

**PREDICTION OF HEALTH INSURANCE PREMIUM COST WITH
IBM AUTO AI SERVICE**

A MINI PROJECT REPORT

Submitted to

**JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY,
HYDERABAD**

In partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

BY

P. PREETHI (18UK1A0593)

MD. MAQSOOD ALI (18UK1A0587)

K. SADA PRIYAN (18UK1A05E7)

K. VENKATESH (18UK1A0581)

Under the esteemed guidance of

Mr. MAHIPALPAL REDDY

Assistant Professor



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VAAGDEVI ENGINEERING COLLEGE

Affiliated to JNTU, HYDERABAD

BOLLIKUNTA, WARANGAL-506005

2018-2022

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VAAGDEVI ENGINEERING COLLEGE

Affiliated to JNTU, HYDERABAD

BOLLIKUNTA, WARANGAL-506005.

2018-2022



CERTIFICATE

This is to certify that the Mini Project Report entitled “**PREDICTION OF HEALTH INSURANCE PREMIUM COST WITH IBM AUTO AI SERVICE**” is being submitted by **P.PREETHI (18UK1A0593), MD.MAQSOOD (18UK1A0587), K.SADA PRIYAN (18UK1A05E7), K.VENKATESH (18UK1A0581)**, in partial fulfillment of the requirement for the award of the Degree in Bachelor of Technology in computer science and engineering during the engineering during the academic year 2018-2022.

Mr. MAHIPAL REDDY

Project Guide

Dr. R. NAVEEN KUMAR

Head of The Department

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

We wish to take this opportunity to express our sincere gratitude and deep sense of respect to our beloved **Dr. P. Prasad Rao**, Principal, Vaagdevi Engineering College for making us available all the required assistance and for his support and inspiration to carry out this mini project in the institute.

We extend our heartfelt thanks to **Dr. R. Naveen Kumar**, Head of the Department of CSE, Vaagdevi Engineering College for providing us necessary infrastructure and thereby giving us freedom to carry out the mini project.

We express heartfelt thanks to the Mini Project Coordinator, **Ms. G. Arunakranthi**, Assistant Professor, Department of CSE for her constant support and giving necessary guidance for completion of this mini project.

We express heartfelt thanks to the guide, **Mr. Mahipal Reddy**, Assistant Professor, Department of CSE for her constant support and giving necessary guidance for completion of this mini project.

We express our sincere thanks and gratitude to **The Smart Bridge**, for providing internship.

Finally, we express our sincere thanks and gratitude to our family members, friends for their encouragement and outpouring their knowledge and experiencing throughout thesis.

ABSTRACT

This project aims at building a web App that automatically estimates premium cost by taking the input values. Using IBM AutoAI, we automate all of the tasks involved in building predictive models for different requirements. You create a model from a data set that includes the age, gender, BMI, number of children, smoking preferences, region, and charges to predict the health insurance premium cost that an individual pays.

CONTENTS

LIST OF CHAPTERS	PAGE NO
1.INTRODUCTION.....	6
1.1 INTRODUCTION	
1.2 OVERVIEW	
1.3 PURPOSE	
2.PROJECT OBJECTIVES.....	7
3.LITERATURE SURVEY	8
3.1EXISTING PROBLEM	
3.2PROPOSED SOLUTION	
4.THEORITICAL ANALYSIS.....	9
4.1BLOCK DIAGRAM	
4.2HARDWARE / SOFTWARE DESIGNING	
5.EXPERIMENTAL INVESTIGATIONS.....	10
5.1METHODS AND IMPLEMENTATION	
6.APPLICATIONS.....	12
7.CONCLUSION.....	18
8.BIBILOGRAPHY.....	18

INTRODUCTION

1.1 INTRODUCTION

As we see the value of gross insurance premiums worldwide continue to skyrocket past 5 trillion dollars, we know that most of these costs are preventable. For example, just by eliminating smoking and lowering your BMI by a few points might mean shaving thousands of dollars from your premium charges. In this application, we study the effects of age, smoking, BMI, gender, and region to determine how much of a difference these factors can make on your insurance premium. By using our application, customers see the radical difference their lifestyle choices make on their insurance charges. By leveraging artificial intelligence (AI) and machine learning, we help customers understand just how much smoking increases their premium by predicting how much they will have to pay within seconds.

