

PROJECT TITLE:**PREDICTING LIFE EXPECTANCY USING MACHINE LEARNING****Project Scope, Schedule, Team & Deliverables:****Project Summary:**

Life expectancy refers to the number of years a person is expected to live based on the statistical average. This problem statement 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 are given.

Project Requirements:**Functional Requirements:**

Predicting LifeExpectancy rate of a country.

Technical Requirements:

Python

IBM Cloud

IBM Watson

Node-Red

Hardware Requirements:

1.Processor - i3 7th gen or higher

2.Speed - 2GHz or more

Project Deliverable:

Given a form where several inputs are entered by the user (GDP, Status, education, mortality etc);the system predicts the Life Expectancy of the country.

Project Team:

Gidge. Saisri

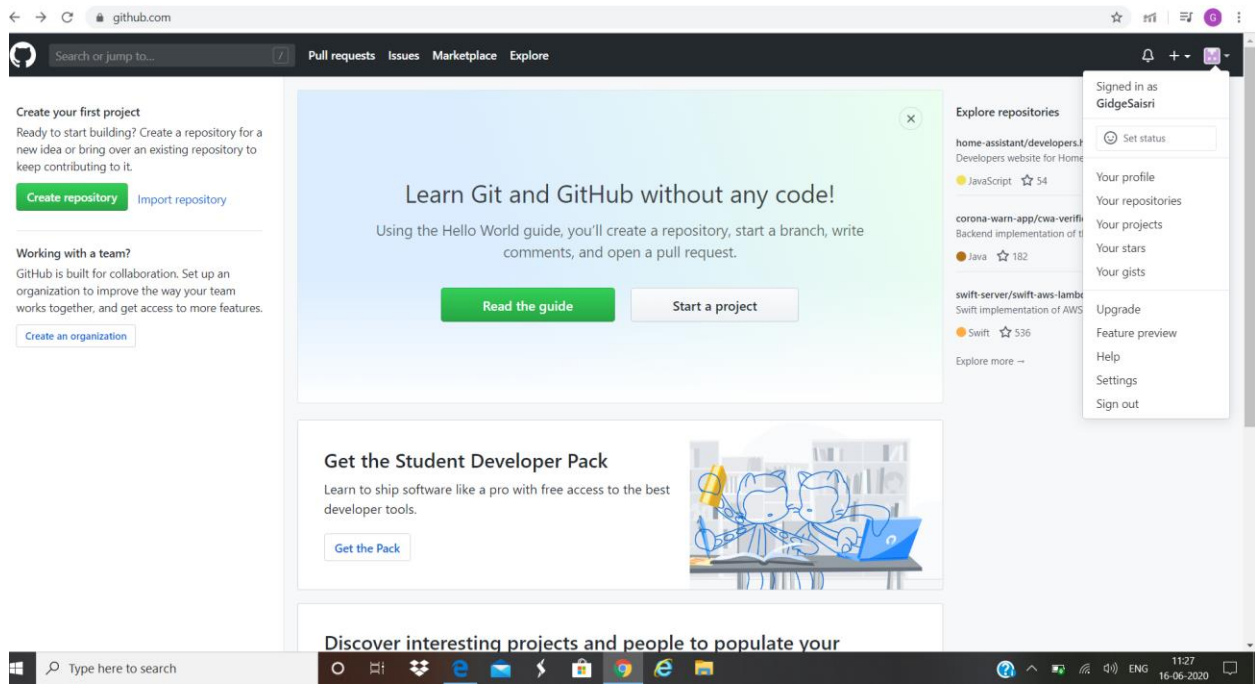
gidgesaisri22@gmail.com

PROJECT SCHEDULE:

