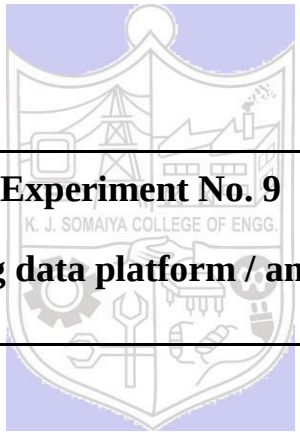


Experiment No. 9

Title: Case study: Big data platform / analytics as business need)

Batch: A1 Roll No.: 1714011 **Experiment No.:9**

Title: Case study: Income Prediction

Resources needed: Internet Connection, Microsoft Azure Account

---

Describe the following points with respect to the business under consideration,

1. **Problem faced by the business**

Answer: The system is so designed that it helps to estimate the salary the employee deserves, so it helps the business to give the correct amount of salary to their employees, so that the employee does not resign because of high salary expectations nor the company will face the issue of giving more salary then required, and thus saving the total expenditure.

2. **Approach/ Methodology followed by the business**

Answer: The methodology which we followed was that we took the dataset of incomes. Then we split the dataset into train and test. Then we trained the machine learning model with that dataset and in the end we tested that model. Then we created a web service to use predictive analysis to predict the salary.

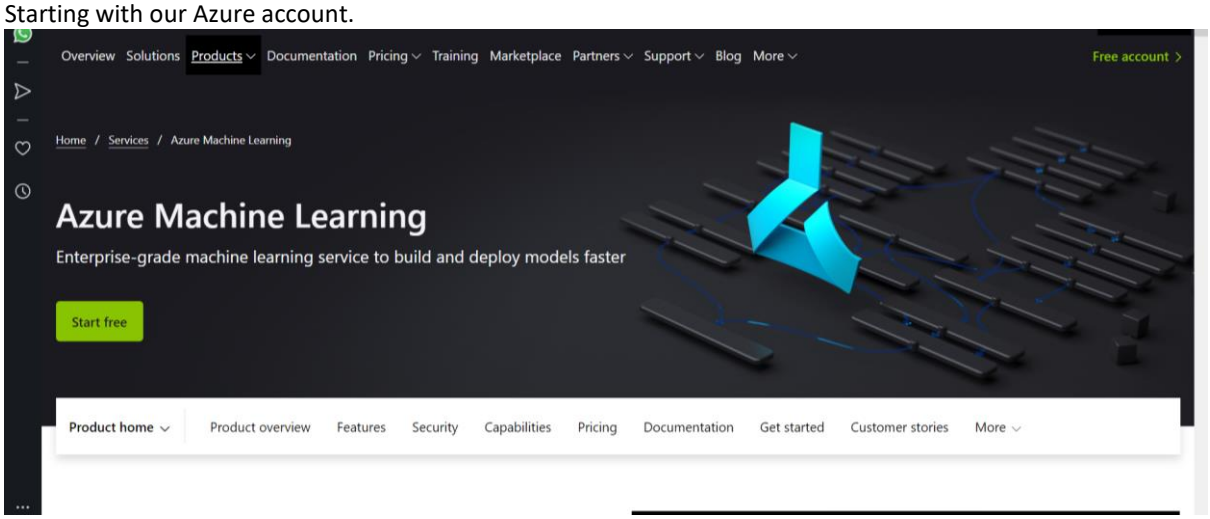
3. **Skillsets, infrastructure and other impact on the business during implementation**

Answer: Skillset: Cloud computing
Infrastructure: Microsoft Azure Cloud services
There were no as such impacts on the business, as any of the business service lines were not being used.

