

Project Name : Predicting Life Expectancy using Machine Learning

Project Member : Arun G K

Project Mentor : Hemant Kumar Gahlot (Smartbridge)

Date : 12.06.2020

TSK-50581: Project Scope, Schedule, Team & Deliverables :

Project Summary :

A typical Regression Machine Learning project leverages historical data to predict insights into the future. This problem statement is aimed at predicting Life Expectancy rate of a country given various features.

Life expectancy is a statistical measure of the average time a human being is expected to live, Life expectancy depends on various factors: Regional variations, Economic Circumstances, Sex Differences, Mental Illnesses, Physical Illnesses, Education, Year of their birth and other demographic factors. 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 :

- Zoho Writer
- Github repository
- Slack channel

Technical Requirements :

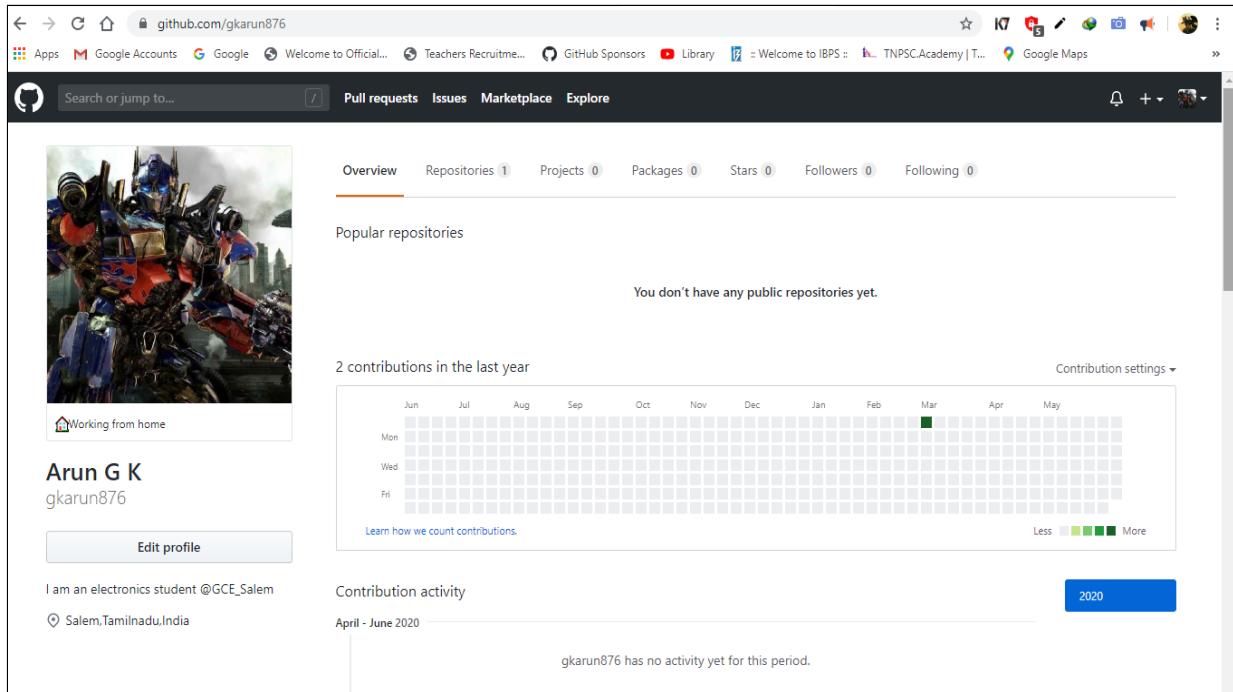
- Python
- IBM Cloud
- IBM Watson
- Node-Red
- Machine Learning

Project Team and Duration :

- Arun G K (30 days)

TSK-50582 - Setup The Development Environment :

1) Github account URL : <https://github.com/gkarun876>



The screenshot shows the GitHub profile page for user gkarun876. The profile picture is a Transformer character. The bio states: "I am an electronics student @GCE_Salem". The location is "Salem,Tamilnadu,India". The page shows 2 contributions in the last year, with a calendar view indicating activity in March. The contribution activity section shows no activity for April - June 2020.

Overview Repositories 1 Projects 0 Packages 0 Stars 0 Followers 0 Following 0

Popular repositories

You don't have any public repositories yet.

2 contributions in the last year Contribution settings

Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May

Mon
Wed
Fri

Learn how we count contributions.

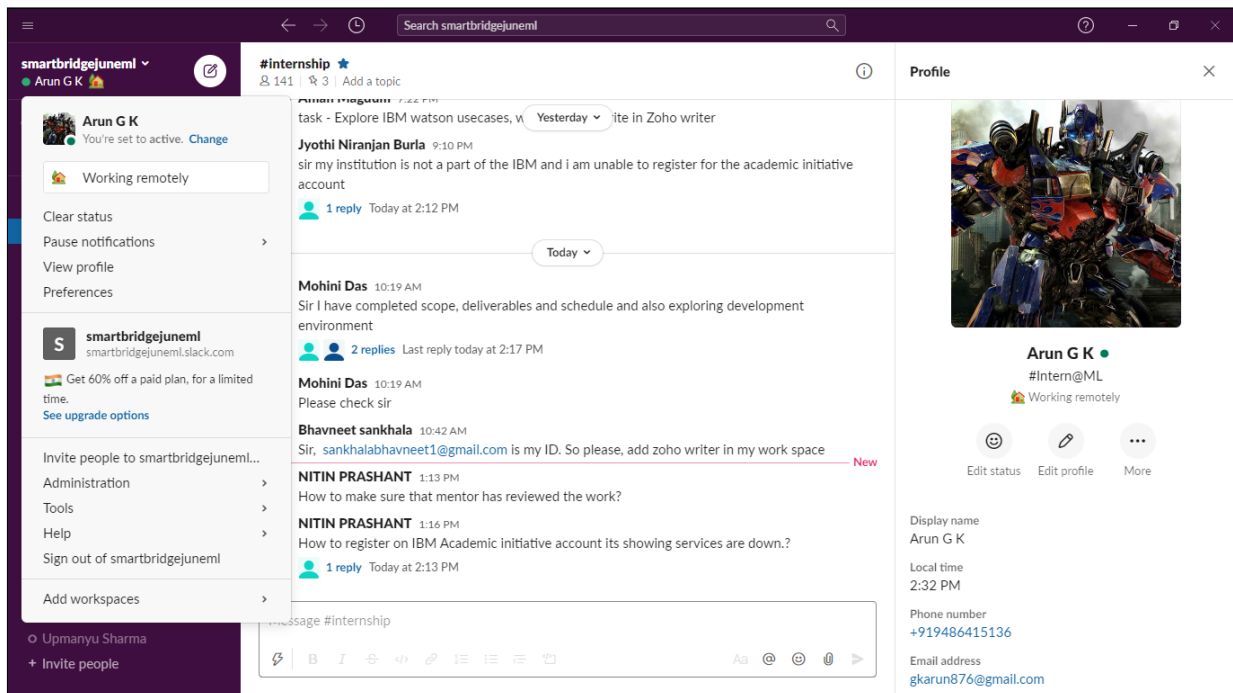
Less More

Contribution activity

April - June 2020

gkarun876 has no activity yet for this period.

2) Slack channel :