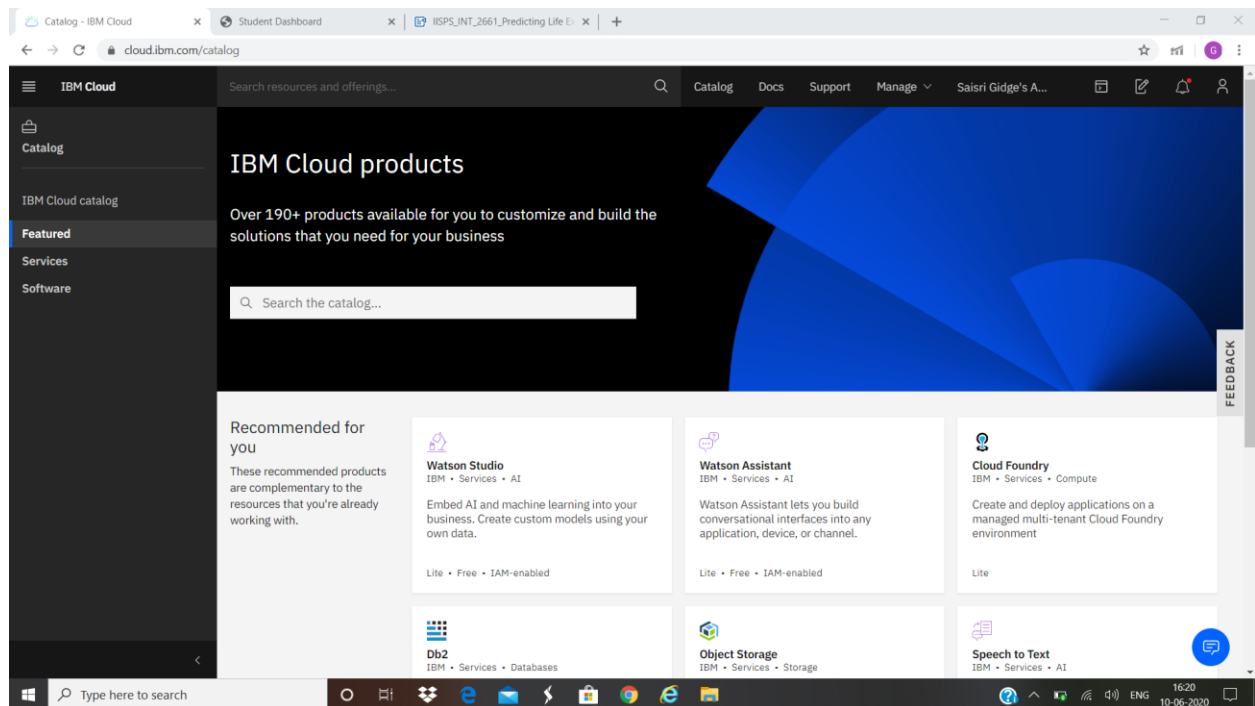
The project is to be completed in 1 month and work for atleast 5 days a week. The project can be divided into two phases, one to build the model and the other for UI.

SET UP THE DEVELOPMENT ENVIRONMENT:

GITHUB Account:



IBM cloud Account :



Node-RED Starter Application :

This screenshot shows the IBM Cloud console interface for a Node RED 100 application. The top navigation bar includes the IBM Cloud logo, a search bar, and links to Catalog, Docs, Support, and Manage. The main content area is titled "Node RED 100" and includes a "hello:sa" button and an "Actions..." dropdown. The "Details" section on the left lists the App URL, Source, Resource group, Deployment target, and Created date. The "Services" section on the left shows the Cloudant service with links to the dashboard and documentation, and a "Credentials" dropdown. The "Deployment Automation" section on the right shows the Name, Location, Tool integrations, and Delivery Pipelines. A "Getting started quickly" link is visible at the bottom of the main content area.

Node RED 100 [hello:sa](#) [Actions...](#)

Details

App URL: <https://node-red-100.eu-gb.mybluemix.net>

Source: <https://eu-gb.git.cloud.ibm.com/gidgesaisri22/NodeRED100>

Resource group: Default

Deployment target: Node RED 100

Created: 16/06/2020

Services

Cloudant

[Open dashboard](#) [Documentation](#)

Credentials

[Connect existing services](#) [Create service](#)

Deployment Automation

Name: NodeRED100

Location: London

Tool integrations:

Delivery Pipelines

Name: NodeRED100

Status: Success

Last input: Last commit by Saisri Gidge (13 minutes ago)

[Clone from zip](#)

[Getting started quickly](#)

This screenshot shows the IBM Cloud console interface for the NodeRED100 | Delivery Pipeline. The top navigation bar includes the IBM Cloud logo, a search bar, and links to Catalog, Docs, Support, and Manage. The main content area is titled "NodeRED100 | Delivery Pipeline" and includes an "Add Stage" button and an "Actions..." dropdown. A notification banner at the top indicates an upcoming change in the 2.7 pipeline base image. The "BUILD" stage on the left shows the last input, jobs, and last execution result. The "DEPLOY" stage on the right shows the last input, jobs, and last execution result.