1.2 OVERVIEW

Using IBM AutoAI, we automate all of the tasks involved in building predictive models for different requirements. You see how AutoAI generates great models quickly, which saves time and effort, and aids in a faster decision-making process. You create a model from a data set that includes the age, gender, BMI, number of children, smoking preferences, region, and charges to predict the health insurance premium cost that an individual pays.

1.3 PURPOSE

This project aims at building a web App that automatically estimates premium cost by taking the input values. Using IBM AutoAI, we automate all of the tasks involved in building predictive models for different requirements. You create a model from a data set that includes the age, gender, BMI, number of children, smoking preferences, region, and charges to predict the health insurance premium cost that an individual pays.

PROJECT OBJECTIVES

After completing this project, you will learn how to

- Work with Watson Studio
- Create a project in Watson Studio
- Use Auto Ai experiment to create a model
- Deploy the ML model as a webserver
- Integrating Model and Node-RED Service
- Build an Application using Node-RED which takes inputs from the user and showcases the prediction on UI

LITERATURE SURVEY

EXISTING PROBLEM

The health insurance industry has always been a slow-moving industry when it comes to adopting the data analytics practices into its business models. With the advent of advanced data analytics technologies, it has become important more than ever to take advantage of such sophisticated analytics to accurately assess and predict the insurance premiums for the insured.

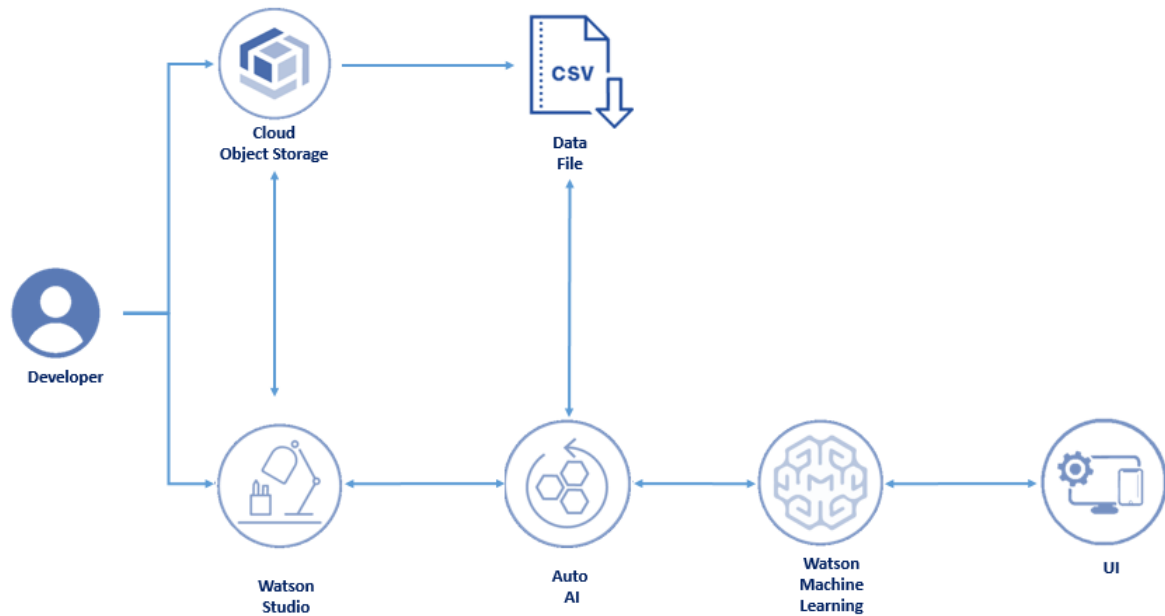
Thus, one of the important tasks for health insurance companies is to determine the policy premiums. By using predictive modelling, the insurers can determine the policy premium for the insured based on their behaviors which are indicated by attributes such as age, BMI (Body Mass Index), smoking habits, number of children etcetera.