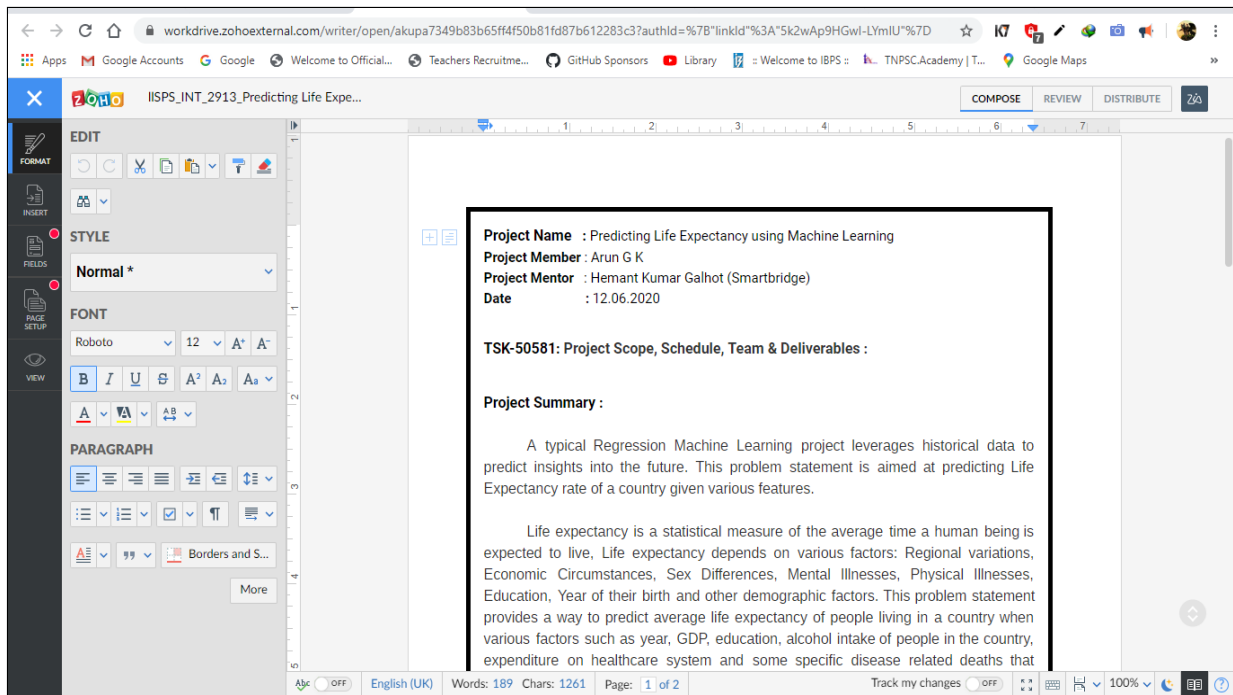


The screenshot shows a Slack channel named #internship. The channel has 141 members. The profile of Arun G K is visible on the right, showing a Transformer character profile picture, the name Arun G K, the handle #Intern@ML, and the status Working remotely. The channel messages include:

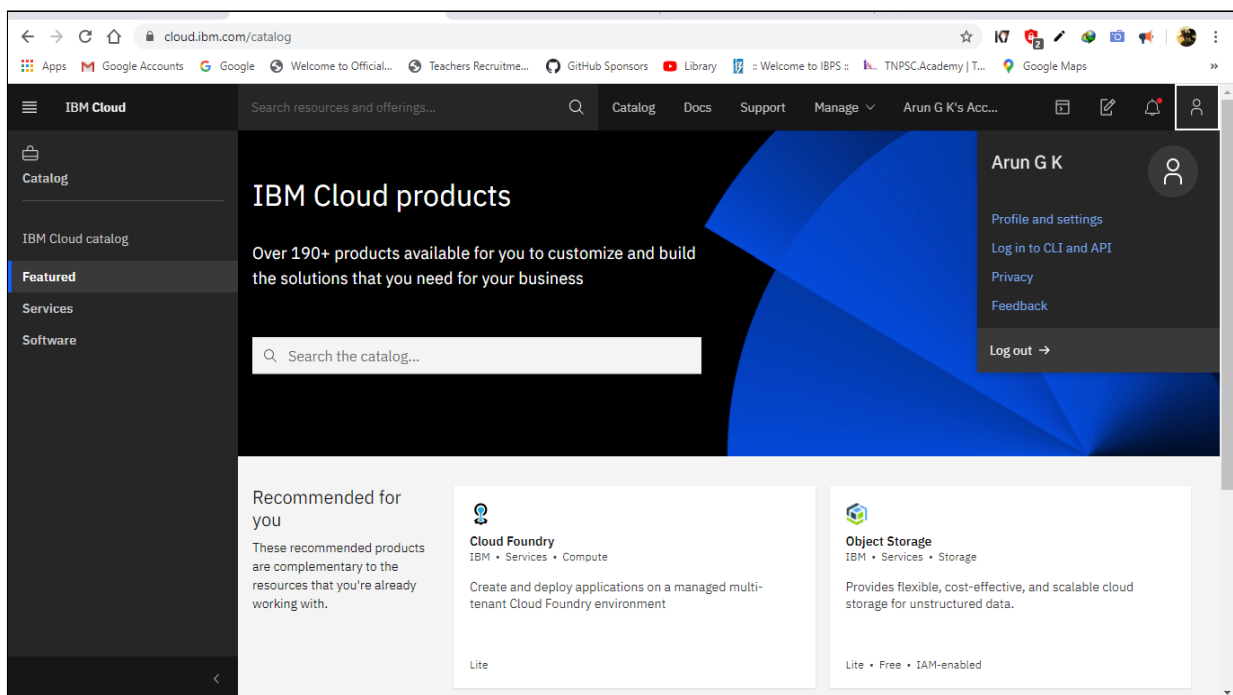
- Jyothi Niranjan Burla: sir my institution is not a part of the IBM and i am unable to register for the academic initiative account. 1 reply Today at 2:12 PM
- Mohini Das: Sir I have completed scope, deliverables and schedule and also exploring development environment. 2 replies Last reply today at 2:17 PM
- Mohini Das: Please check sir
- Bhavneet sankhala: Sir, sankhalabhavneet1@gmail.com is my ID. So please, add zoho writer in my work space. New
- NITIN PRASHANT: How to make sure that mentor has reviewed the work?
- NITIN PRASHANT: How to register on IBM Academic initiative account its showing services are down.? 1 reply Today at 2:13 PM

The left sidebar shows the user's profile, status (Working remotely), and a list of channels and workspaces.

3) Work With Document Writer :



TSK-50583 - Create IBM Cloud Account :



TSK-50584-Create a NODE-RED Starter Application :

The screenshot shows the IBM Cloud Developer console interface. The main heading is "Node RED LKWMG" with an "Add tags" link and an "Actions..." dropdown menu. The "Details" section on the left lists the following information:

- App URL: <https://node-red-lkwmg.eu-gb.mybluemix.net>
- Source: <https://eu-gb.git.cloud.ibm.com/gkarun876/NodeREDLKWMG>
- Resource group: Default
- Deployment target: Node RED LKWMG
- Created: 13/06/2020

The "Services" section shows a "Cloudant" service with links to "Open dashboard" and "Documentation", and a "Credentials" dropdown. Below this are buttons for "Connect existing services" and "Create service".

The "Deployment Automation" section on the right provides more details:

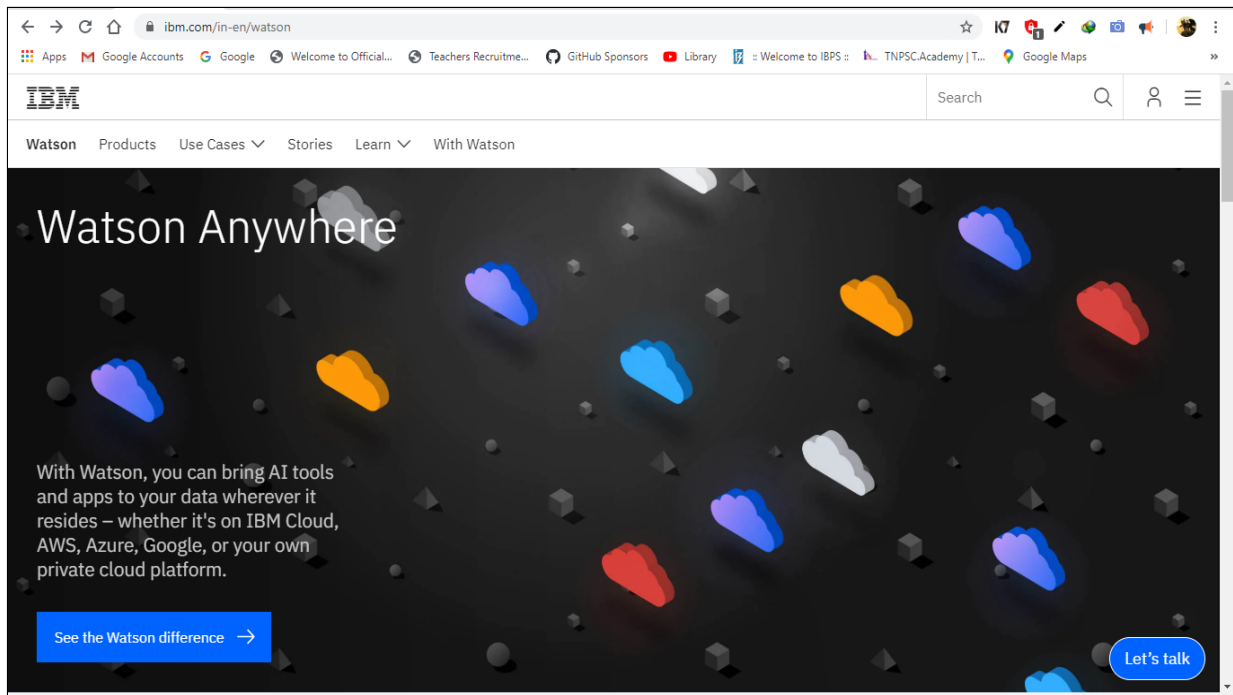
- Name: NodeREDLKWMG
- Location: London
- Tool integrations: (Icons for GitHub, Docker, etc.)
- Delivery Pipelines: Name: NodeREDLKWMG, Status: Success (green checkmark), Last input: Last commit by Arun G K (11 minutes ago), Clone from zip link.