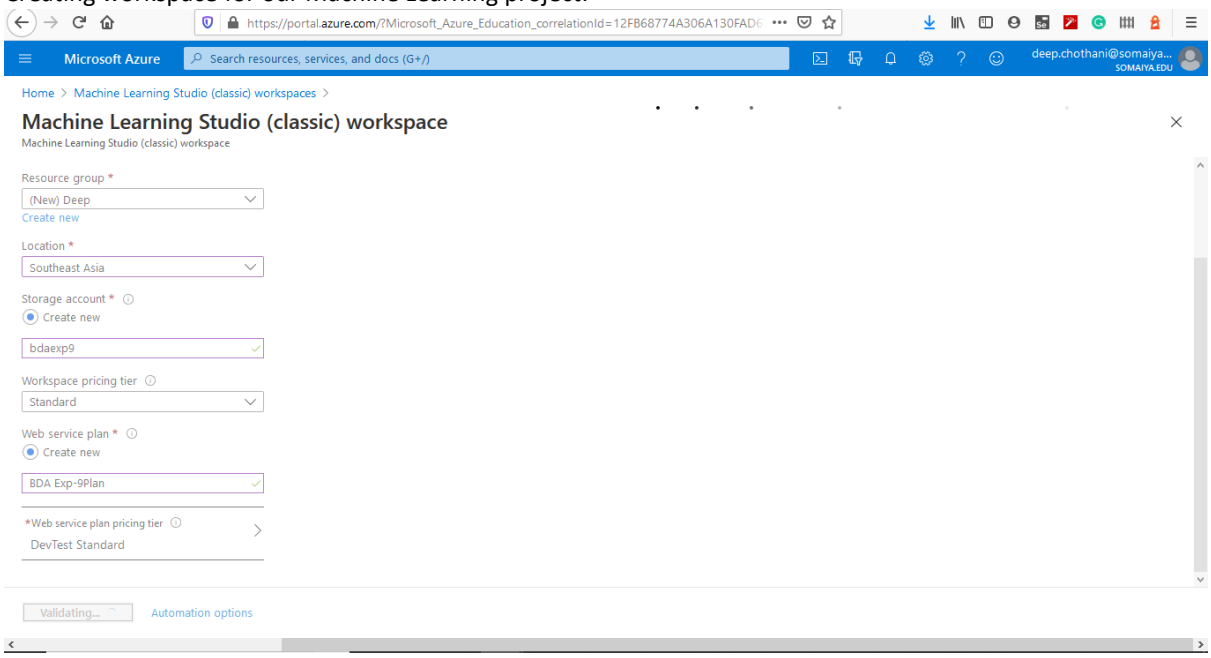
4. **Similar approaches followed by other businesses**

Answer: Other businesses is still not using ML in order to estimate their employee salaries, and thus face problems which we were facing initially like employees dis-satisfaction or company's over expenditure by giving more salary then required.

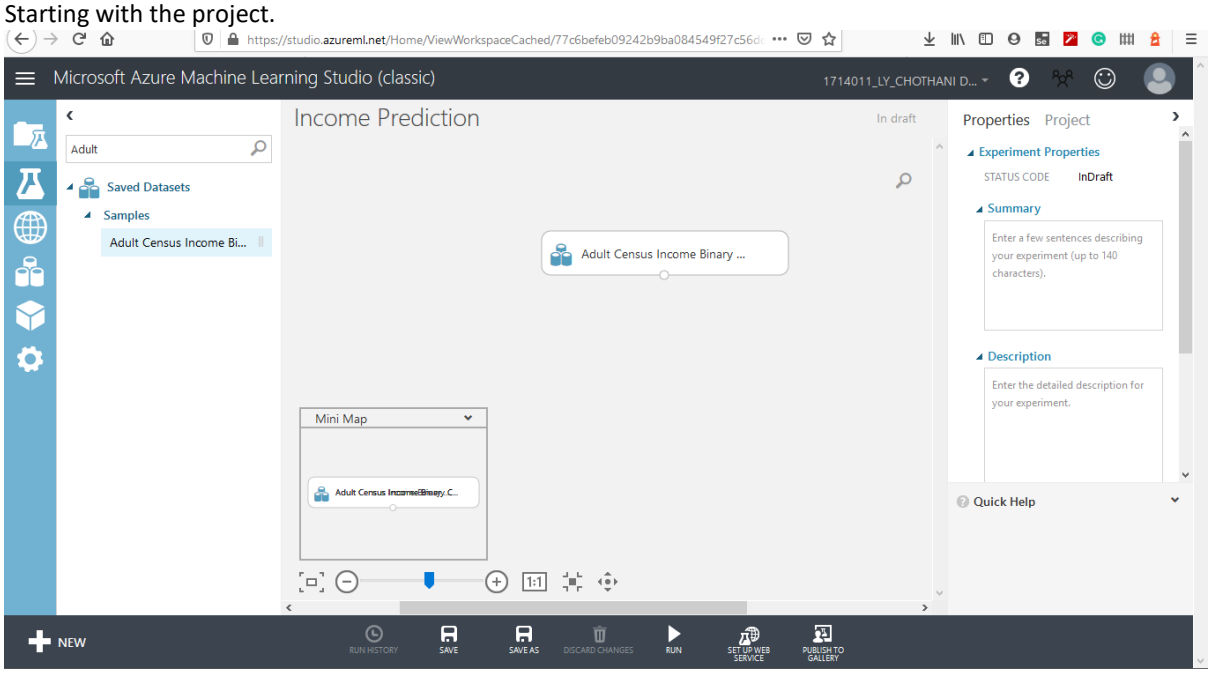
Starting with our Azure account.