NodeRED100 | Delivery Pipeline [Add Stage](#) [Actions...](#)

Upcoming change in 2.7 pipeline base image:
Java™ JVM has been upgraded to Java™ 11. [Learn more.](#)

BUILD

STAGE PASSED

LAST INPUT: Git URL [View logs and history](#)

Last commit by Saisri Gidge 8m ago

[Update package.json](#)

JOBS

Build Passed 5m ago

[View logs and history](#)

LAST EXECUTION RESULT

Build 3

DEPLOY

STAGE PASSED

LAST INPUT: Stage: BUILD / Job: Build

Build 3

JOBS

Rolling Deploy Passed now

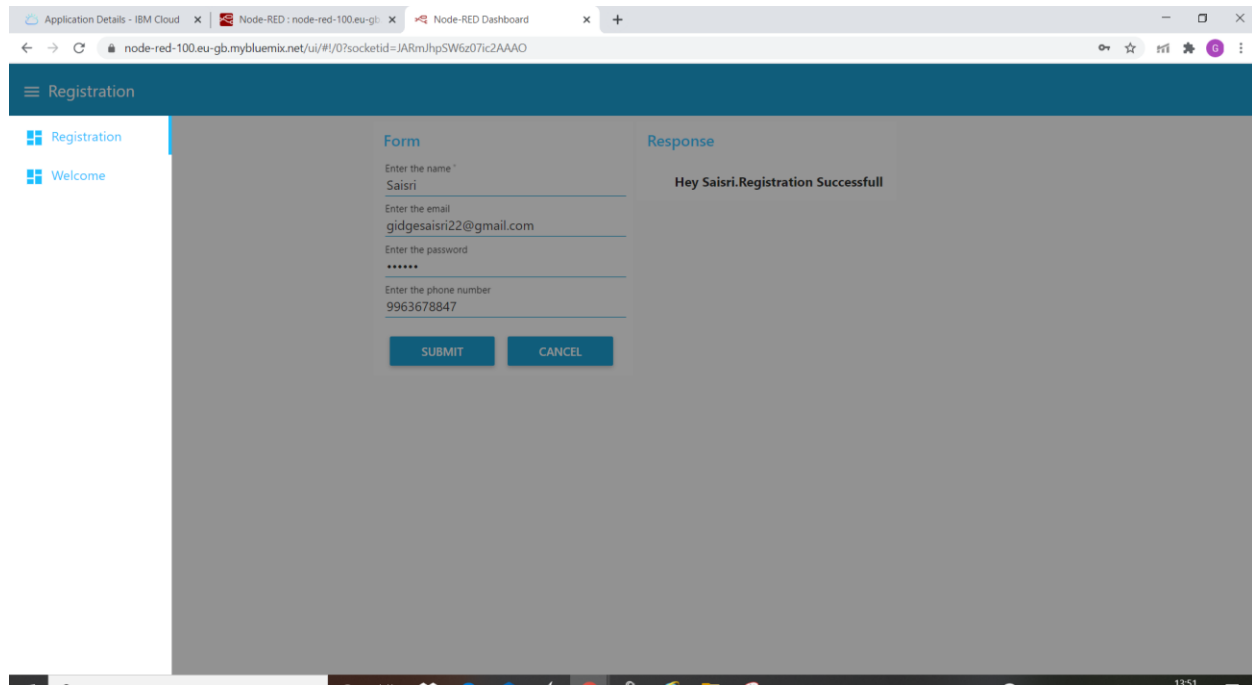
[View logs and history](#)

LAST EXECUTION RESULT

Node RED 100 [View console](#)

Build 3

Creating a web page using node-red



BUILDING OWN MODEL IN IBM WATSON STUDIO: salary data:

The screenshot shows an Excel spreadsheet titled 'Salary_Data'. The spreadsheet contains a table with 28 rows of data. The first column is labeled 'YearsExpe' and the second column is labeled 'Salary'. The data shows a positive correlation between experience and salary.

YearsExpe	Salary
1.1	39343
1.3	46205
1.5	37731
2	43525
2.2	39891
2.9	56642
3	60150
3.2	54445
3.2	64445
3.7	57189
3.9	63218
4	55794
4	56957
4.1	57081
4.5	61111
4.9	67938
5.1	66029
5.3	83088
5.9	81363
6	93940
6.8	91738
7.1	98273
7.9	101302
8.2	113812
8.7	109431
9	105582
9.5	116060

Service Details - IBM Cloud x IBM Watson Studio x Student Dashboard x IISPS_INT_2661_Predicting Life E: x +

eu-gb.dataplatform.cloud.ibm.com/projects/77d3ba46-14d8-4c9f-8c1f-a96490a7e3fd/assets?context=wdp

IBM Watson Studio Upgrade Saisri Gidge's Account SG

My projects / Salary_data

Launch IDE Add to project +

Notebooks

New notebook +

Name	Shared	Scheduled	Status	Language	Last editor	Last modified
Linear Regression Actual (2)				Python 3.6	Saisri Gidge	Jul 03, 2020

Deep learning experiments

New deep learning experiment +

Name	Last Modified
You don't have any Deep learning experiments yet	

Models

Watson Machine Learning model Import model +

Name	Type	Runtime	Last modified
SalaryAutoai - P4 ExtraTreesRegressorEstimator	wml-hybrid_0.1	hybrid_0.1	Jul 03, 2020

Data Load Files Catalog

Drop files here or [browse](#) for files to upload.