A vertical "ASK A QUESTION" button is visible on the right side of the console.

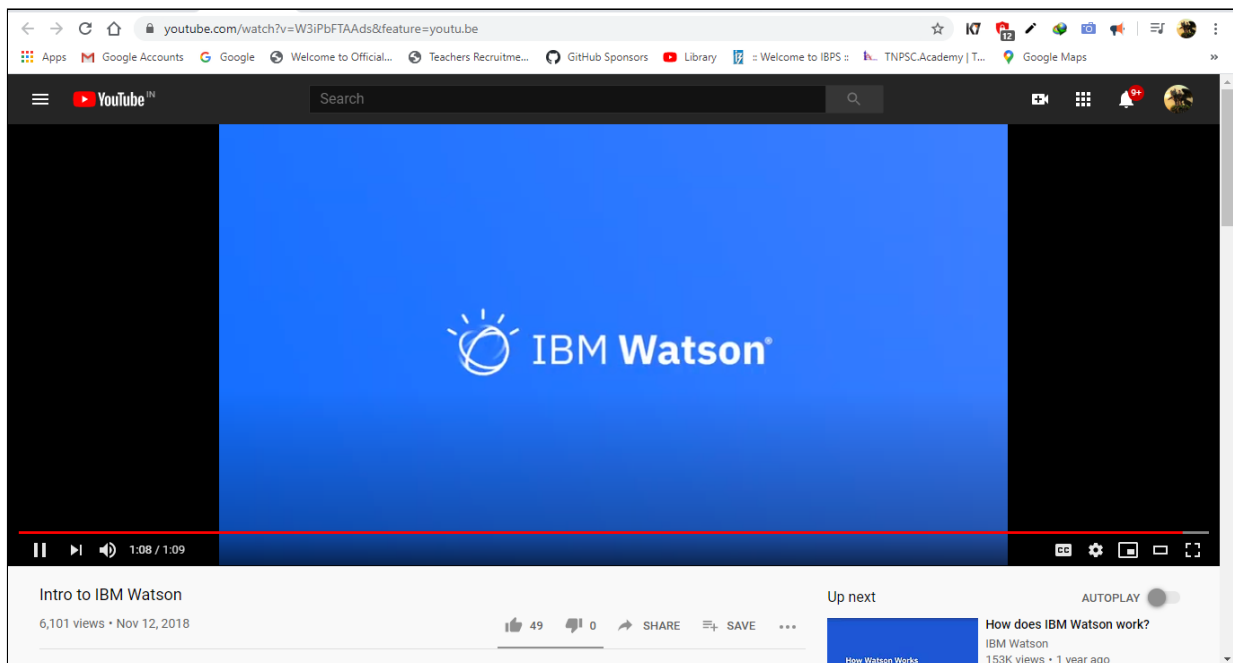
The screenshot shows the Node-RED web interface in a browser. The URL is <https://node-red-lkwmg.eu-gb.mybluemix.net/red/#flow/13bfeaa.057cf15>. The interface displays a flow named "Flow 1" on a grid. The flow consists of two nodes: a "Hello Node-RED!" message node (blue) and a "msg.payload" output node (green). The left sidebar shows a "common" section with nodes like inject, debug, complete, catch, status, link in, link out, and comment, and a "function" section with nodes like function, switch, and change. The right sidebar shows the "Deploy" button and a panel for the selected flow, displaying its ID ("13bfeaa.057cf15"), name ("Flow 1"), and status ("Enabled"). A tooltip at the bottom right of the right sidebar reads: "Hold down ctrl when you click on a node to add or remove it from the current selection".

TSK-50585-Explore IBM Watson Usecases :

Watson products and services :



Watson at work :



TSK-50586-Explore IBM Watson Machine Learning :

Introduction to Machine Learning :

The screenshot shows the 'Learning path: Get started with machine learning' page on the IBM Developer website. The page is for developers and provides a structured path to learn machine learning. It includes a sidebar with navigation links, a main content area with a table of learning levels, and a right sidebar with social media links and technology tags.

Articles
Models
Code Patterns
Podcasts
Open Project
Series
Tutorials
Videos

Community
Blog Posts
Announcements
Events

More resources
Data and analytics architecture
Data and analytics solutions
AI newsletters

Learning path: Get started with machine learning

Understand the principles of machine learning and get working knowledge of the different phases and tasks

By [Samaya Madhavan](#), [Mark Sturdevant](#)
Published December 4, 2019

Level	Topic	Type
100	Introduction to machine learning	Article
101	Build and test your first machine learning model using Python and scikit-learn	Tutorial+Notebook
201	Learn regression algorithms using Python and scikit-learn	Tutorial+Notebook
202	Learn classification algorithms using Python and scikit-learn	Tutorial+Notebook

Technologies (4)

- Artificial intelligence
- Data science
- Deep learning
- Machine learning

Table of Contents

Site feedback

About IBM Watson Machine Learning :

The screenshot shows a YouTube video player for the video 'IBM Watson Data and AI: Get Started With IBM Watson Machine Learning in IBM Cloud'. The video is from IBM Watson and has 1,425 views as of February 10, 2020. The video content shows a title card with the IBM logo and a description of the video's purpose. The video player includes standard controls like play, pause, and volume, as well as a description box below the video.

IBM Watson Data and AI

Get Started With IBM Watson Machine Learning in IBM Cloud

This video shows you how to provision the IBM Cloud services necessary to work with the Watson Machine Learning service.

© Copyright IBM Corporation 2020

IBM Watson Machine Learning: Get Started in IBM Cloud

1,425 views • Feb 10, 2020

Up next

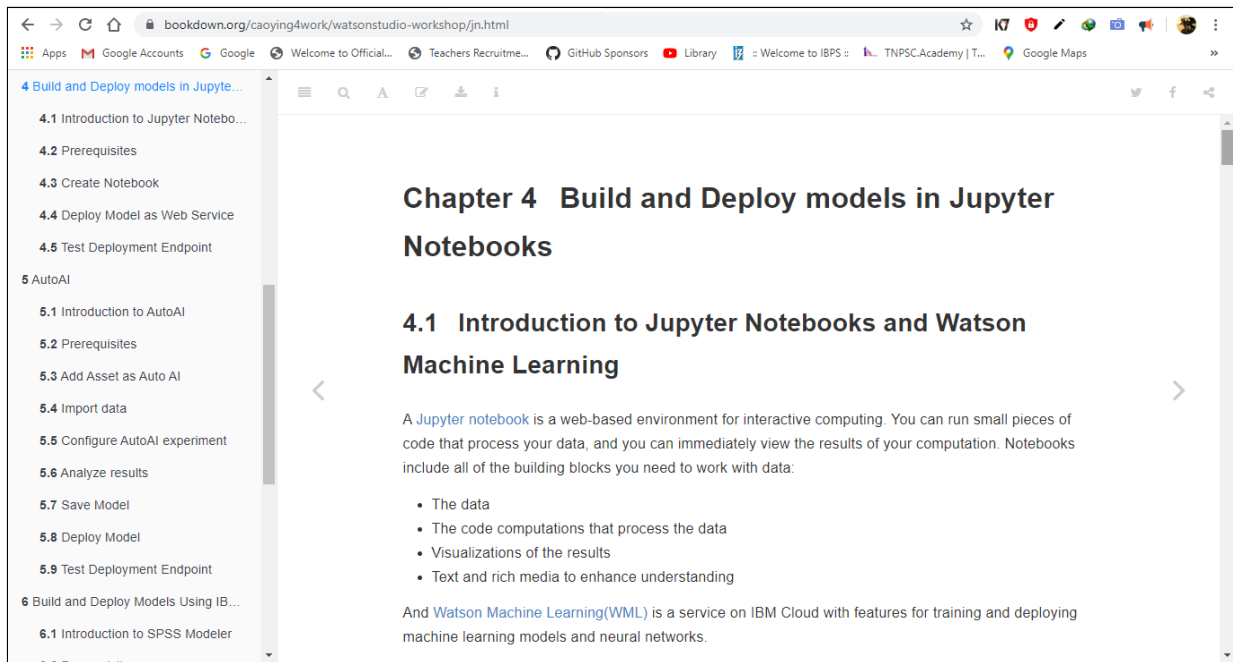
How does IBM Watson work?