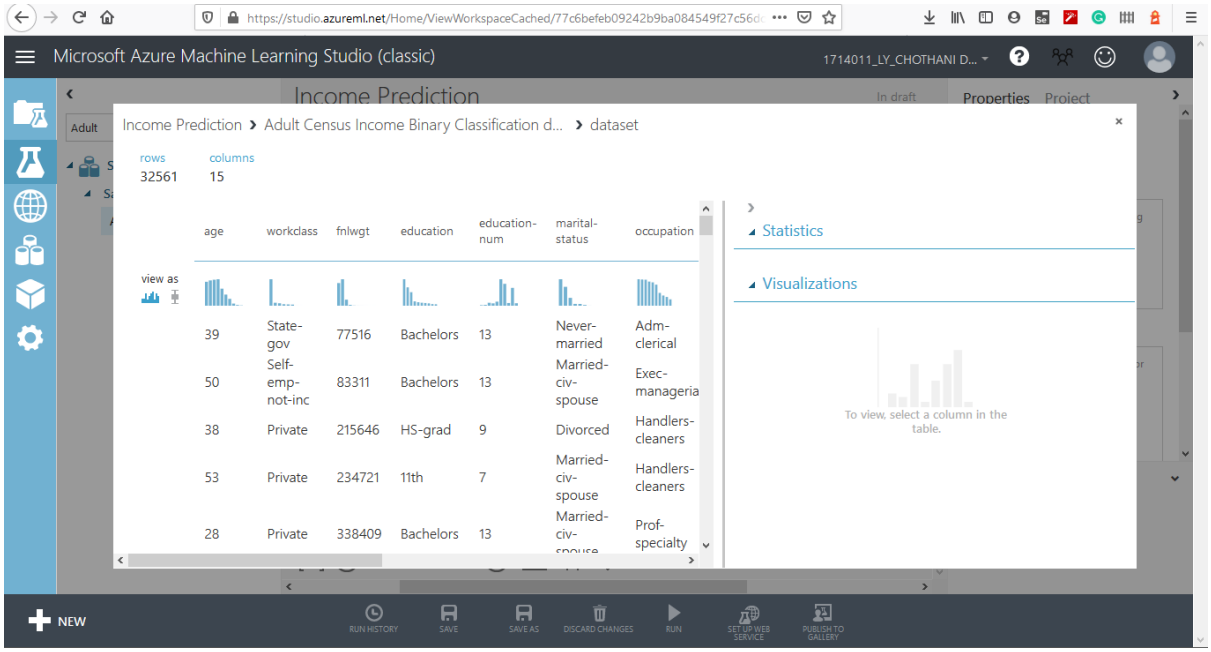
Creating workspace for our Machine Learning project.