Type here to search

PIPELINES:

Service Details - IBM Cloud x IBM Watson Studio x IBM Watson Studio x IISPS_INT_2661_Predicting Life E: x +

eu-gb.dataplatform.cloud.ibm.com/ml/auto-ml/8436e140-0997-4772-9c88-d8fda11a36af/train?projectId=77d3ba46-14d8-4c9f-8c1f-a96490a7e3fd&mlInstanceGuid=a4a8cfa2-cdcf-4784-bb...

IBM Watson Studio Upgrade Saisri Gidge's Account SG

My projects / Salary_data / SalaryAutoai

Experiment summary Pipeline comparison Rank by: Root mean squared err... Score: Cross validation Holdout

Experiment completed 8 PIPELINES GENERATED

8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.

Time elapsed: 3 minutes

View full log

Pipeline leaderboard

Rank	Name	Algorithm	RMSE (Optimized)	Enhancements	Build time
1	Pipeline 4	Extra Trees Regressor	4954.377	HPO-1 FE HPO-2	00:00:12

Type here to search

AUTOMATE MACHINELEARNING MODEL:

The screenshot shows the IBM Watson Studio interface for a project named "Salary_deploy". The "Test" tab is active, displaying a form to "Enter input data". The "YearsExperience" field is set to 10. A "Predict" button is visible. The output shows a JSON response with a prediction value of 117242.85357142857.

```
{
  "predictions": [
    {
      "fields": [
        "prediction"
      ],
      "values": [
        117242.85357142857
      ]
    }
  ]
}
```

node red deploy for salary data:

The screenshot shows the Node-RED interface with a flow titled "Flow 1". The flow includes a "form" node, a "timestamp" node, a "PreToken" function node, an "http request" node, a "Pre Prediction" function node, and a "msg.payload" node. The "PreToken" function node is connected to the "http request" node, which is connected to the "Pre Prediction" function node. The "Pre Prediction" function node is connected to the "msg.payload" node. The "PreToken" function node is also connected to a "Salary" node. The "Pre Prediction" function node is connected to a "Salary" node. The "PreToken" function node is connected to a "Salary" node. The "Pre Prediction" function node is connected to a "Salary" node.

PREDICTING LIFE EXPECTANCY USING MACHINE LEARNING:

Collection of dataset:

Country	Year	Status	Life expect	Adult Mor	infant dea	r Alcohol	percentag	Hepatitis E	Measles	BMI	under-five	Polio	Total expe	Diphtheria	HIV/AIDS	GDP	Population	thinness	thinness 5	Income co	Schooling
Afghanistan	2015	Developing	65	263	62	0.01	71.27962	65	1154	19.1	83	6	8.16	65	0.1	584.2592	33736494	17.2	17.3	0.479	10.1
Afghanistan	2014	Developing	59.9	271	64	0.01	73.52358	62	492	18.6	86	58	8.18	62	0.1	612.6965	327582	17.5	17.5	0.476	10
Afghanistan	2013	Developing	59.9	268	66	0.01	73.21924	64	430	18.1	89	62	8.13	64	0.1	631.745	31731688	17.7	17.7	0.47	9.9
Afghanistan	2012	Developing	59.5	272	69	0.01	78.18422	67	2787	17.6	93	67	8.52	67	0.1	669.959	3696958	17.9	18	0.463	9.8
Afghanistan	2011	Developing	59.2	275	71	0.01	7.097109	68	3013	17.2	97	68	7.87	68	0.1	63.53723	2978599	18.2	18.2	0.454	9.5
Afghanistan	2010	Developing	58.8	279	74	0.01	56.76937	66	1989	16.7	102	66	9.2	66	0.1	553.3289	2883167	18.4	18.4	0.448	9.2
Afghanistan	2009	Developing	58.6	281	77	0.01	56.76222	63	2861	16.2	106	63	9.42	63	0.1	445.8933	284331	18.6	18.7	0.434	8.9
Afghanistan	2008	Developing	58.1	287	80	0.03	25.87393	64	1599	15.7	110	64	8.33	64	0.1	373.3611	2729431	18.8	18.9	0.433	8.7
Afghanistan	2007	Developing	57.5	295	82	0.02	10.91016	63	1141	15.2	113	63	6.73	63	0.1	369.8358	26616792	19	19.1	0.415	8.4
Afghanistan	2006	Developing	57.3	295	84	0.03	17.17152	64	1990	14.7	116	58	7.43	58	0.1	272.5638	2589345	19.2	19.3	0.405	8.1
Afghanistan	2005	Developing	57.3	291	85	0.02	1.388648	66	1296	14.2	118	58	8.7	58	0.1	25.29413	257798	19.3	19.5	0.396	7.9
Afghanistan	2004	Developing	57	293	87	0.02	15.29607	67	466	13.8	120	5	8.79	5	0.1	219.1414	24118979	19.5	19.7	0.381	6.8
Afghanistan	2003	Developing	56.7	295	87	0.01	11.08905	65	798	13.4	122	41	8.82	41	0.1	198.7285	2364851	19.7	19.9	0.373	6.5
Afghanistan	2002	Developing	56.2	3	88	0.01	16.88735	64	2486	13	122	36	7.76	36	0.1	187.846	21979923	19.9	2.2	0.341	6.2
Afghanistan	2001	Developing	55.3	316	88	0.01	10.57473	63	8762	12.6	122	35	7.8	33	0.1	117.497	2966463	2.1	2.4	0.34	5.9
Afghanistan	2000	Developing	54.8	321	88	0.01	10.42496	62	6532	12.2	122	24	8.2	24	0.1	114.56	293756	2.3	2.5	0.338	5.5
Albania	2015	Developing	77.8	74	0	4.6	364.9752	99	0	58	0	99	6	99	0.1	3954.228	28873	1.2	1.3	0.762	14.2
Albania	2014	Developing	77.5	8	0	4.51	428.7491	98	0	57.2	1	98	5.88	98	0.1	4575.764	288914	1.2	1.3	0.761	14.2
Albania	2013	Developing	77.2	84	0	4.76	430.877	99	0	56.5	1	99	5.66	99	0.1	4414.723	289592	1.3	1.4	0.759	14.2
Albania	2012	Developing	76.9	86	0	5.14	412.4434	99	9	55.8	1	99	5.59	99	0.1	4247.614	2941	1.3	1.4	0.752	14.2
Albania	2011	Developing	76.6	88	0	5.37	437.0621	99	28	55.1	1	99	5.71	99	0.1	4437.179	295195	1.4	1.5	0.738	13.3
Albania	2010	Developing	76.2	91	1	5.28	41.82276	99	10	54.3	1	99	5.34	99	0.1	494.3588	291321	1.4	1.5	0.725	12.5
Albania	2009	Developing	76.1	91	1	5.79	348.056	98	0	53.5	1	98	5.79	98	0.1	4114.137	2927519	1.5	1.6	0.721	12.2
Albania	2008	Developing	75.3	1	1	5.61	36.62207	99	0	52.6	1	99	5.87	99	0.1	437.5396	2947314	1.6	1.6	0.713	12
Albania	2007	Developing	75.0	9	1	5.58	32.24655	98	22	51.7	1	99	6.1	98	0.1	363.1369	29717	1.6	1.7	0.703	11.6
Albania	2006	Developing	74.2	99	1	5.31	3.302154	98	68	5.8	1	97	5.86	97	0.1	35.1293	2992547	1.7	1.8	0.696	11.4
Albania	2005	Developing	73.5	15	1	5.16	36.90412	98	6	49.9	1	97	6.12	98	0.1	779.1470	311487	1.8	1.8	0.685	10.8

Creating Necessary cloud service:

Name	Group	Location	Offering	Status	Tags
node-red-100-cloudant-1592300546824-65761	gldgesaisri22@gmail.com / dev	London	Cloudant	Provisioned	-
Continuous Delivery	Default	London	Continuous Delivery	Active	-
Watson Studio-12	Default	London	Watson Studio	Active	anuc...
node-red-100-cloudant-1592300546824	Default	London	Cloudant	Active	-
pm-20-lifec	Default	London	Machine Learning	Active	cpda...