IBM Watson

153K views • 1 year ago

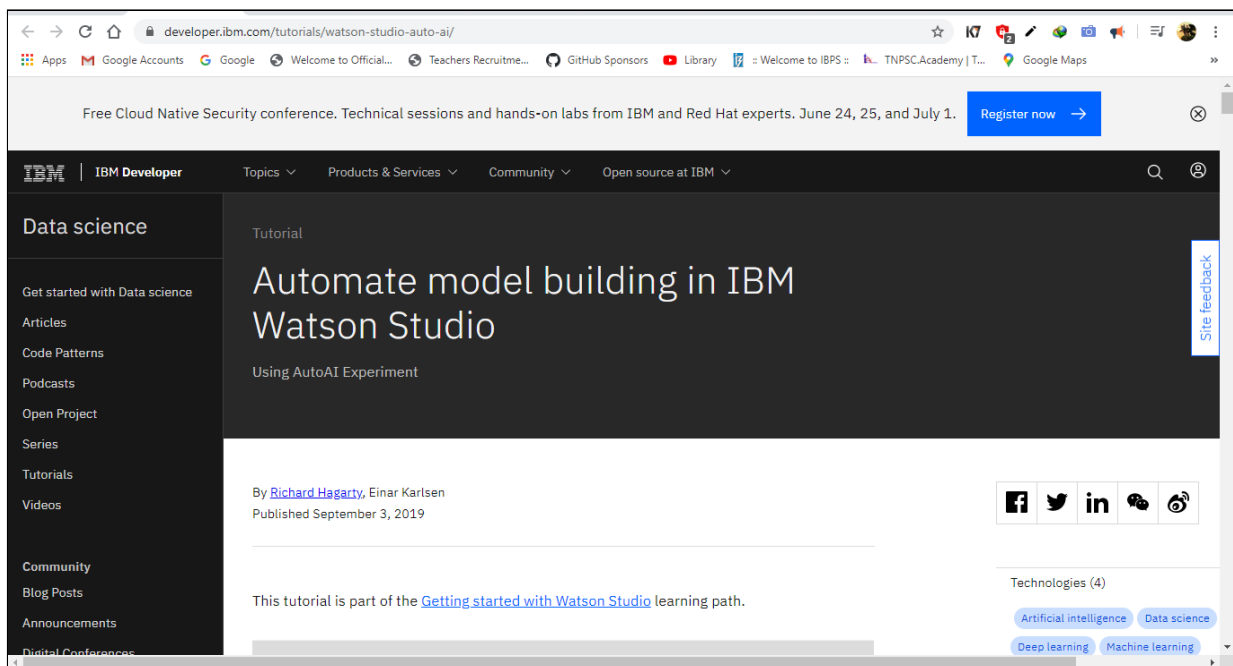
TSK-50587-Build your own ML Model in IBM Watson Studio :

Build Your Own ML Model IN IBM Watson Studio Using Machine Learning Service :



TSK-50588-Automate Your ML Model :

AutoAI Reference 1 :



AutoAI Reference 2 :

IBM Watson Studio

Create an AutoAI experiment

Define details

Experiment type

☒ From blank ☐ From sample

Name *

Name of AutoAI experiment

Description

Description of AutoAI experiment

Associate services

Machine Learning Service

No Machine Learning service instances associated with your project. Reassociate a Machine Learning service instance with your project on the project settings page, then click the reload button below to refresh the instances available for association with our new model builder instance.

Reload

Compute configuration *

8 vCPU and 32 GB RAM

This compute configuration consumes 20 capacity units per hour. [Learn more](#) about capacity unit hours and Watson Machine Learning pricing plans.

To run an AutoAI experiment, you'll need the Watson Machine Learning service.

IBM Watson Machine Learning: Run a sample AutoAI experiment to create a machine learning model

2,356 views • Feb 5, 2020

Up next

NLP Examples

Natural Language Processing in Python PyOhio

TSK-50589-Collect The Dataset For The Project :

kaggle

Search

Home

Compete

Data

Notebooks

Discuss

Courses

More

Recently Viewed

Life Expectancy (WHO)

Dataset

Life Expectancy (WHO)

Statistical Analysis on factors influencing Life Expectancy

KumarRajarshi • updated 2 years ago (Version 1)

Data Tasks Kernels (31) Discussion (10) Activity Metadata

Download (326 KB) New Notebook

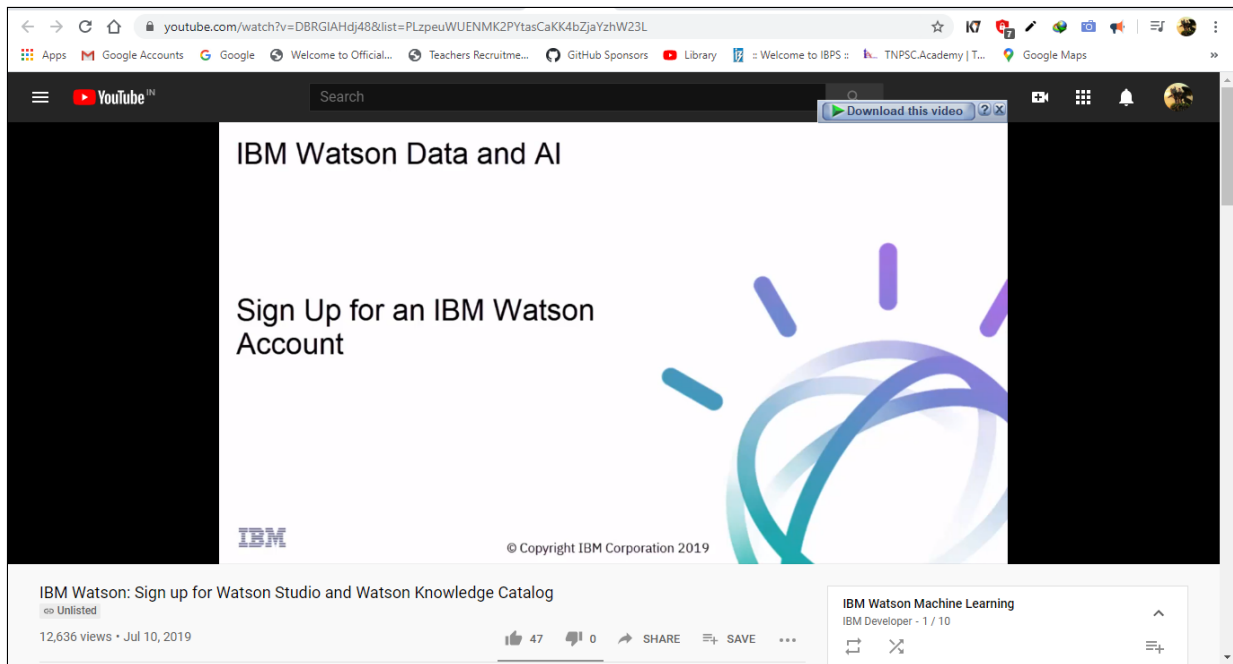
Usability 8.2 License Other (specified in description) Tags statistics, reference, society, health, social sciences and 4 more