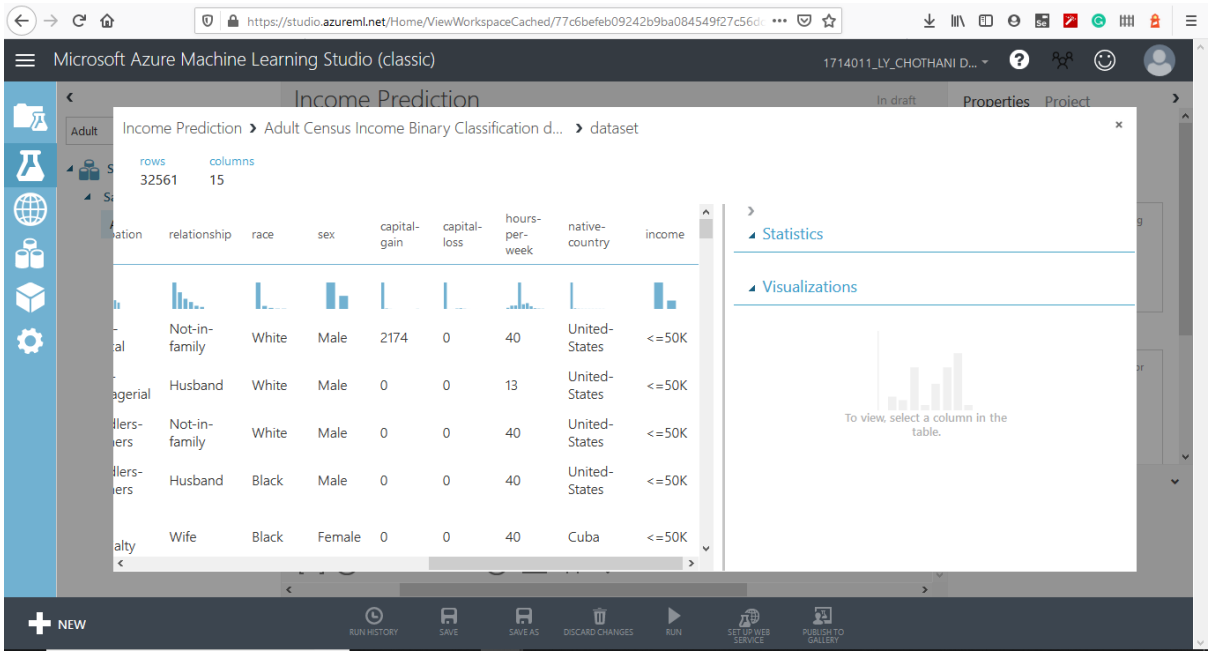
Starting with the project.



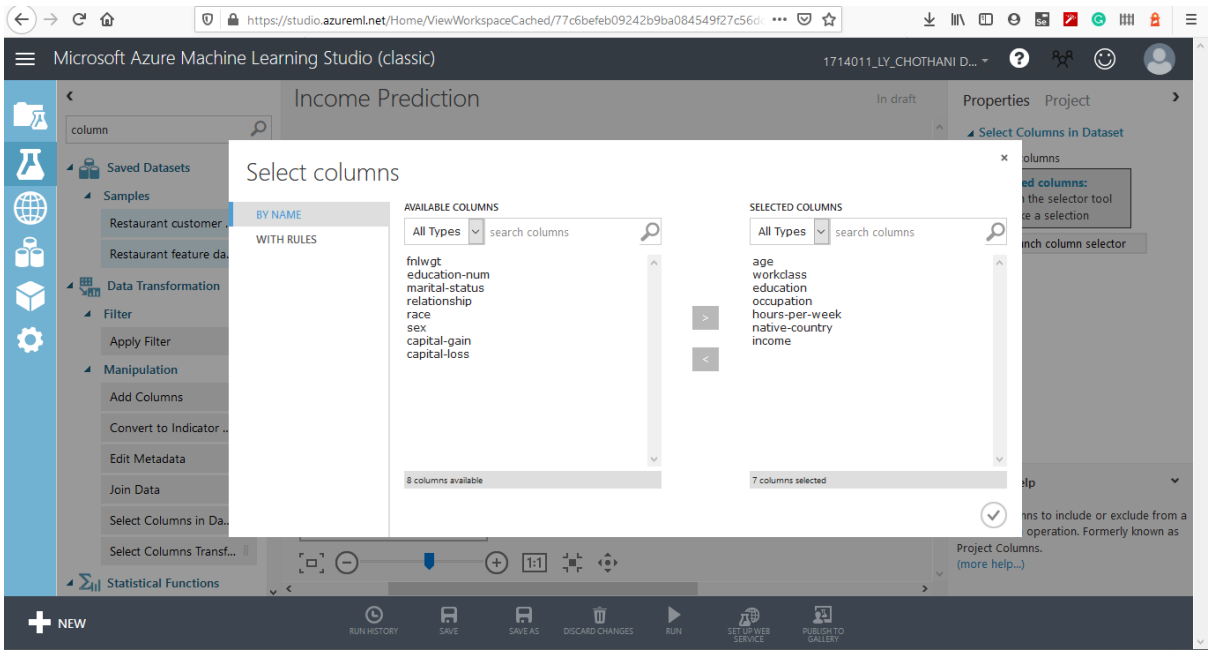
Selecting the dataset Adult Census Income- which has the data of income of employees with different inputs.