PROPOSED SOLUTION

Using IBM AutoAI, we automate all of the tasks involved in building predictive models for different requirements. You see how AutoAI generates great models quickly, which saves time and effort, and aids in a faster decision-making process. You create a model from a data set that includes the age, gender, BMI, number of children, smoking preferences, region, and charges to predict the health insurance premium cost that an individual pays.

THEORITICAL ANALYSIS

BLOCK DIAGRAM



HARDWARE / SOFTWARE DESIGNING

We use only software designing in this project there are

- IBM Cloud
- IBM Watson studio
- Auto AI Experiment
- Model Deployment
- IAM Access Token
- Node red service

EXPERIMENTAL INVESTIGATIONS

PROJECT FLOW

- Log in to IBM account
- Create IBM Watson Studio and Node-RED Service
- Create a Watson studio project
- ADD Auto AI Experiment
- Run the Auto AI Experiment to build a Machine learning model on the desired dataset
- Save the model
- Deploy the model as a web server and generate scoring End Point
- Create a WEB application Using Node-RED to take user input and showcase Prediction on UI

PREREQUISITES

To complete this project you must have an IBM account

- Please register for IBM
- Please log in to IBM Account

Prior Knowledge

you should have knowledge of the following Concepts

- Watson Studio

Create IBM Services

In this activity, you will be creating the Necessary IBM services. The following are the services that you have to create.

- Watson studio
- Node-RED
- Cloud Object Storage service (COS)
- Machine Learning service (ML)

COS and ML services will e created while creating a Watson Studio Project

Build Machine Learning Model

In this activity, you will learn how you can save a pipeline as a Watson Machine Learning model, deploy the model, and score it to view a prediction.

This Activity contains the Following Tasks

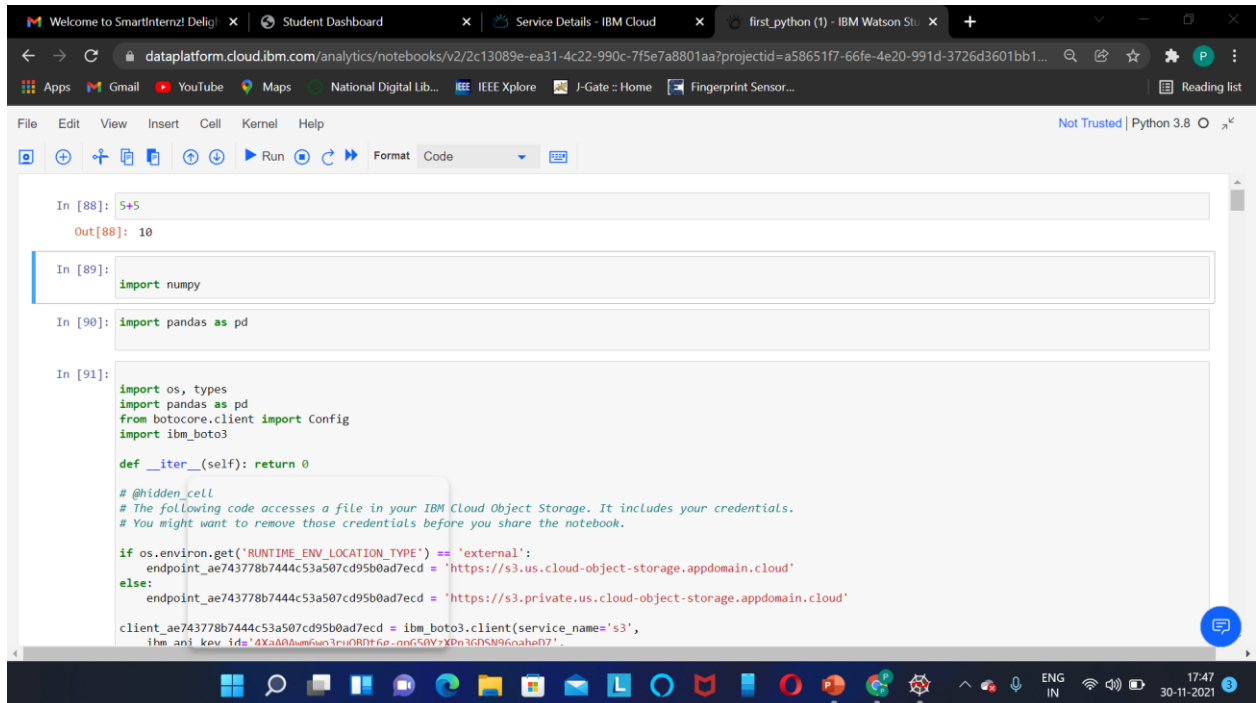
- Collect the data set
- Create Watson Studio project
- Add Auto AI experiment
- Run AI Experiment
- Save the model
- Deploy the model