Description

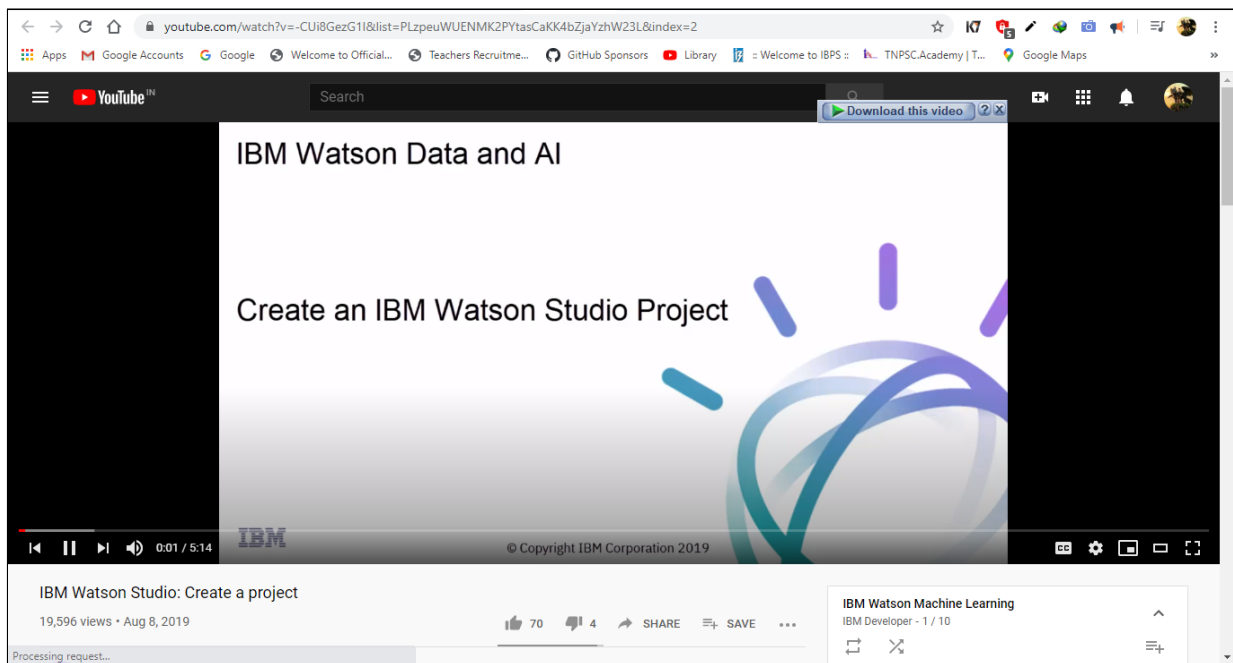
Context

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 affect 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 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

TSK-50590-Create Necessary IBM Cloud Services :



TSK-50591-Create a Watson Studio Project :



TSK-50592-Configure Watson Studio :

eu-gb.dataplatform.cloud.ibm.com/home?context=wdp

IBM Watson Studio

Create a project

Create a project, and then add the tools and assets you need.

Recently updated projects [View all \(1\)](#) [New project +](#)

Name	Role	Collaborators	Date created	Last updated
Predicting Life Expectancy using Machine Learning - SB1078	Admin	AK	Jun 14, 2020	Jun 15, 2020

Watson services [View all \(1\)](#) [Add service +](#)

Instance name	Service	Plan	Tool
Machine Learning-60	Machine Learning		

New in gallery [Explore](#)

NOTEBOOK + NOTEBOOK + SAMPLE PROJECT

TSK-50593-Create Machine Learning Service :

eu-gb.dataplatform.cloud.ibm.com/projects/b134b514-0b6a-4039-9f9e-d03ef1e1f3ea/settings?context=wdp

IBM Watson Studio

My projects / [Predicting Life Expectancy using ...](#) [Launch IDE](#) [Add to project +](#)

0 Byte used
Cloud Object Storage [Manage in IBM Cloud](#)

Associated services [Add service](#)

Name	Service type	Plan
Machine Learning-60	Watson - Machine Learning	

Access tokens [New token +](#)

Name	Role	Created	Last used
Token	Viewer	Jun 15, 2020, 09:42 AM	Never used

Integrations

<https://eu-gb.dataplatform.cloud.ibm.com/projects/b134b514-0b6a-4039-9f9e-d03ef1e1f3ea/assets?context=wdp#>

TSK-50594-Create A Jupyter Notebook In IBM Watson And Import Data :

eu-gb.dataplatform.cloud.ibm.com/analytics/notebooks/new-notebook?projectid=b134b514-0b6a-4039-9f9e-d03ef1e1f3ea&context=wdp

IBM Watson Studio

New notebook

Blank From file From URL

Name: Jupyter notebook

Select runtime: Default Python 3.6 XS (2 vCPU 8 GB RAM)

Description (optional): Notebook to predict life expectancy

The selected runtime has 2 vCPU and 8 GB RAM. It consumes 1 capacity unit per hour. [Learn more](#) about capacity unit hours and Watson Studio pricing plans.

Language: ☒ Python 3.6

Cancel Create

eu-gb.dataplatform.cloud.ibm.com/analytics/notebooks/v2/d8ca2f6c-b660-4435-a382-a90b8005f28f?projectid=b134b514-0b6a-4039...