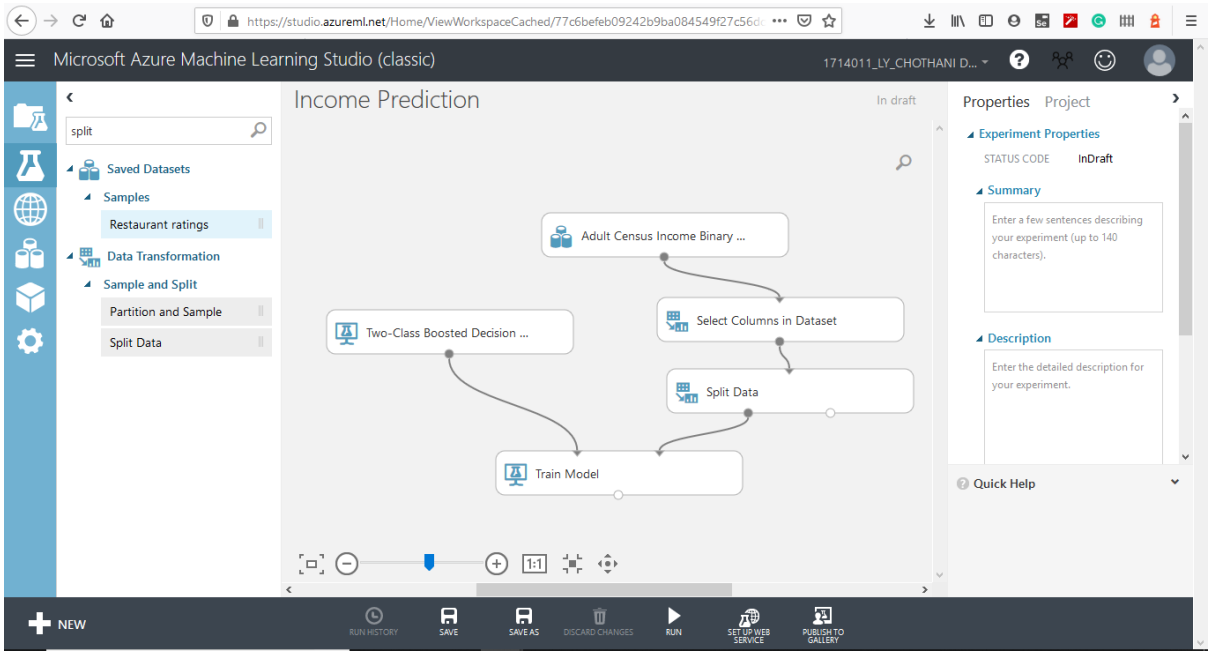
Visualizing the data.