COLLECT THE DATASET

In this Activity, We gonna build a machine learning model that predicts insurance-premium-prediction based on the following parameters

- Age
- BMI
- Children and expenses

- Sex
- Smoker and
- Region



APPLICATIONS

CREATE WATSON STUDIO PROJECT

A project is how you organize your resources to achieve a particular goal.

ADD AUTO AI EXPERIMENT

The AutoAI graphical tool in Watson Studio automatically analyzes your data and generates candidate model pipelines customized for your predictive modeling problem. These model pipelines are created iteratively as AutoAI analyzes your dataset and discovers data transformations, algorithms, and parameter settings that work best for your problem setting. Results are displayed on a leaderboard, showing the automatically generated model pipelines ranked according to your problem optimization objective.

```

ibm_api_key_id='4XaA0Awm6wo3ru0B0T6g-qpG50YzXPn3GDSN96oaheD7',
ibm_auth_endpoint='https://iam.cloud.ibm.com/oidc/token',
config=Config(signature_version='oauth'),
endpoint_url=endpoint_ae743778b7444c53a507cd95b0ad7ecd)

body = client_ae743778b7444c53a507cd95b0ad7ecd.get_object(Bucket='health1-donotdelete-pr-bzdopbktc3euxx',Key='insurance-Copy1.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[91]:

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520

```

In [92]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn import metrics
body = client_ae743778b7444c53a507cd95b0ad7ecd.get_object(Bucket='health1-donotdelete-pr-bzdopbktc3euxx',Key='insurance-Copy1.csv')['Body']

```

```

In [95]: df_data_2.shape
Out[95]: (1338, 7)

In [96]: df_data_2.head()
Out[96]:

```

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520

Out[94]:

	age	bmi	children	charges
count	1338.000000	1338.000000	1338.000000	1338.000000
mean	39.207025	30.663397	1.094918	13270.422265
std	14.049960	6.098187	1.205493	12110.011237
min	18.000000	15.960000	0.000000	1121.873900
25%	27.000000	26.296250	0.000000	4740.287150
50%	39.000000	30.400000	1.000000	9382.033000
75%	51.000000	34.693750	2.000000	16639.912515
max	64.000000	53.130000	5.000000	63770.428010

To add the project

- Click on Add Project from Watson Studio project
- Select Auto AI
- Specify a name and description for your experiment.
- Select a machine learning service instance and click Create.

```
Welcome to SmartInternz! Deligi | Student Dashboard | Service Details - IBM Cloud | first_python (1) - IBM Watson Stu | +
dataplatfom.cloud.ibm.com/analytics/notebooks/v2/3164d01e-e35d-46d7-8282-6153e231415e?projectid=a58651f7-66fe-4e20-991d-3726d3601bb...
File Edit View Insert Cell Kernel Help Not Trusted | Python 3.8
Run Format Code

