body as file-like object
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType(__iter__, body)
df_data_1 = pd.read_csv(body)
df_data_1.head()
Out[1]:
Adult
infant
percentage
Hepatitis
Total
thinness
thini