IBM Watson Studio

My projects / Predicting Life Expectancy using ... / Jupyter notebook

File Edit View Insert Cell Kernel Help

Format Code

```
from urllib.request import urlopen
import json
from io import StringIO
import pandas as pd
from ibm_cloud_sdk_core.authenticators import OAuthenticator
from ibm_cloud_sdk_core import ApiException
from ibm_cloud_sdk_core import Config

body = client_3f6ed7db9e15428897097c8cc69bb2f8.get_object(Bucket='predictinglifeexpectancyusingmach-donotdelete-')
# 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]:

	Country	Year	Status	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	...	Polio	Total expenditure	Diphth
0	Afghanistan	2015	Developing	65.0	263.0	62	0.01	71.279624	65.0	1154	...	6.0	8.16	
1	Afghanistan	2014	Developing	59.9	271.0	64	0.01	73.523582	62.0	492	...	58.0	8.18	
2	Afghanistan	2013	Developing	59.9	268.0	66	0.01	73.219243	64.0	430	...	62.0	8.13	
3	Afghanistan	2012	Developing	59.5	272.0	69	0.01	78.184215	67.0	2787	...	67.0	8.52	
4	Afghanistan	2011	Developing	59.2	275.0	71	0.01	7.097109	68.0	3013	...	68.0	7.87	

5 rows x 22 columns

Data

Files

Upload one file at a time. All file types accepted. 5 GB max file size.

Drag and drop files here or upload.

Life Expectancy Data.csv

Insert to code

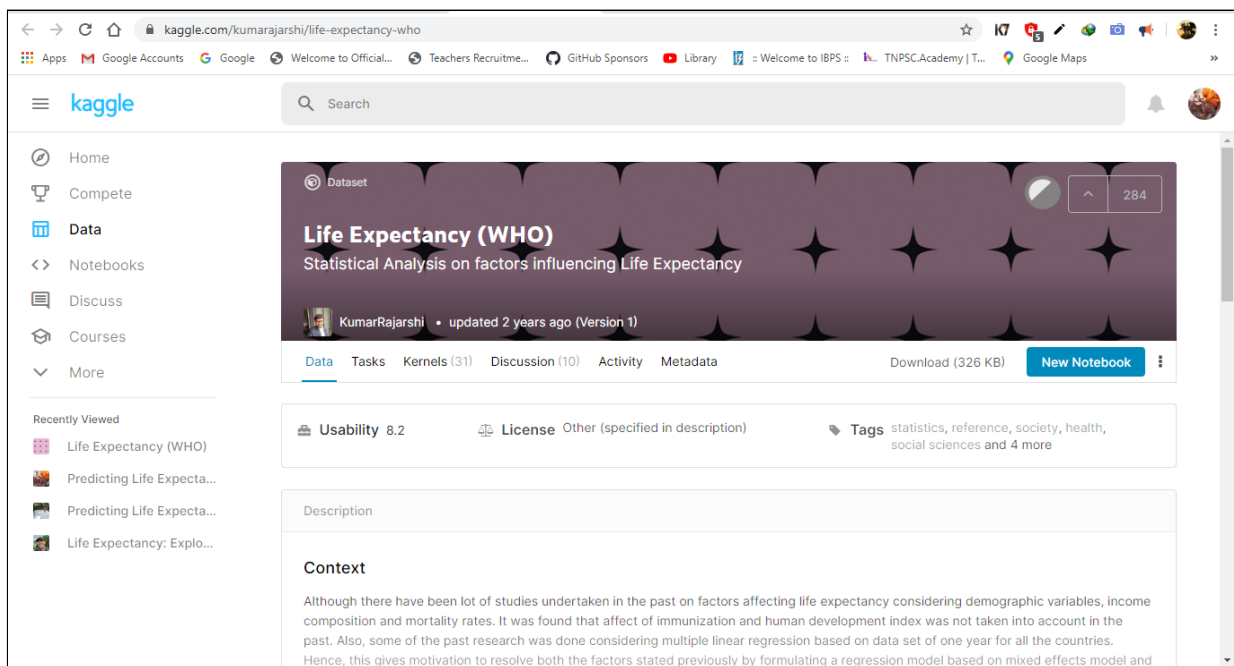
TSK-50595-Build A Machine Learning Model And Create Endpoints For Node-Red Integration:

I am going to build the model using Auto AI experiment.

TSK-50596-Build Node-Red Flow to Integrate Services :

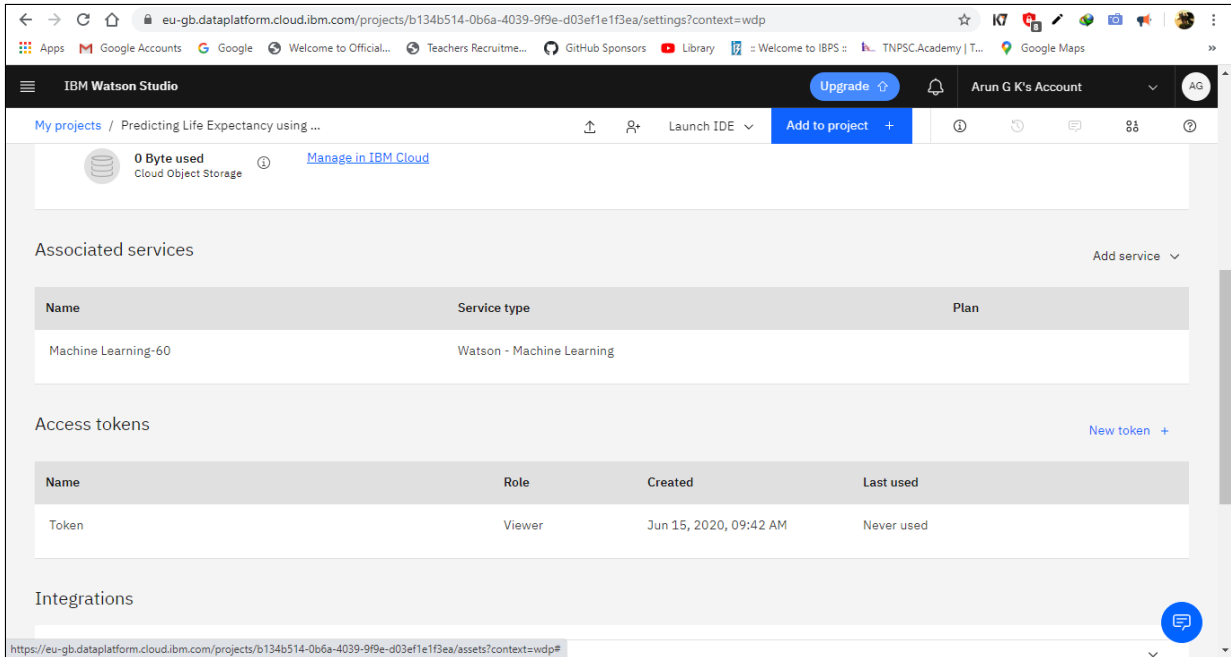
I am going to build the model using Auto AI experiment.

TSK-50597-Collect The Dataset For The Project :



The screenshot shows the Kaggle website interface. The browser address bar displays the URL: kaggle.com/kumarajarshi/life-expectancy-who. The Kaggle logo is in the top left, and a search bar is in the top right. On the left sidebar, navigation links include Home, Compete, Data, Notebooks, Discuss, Courses, and More. Below these is a 'Recently Viewed' section listing several datasets related to life expectancy. The main content area features a dataset card for 'Life Expectancy (WHO)' by KumarRajarshi, updated 2 years ago (Version 1). The card includes a title, subtitle 'Statistical Analysis on factors influencing Life Expectancy', and a 'Data' tab selected among others like Tasks, Kernels (31), Discussion (10), Activity, and Metadata. It also shows a download size of 326 KB and a 'New Notebook' button. Below the card, there are sections for 'Usability 8.2', 'License Other (specified in description)', and 'Tags' including statistics, reference, society, health, social sciences, and 4 more. The 'Description' section is partially visible, starting with 'Context' and mentioning studies on factors affecting life expectancy.

TSK-50598-Create Necessary IBM Cloud Services :

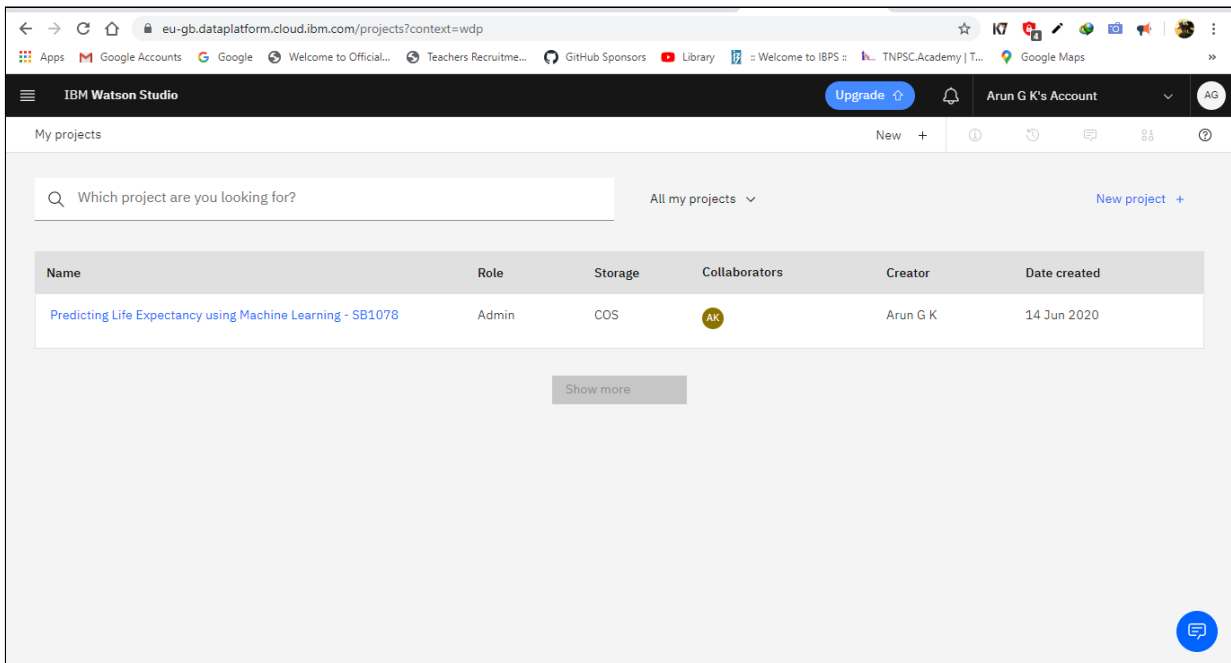


The screenshot shows the IBM Watson Studio interface for a project named "Predicting Life Expectancy using ...". The top navigation bar includes the IBM Watson Studio logo, an "Upgrade" button, and the user's account "Arun G K's Account". The main content area is divided into several sections:

- Cloud Object Storage:** Shows "0 Byte used" and a link to "Manage in IBM Cloud".
- Associated services:** A table with columns "Name", "Service type", and "Plan". It lists "Machine Learning-60" with "Watson - Machine Learning" as the service type.
- Access tokens:** A table with columns "Name", "Role", "Created", and "Last used". It shows a "Token" with a "Viewer" role, created on "Jun 15, 2020, 09:42 AM", and "Never used".
- Integrations:** A section with a blue chat icon.

The URL in the address bar is: <https://eu-gb.dataplatform.cloud.ibm.com/projects/b134b514-0b6a-4039-9f9e-d03ef1e1f3ea/settings?context=wdp>

TSK-50599-Create a Watson Studio Project :



The screenshot shows the IBM Watson Studio interface for the "My projects" page. The top navigation bar includes the IBM Watson Studio logo, an "Upgrade" button, and the user's account "Arun G K's Account". The main content area includes a search bar and a table of projects:

Search: Which project are you looking for? All my projects ▾ New project +

Name	Role	Storage	Collaborators	Creator	Date created
Predicting Life Expectancy using Machine Learning - SB1078	Admin	COS	AK	Arun G K	14 Jun 2020

Below the table is a "Show more" button. The URL in the address bar is: <https://eu-gb.dataplatform.cloud.ibm.com/projects?context=wdp>

TSK-50600-Configure Watson Studio :

eu-gb.dataplatform.cloud.ibm.com/home?context=wdp

IBM Watson Studio Upgrade Arun G K's Account AG

Create a project
Create a project, and then add the tools and assets you need.

Recently updated projects [View all \(1\)](#) [New project +](#)

Name	Role	Collaborators	Date created	Last updated
Predicting Life Expectancy using Machine Learning - SB1078	Admin	AK	Jun 14, 2020	Jun 15, 2020

Watson services [View all \(1\)](#) [Add service +](#)

Instance name	Service	Plan	Tool
Machine Learning-60	Machine Learning		

New in gallery [Explore](#)

NOTEBOOK + NOTEBOOK + SAMPLE PROJECT

TSK-50601-Create Machine Learning Service :

eu-gb.dataplatform.cloud.ibm.com/projects/b134b514-0b6a-4039-9f9e-d03ef1e1f3ea/settings?context=wdp

IBM Watson Studio Upgrade Arun G K's Account AG

My projects / Predicting Life Expectancy using ... Launch IDE Add to project +

0 Byte used Cloud Object Storage Manage in IBM Cloud

Associated services [Add service v](#)

Name	Service type	Plan
Machine Learning-60	Watson - Machine Learning	

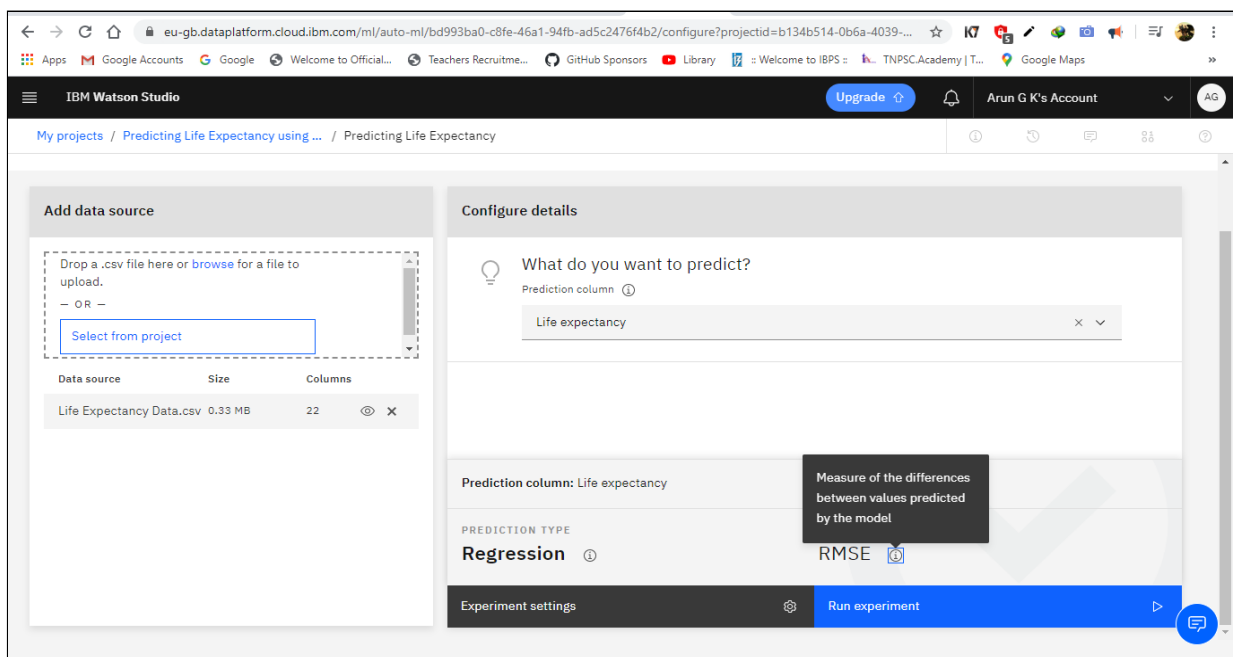
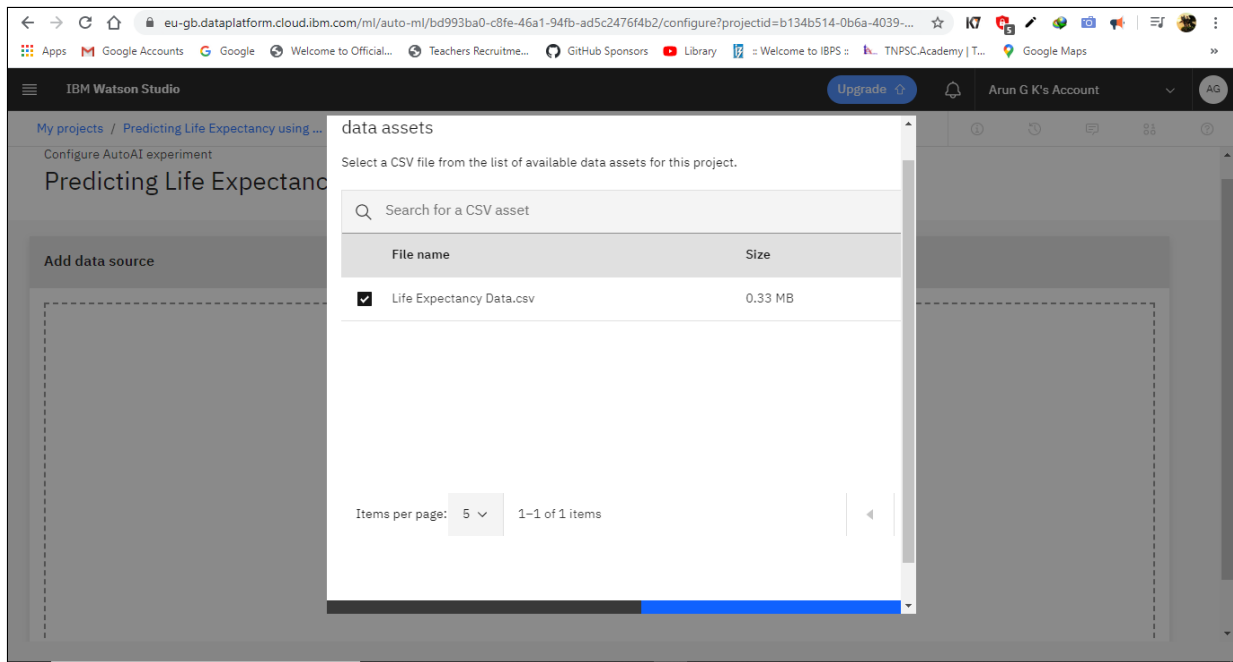
Access tokens [New token +](#)

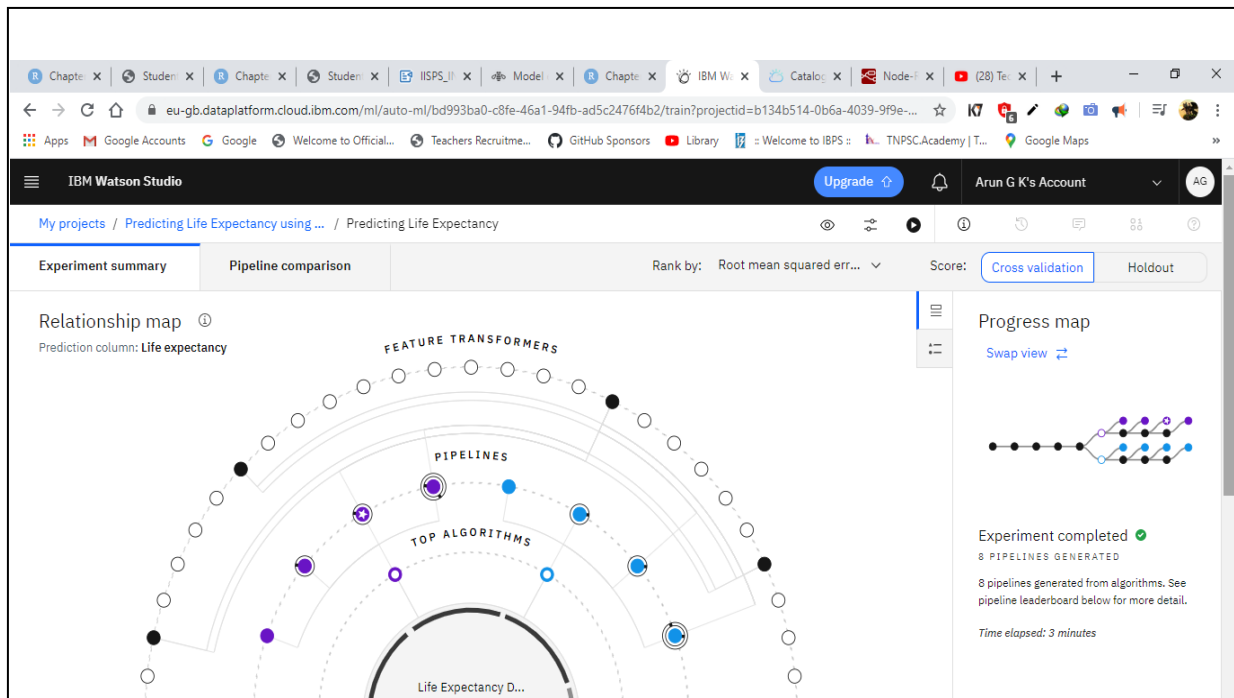
Name	Role	Created	Last used
Token	Viewer	Jun 15, 2020, 09:42 AM	Never used

Integrations

<https://eu-gb.dataplatform.cloud.ibm.com/projects/b134b514-0b6a-4039-9f9e-d03ef1e1f3ea/assets?context=wdp#>

TSK-50602 - Import Dataset And Create Auto AI experiment :





IBM Watson Studio

My projects / Predicting Life Expectancy using ... / Predicting Life Expectancy

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

Rank	↑	Name	Algorithm	RMSE (Optimized)	Enhancements	Build time
>	★ 1	Pipeline 3	Extra Trees Regressor	2.003	HPO-1 FE	00:00:46
>	2	Pipeline 4	Extra Trees Regressor	2.003	HPO-1 FE HPO-2	00:00:24
>	3	Pipeline 1	Extra Trees Regressor	2.070	None	00:00:01
>	4	Pipeline 2	Extra Trees Regressor	2.070	HPO-1	00:00:10