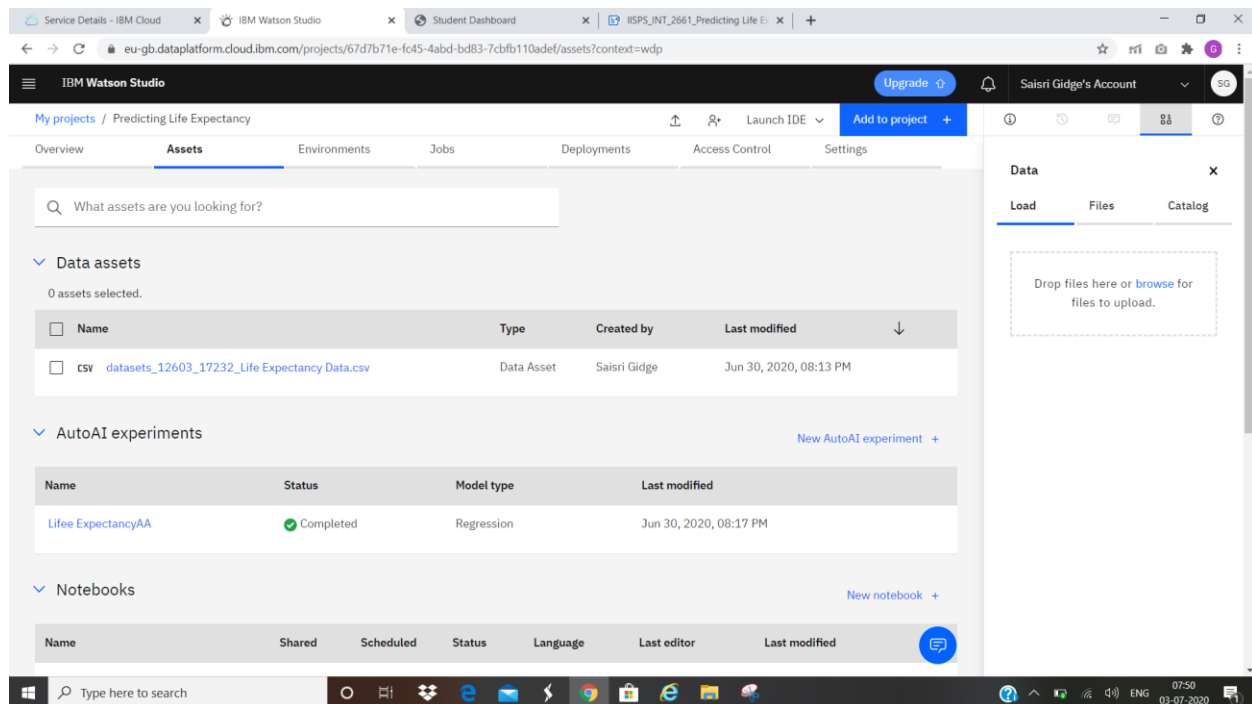
CREATE A WATSON STUDIO PROJECT:

The screenshot shows the IBM Watson Studio web interface. The browser address bar displays the URL: `eu-gb.dataplatform.cloud.ibm.com/projects/67d7b71e-fc45-4abd-bd83-7cbb110adef/assets?context=wdp`. The page header includes the IBM Watson Studio logo, an 'Upgrade' button, and the user's account name 'Saisri Gidge's Account'. The main navigation bar shows tabs for 'Overview', 'Assets', 'Environments', 'Jobs', 'Deployments', 'Access Control', and 'Settings'. The 'Assets' tab is active, displaying a search bar and a list of data assets. A table lists one asset: 'datasets_12603_17232_Life Expectancy Data.csv', which is a 'Data Asset' created by 'Saisri Gidge' on 'Jun 30, 2020, 08:13 PM'. Below the assets, there is a section for 'AutoAI experiments' showing a completed experiment named 'Life ExpectancyAA' with a status of 'Completed' and a model type of 'Regression'. A 'Notebooks' section is also visible at the bottom. On the right side, a 'Data' panel is open, showing options to 'Load', 'Files', or 'Catalog' data, with a prompt to 'Drop files here or browse for files to upload.'

CREATE A MACHINELEARNING SERVICE:

The screenshot shows the IBM Cloud console interface. The browser address bar displays the URL: `cloud.ibm.com/services/pm-20/crn%3Av1%3Abluemix%3Apublic%3Apm-20%3Aeu-gb%3Aa%2F7f49fbdaed34ac7b5730034f69e6de7%3Aa4a8cfa2-cdcf-4784-bb82-1714040a0dd5%3A...`. The page header includes the IBM Cloud logo, a search bar, and navigation links for 'Catalog', 'Docs', 'Support', and 'Manage'. The user's account name 'Saisri Gidge's Account' is also visible. The main content area shows the 'Watson Machine Learning' service page. It features a large purple logo and the text 'Welcome! Get Started with Watson Machine Learning in Watson Studio.' Below this, there is a button labeled 'Access in Watson Studio'. The page also includes sections for 'Documentation' and 'Community'. The left sidebar shows a 'Resource list' with the service 'pm-20-lifee' listed as 'Active' with a 'cpdaas' icon. The bottom of the page shows a Windows taskbar with various application icons and the system clock displaying '18:27 02-07-2020'.