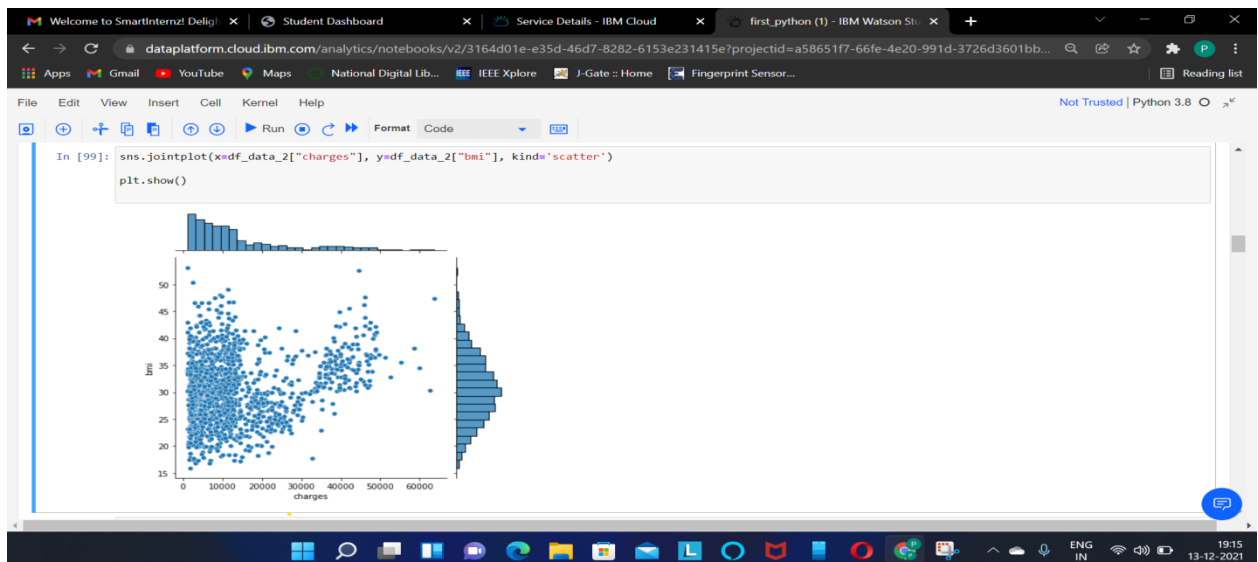
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1338 entries, 0 to 1337
Data columns (total 7 columns):
# Column Non-Null Count Dtype
---
0 age 1338 non-null int64
1 sex 1338 non-null object
2 bmi 1338 non-null float64
3 children 1338 non-null int64
4 smoker 1338 non-null object
5 region 1338 non-null object
6 charges 1338 non-null float64
dtypes: float64(2), int64(2), object(3)
memory usage: 73.3+ KB

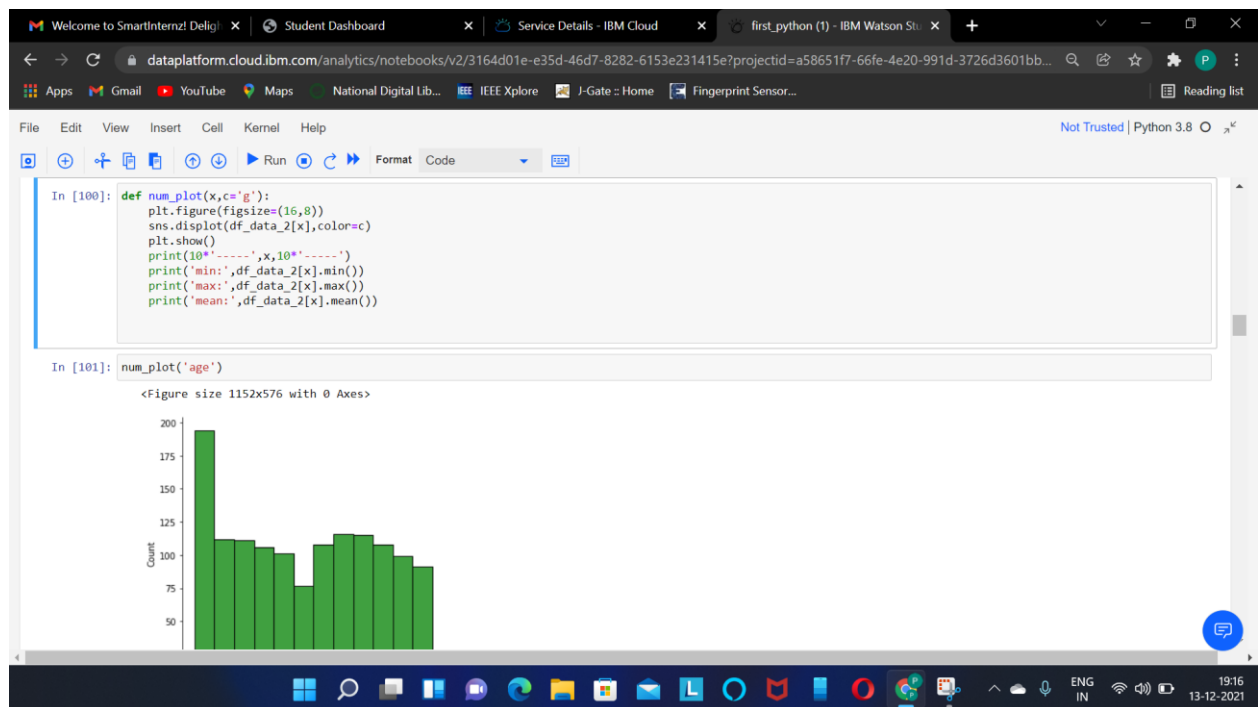
In [98]: df_data_2.dtypes
Out[98]: age          int64
sex            object
bmi            float64
children       int64
smoker         object
region         object
charges        float64
dtype: object