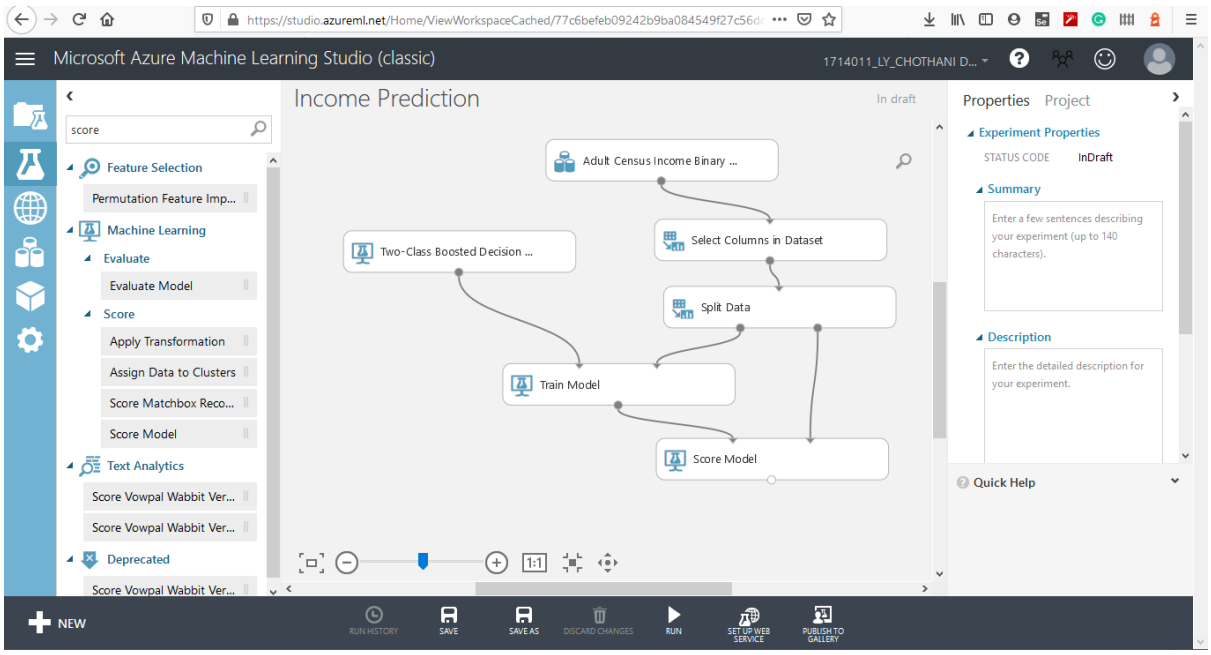
It shows all the columns of the dataset.