>	5	Pipeline 5	Decision Tree Regressor	2.807	None	00:00:01
>	6	Pipeline 6	Decision Tree Regressor	2.807	HPO-1	00:00:01
>	7	Pipeline 7	Decision Tree Regressor	2.811	HPO-1 FE	00:00:37
>	8	Pipeline 8	Decision Tree Regressor	2.811	HPO-1 FE HPO-2	00:00:06

Chapte x Studen x Chapte x Studen x IISPS_I x Model x Chapte x IBM W x Catalo x Node-i x (33) Te x + -

eu-gb.dataplatform.cloud.ibm.com/ml/auto-ml/bd993ba0-c8fe-46a1-94fb-ad5c2476f4b2/visualize?projectid=b134b514-0b6a-4039-9f9e-d...

Apps Google Accounts Google Welcome to Official... Teachers Recruitme... GitHub Sponsors Library Welcome to IBPS TNPSCAcademy | T... Google Maps

IBM Watson Studio Upgrade Arun G K's Account AG

My projects / Predicting Life Expectancy using ... / Predicting Life Expectancy

Back to Predicting Life Expectancy

Rank 1 Pipeline 3 Holdout RMSE (Optimized) 1.941 Algorithm Extra Trees Regressor Enhancements HPO-1 FE Build time 00:00:46 Save as

ExtraTreesRegressor

Model Evaluation Measures ⁱ

TARGET : LIFE EXPECTANCY

EVALUATION

Model Evaluation Measures

MODEL VIEWER

Model Information

Feature Transformations

Feature Importance

	Holdout Score	Cross Validation Score
Root Mean Squared Error (RMSE)	1.941	2.003
R ²	0.956	0.956
Explained Variance	0.956	0.956

Chapte x Studen x Chapte x Studen x IISPS_I x Model x Chapte x IBM W x Catalo x Node-i x (33) Te x + -

eu-gb.dataplatform.cloud.ibm.com/ml/auto-ml/bd993ba0-c8fe-46a1-94fb-ad5c2476f4b2/visualize?projectid=b134b514-0b6a-4039-9f9e-d...

Apps Google Accounts Google Welcome to Official... Teachers Recruitme... GitHub Sponsors Library Welcome to IBPS TNPSCAcademy | T... Google Maps

IBM Watson Studio Upgrade Arun G K's Account AG

My projects / Predicting Life Expectancy using ... / Predicting Life Expectancy

Back to Predicting Life Expectancy

Rank 1 Pipeline 3 Holdout RMSE (Optimized) 1.941 Algorithm Extra Trees Regressor Enhancements HPO-1 FE Build time 00:00:46 Save as

ExtraTreesRegressor

Model Information ⁱ

TARGET : LIFE EXPECTANCY

EVALUATION

Model Evaluation Measures

MODEL VIEWER

Model Information

Feature Transformations

Feature Importance

Label (Target)	Life expectancy
Number of Features	41
Created At	6/16/2020, 9:55:34 AM

IBM Watson Studio

My projects / Predicting Life Expectancy using ... / Predicting Life Expectancy - P3 E... / Auto AI Life expectancy deploym...

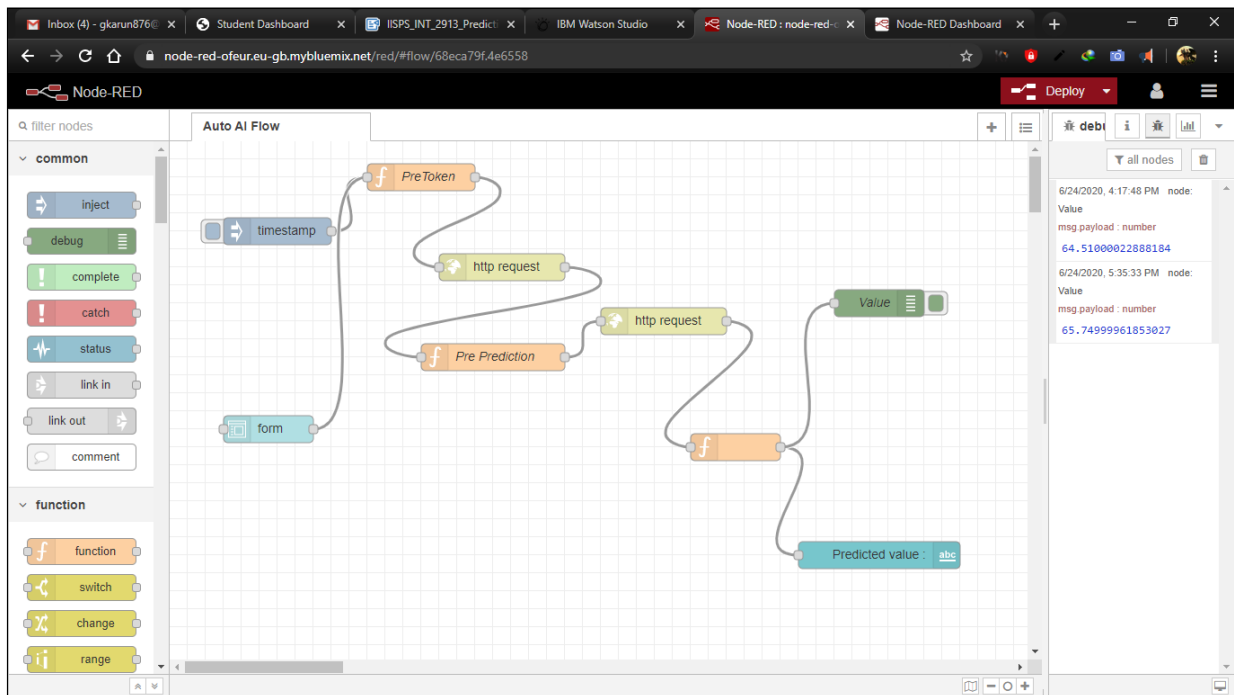
Auto AI Life expectancy deployment

Overview Implementation Test

Deployment

Name	Auto AI Life expectancy deployment
Type	Web Service
Deployment ID	7bfb7668-79ab-4f63-9ed4-f30a384e9a71
Status	Ready
Asset type	Model
Asset name	Predicting Life Expectancy - P3 ExtraTreesRegressorEstimator
Machine learning service	Machine Learning-60
Created	Jun 16, 2020 10:35 AM

TSK-50603-Build Node-Red Flow To Integrate Auto AI :



Inbox (4) - gkarun8766 x Student Dashboard x IISPS_INT_2913_Predict x IBM Watson Studio x Node-RED : node-red x Node-RED Dashboard x

node-red-ofeur.eu-gb.mybluemix.net/ui/#/0?socketid=yePj6Z7rtz8X_V0DAAAAD

☆ 🔍 🗨️ 🌐 🌍

Home

Predicting Life Expectancy

Predicted value : **65.74999961853027**

Country*
India

Year
2020

Status

Adult Mortality

infant deaths

Alcohol

percentage expenditure

Hepatitis B

Measles

BMI
18