In [99]: sns.jointplot(x=df_data_2["charges"], y=df_data_2["bmi"], kind='scatter')
plt.show()
```

RUN AUTO AI EXPERIMENT

Run the auto ai experiment and save the model



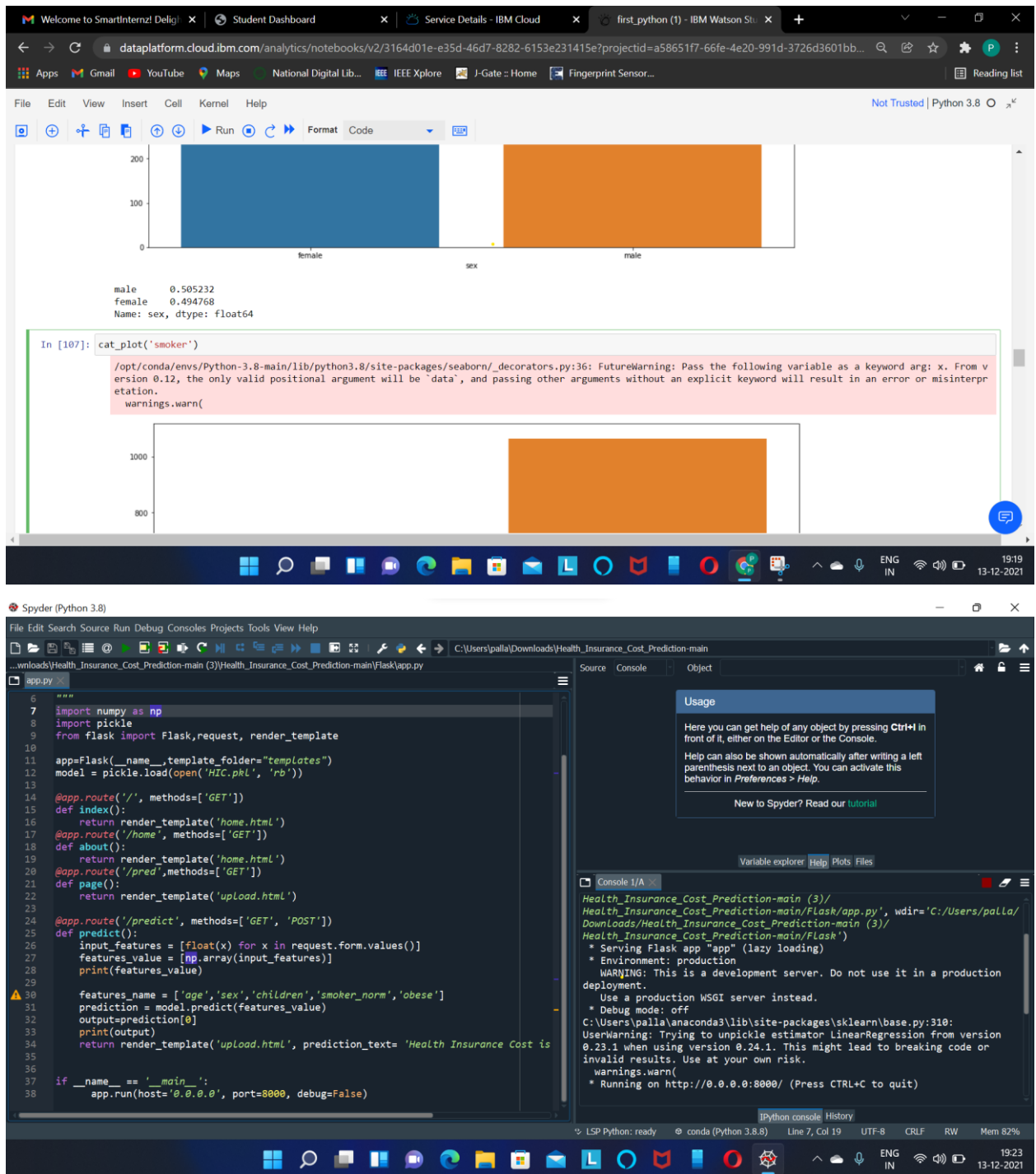


SAVE THE MODEL

Once the pipeline creation is complete, you can view and compare the ranked pipelines in a leaderboard. Choose Save model from the action menu for the pipeline with the highest accuracy or low error rate This saves the pipeline as a Machine Learning asset in your project. A notification gives you the link to view the saved model in your project.

DEPLOY THE MODEL

Before you can use your trained model to make predictions on new data, you must deploy the model.



You can deploy the model from the model details page. You can access the model details page in one of these ways:

- Click on the model name in the notification displayed when you save the model.

- Open the Assets page for the project containing the model and click the model name in the Machine Learning Model section.

From the model details page:

- Click the Promote to deployment space.
- Choose an existing deployment space or create a new one.
- Click Add Deployment.
- In the page that opens, fill in the fields:
 - Specify a name for the deployment.
 - Select “Web service” as the Deployment type.
 - Click Save.

After you save the deployment, click on the deployment name to view the deployment details page.

BUILD NODE RED APPLICATION

Let's build a User interface which takes inputs from the user. The Model Analyses the Inputs and returns the Prediction that is showcased on the User interface

INTEGRATE NODE –RED WITH AUTO AI MODEL

In this task, you will be creating a Node-RED flow that integrated with the model built