IMPORT DATASET AND CREATE AUTOAI EXPERIMENT:

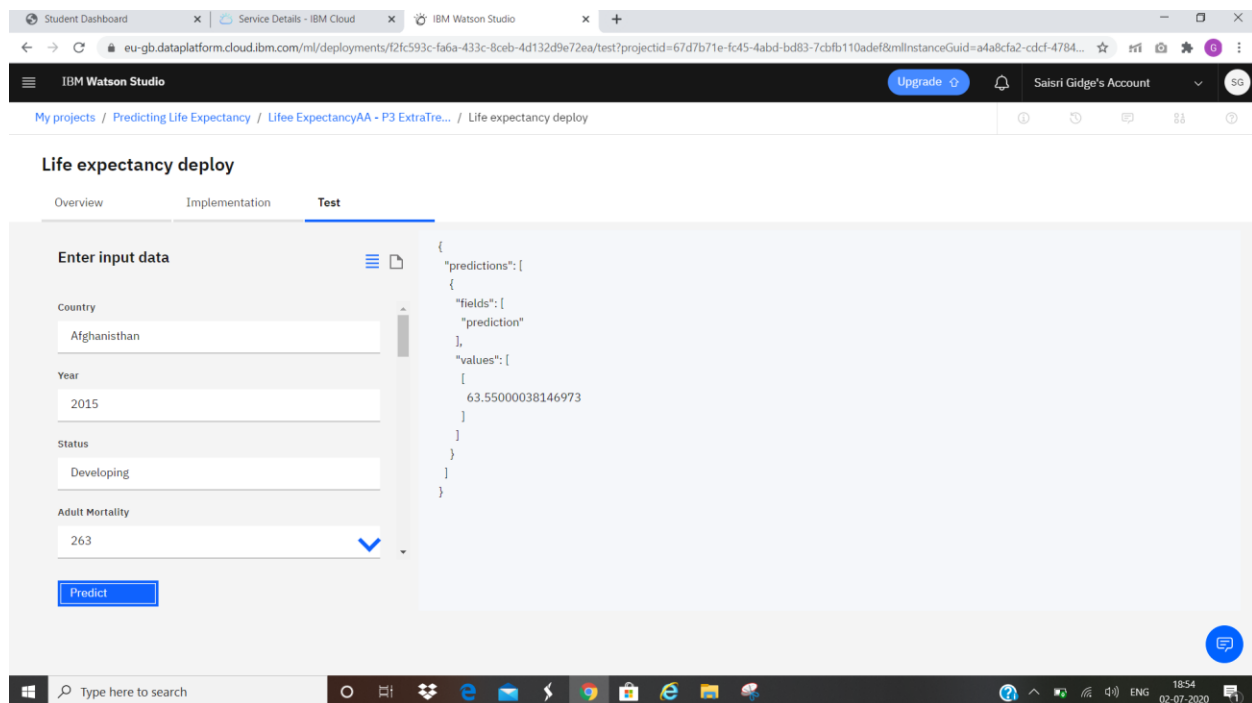


The screenshot shows the IBM Watson Studio interface. The top navigation bar includes 'Service Details - IBM Cloud', 'IBM Watson Studio', 'Student Dashboard', and 'IISPS_INT_2661_Predicting Life E...'. The main header shows 'My projects / Predicting Life Expectancy' with tabs for Overview, Assets, Environments, Jobs, Deployments, Access Control, and Settings. The 'Assets' tab is active, displaying a search bar and a table of data assets. A table titled 'Data assets' shows 0 assets selected. Below it, a table lists assets with columns: Name, Type, Created by, and Last modified. One asset is listed: 'csv datasets_12603_17232_Life Expectancy Data.csv' by 'Saisri Gidge' on 'Jun 30, 2020, 08:13 PM'. Below the assets table, there is a section for 'AutoAI experiments' with a 'New AutoAI experiment +' button. A table lists experiments with columns: Name, Status, Model type, and Last modified. One experiment is listed: 'Life ExpectancyAA' with status 'Completed', model type 'Regression', and last modified 'Jun 30, 2020, 08:17 PM'. Below the experiments table, there is a section for 'Notebooks' with a 'New notebook +' button. A table lists notebooks with columns: Name, Shared, Scheduled, Status, Language, Last editor, and Last modified. On the right side, a 'Data' panel is open, showing 'Load', 'Files', and 'Catalog' tabs. The 'Load' tab is active, displaying a message: 'Drop files here or browse for files to upload.'