CONCLUSION

Using IBM Auto AI, we automate all of the tasks involved in building predictive models for different requirements. You see how Auto AI generates great models quickly, which saves time and effort, and aids in a faster decision-making process. As we see the value of gross insurance premiums worldwide continue to skyrocket past 5 trillion dollars, we know that most of these costs are preventable. For example, just by eliminating smoking and lowering your BMI by a few points might mean shaving thousands of dollars from your premium charges. In this application, we study the effects of age, smoking, BMI, gender, and region to determine how much of a difference these factors can make on your insurance premium. By using our application, customers see the radical difference their lifestyle choices make on their insurance charges. By leveraging artificial intelligence (AI) and machine learning, we help customers understand just how much smoking increases their premium by predicting how much they will have to pay within seconds.

BIBLIOGRAPHY

- <https://developer.ibm.com/patterns/create-an-application-to-predict-your-insurance-premium-cost-with-autoai/>
- <https://github.com/IBM/predict-insurance-charges-with-autoai>
- <https://github.com/SmartPracticeschool/SPS-7583-Health-Insurance-premium-prediction-using-IBM-Auto-AI-Service>
- <https://smartinternz.com/tunisia-project/144>
- <https://www.the-digital-insurer.com/aggregator/predict-your-health-insurance-premium-charges-with-autoai-ibm/>
- <https://www.altexsoft.com/blog/datascience/comparing-machine-learning-as-a-service-amazon-microsoft-azure-google-cloud-ai-ibm-watson/>