Name	Type	Created by	Last modified
csv datasets_12603_17232_Life Expectancy Data.csv	Data Asset	Saisri Gidge	Jun 30, 2020, 08:13 PM

Name	Status	Model type	Last modified
Life ExpectancyAA	Completed	Regression	Jun 30, 2020, 08:17 PM

AUTOAI:



The screenshot shows the IBM Watson Studio interface for the 'Life expectancy deploy' project. The top navigation bar includes 'Student Dashboard', 'Service Details - IBM Cloud', 'IBM Watson Studio', and a project ID. The main header shows 'My projects / Predicting Life Expectancy / Life ExpectancyAA - P3 ExtraTre... / Life expectancy deploy'. The 'Test' tab is active, displaying a form to 'Enter input data' and a JSON output. The form has fields for Country (Afghanistan), Year (2015), Status (Developing), and Adult Mortality (263). A 'Predict' button is at the bottom. The JSON output shows a prediction for the life expectancy of a person in Afghanistan in 2015, with a value of 63.5500038146973.

Life expectancy deploy

Overview Implementation **Test**

Enter input data

Country: Afghanistan

Year: 2015

Status: Developing

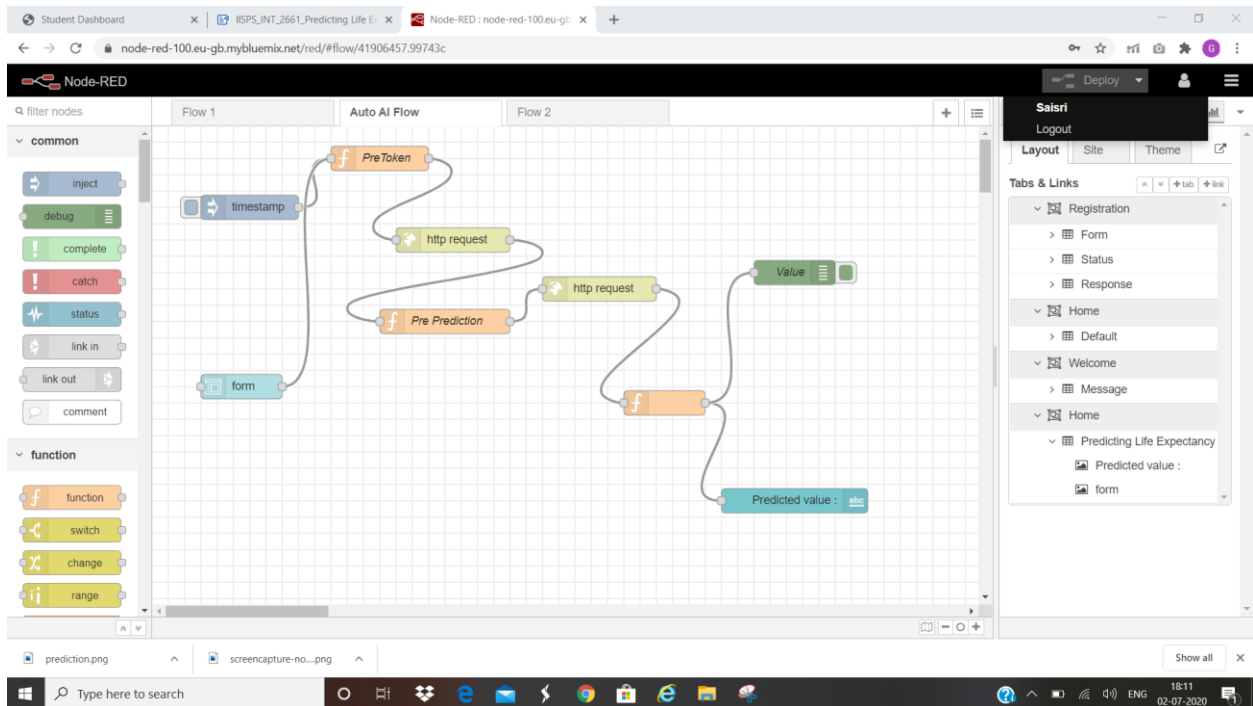
Adult Mortality: 263

Predict

```
{
  "predictions": [
    {
      "fields": [
        "prediction"
      ],
      "values": [
        63.5500038146973
      ]
    }
  ]
}
```

BUILD NODERED FLOW AND INTEGRATE INTO AUTOAI:

NODERED FLOW:



Salary prediction:

Home

Predicting Life Expectancy

Predicted value : 63.55000038146973

Country

Afghanistan

Year

2015

Status

developing

Adult Mortality

263

infant deaths

62

Alcohol

0.01

percentage expenditure

71.23435

Hepatitis B

65

Measles

1154

BMI

19.1

under-five deaths

83

Polio

6

Total expenditure

8.16

Diphtheria

65

HIV/AIDS

0.1

GDP

584.2567

Population

33736494

thinness 1-19 years

17.2

thinness 5-9 years

17.3

Income composition of resources

0.479

Schooling

10.1

PREDICT

CANCEL