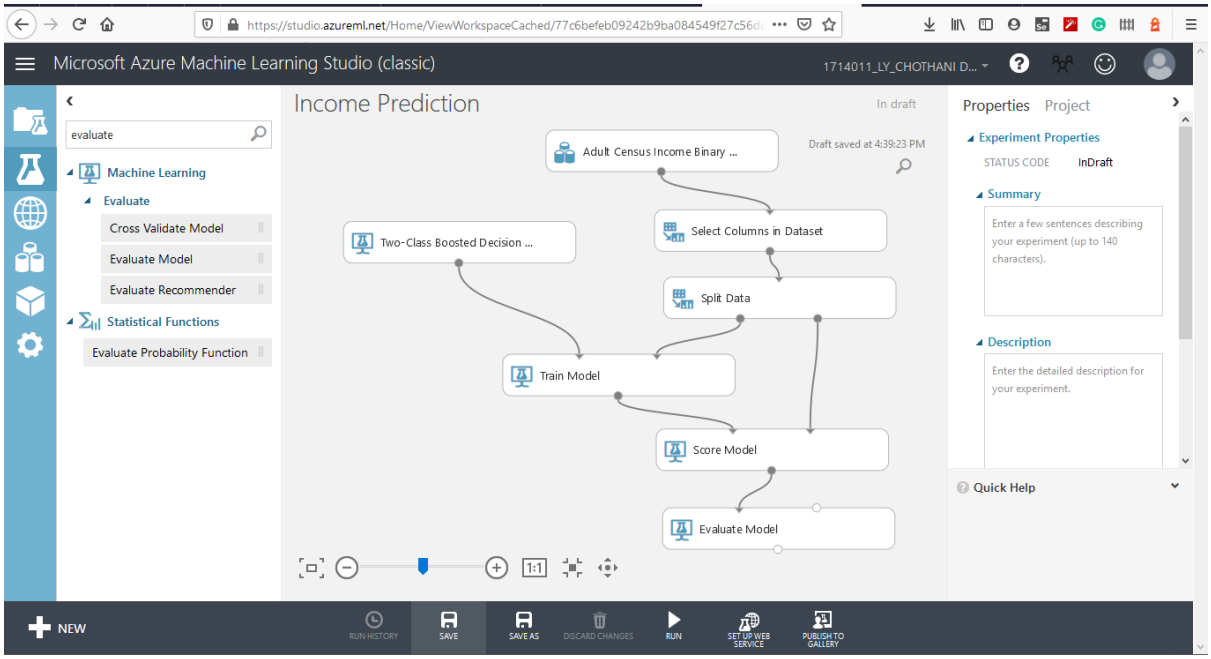


Selecting the columns for training the model.

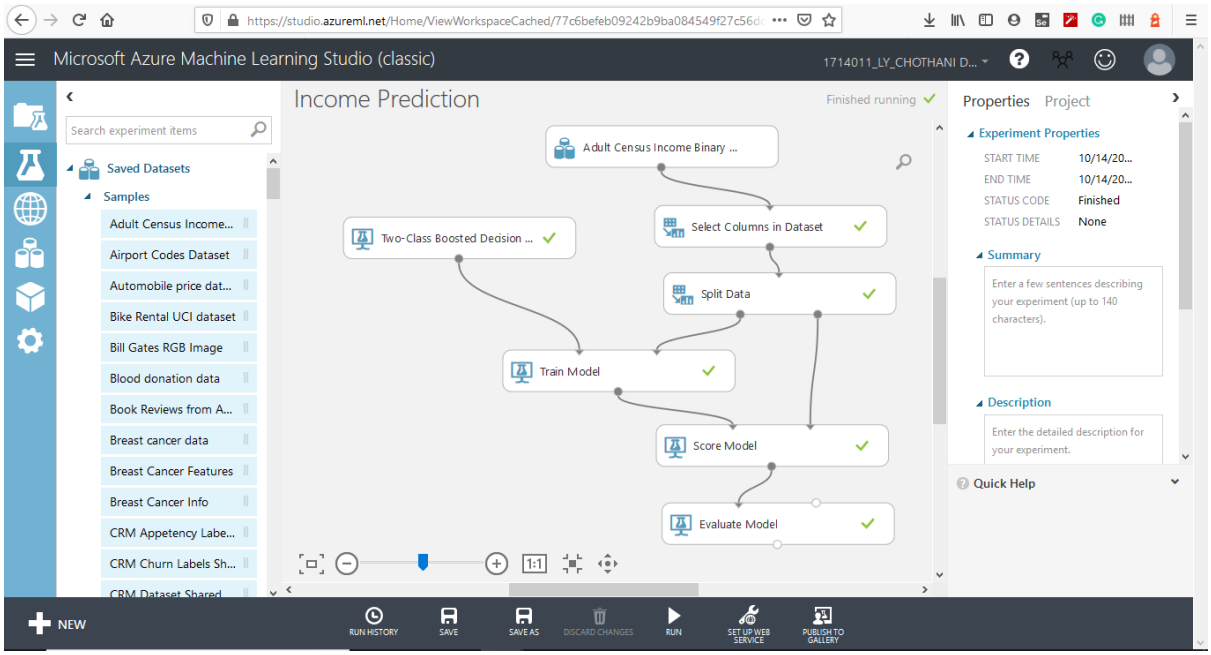


This shows the flowchart of working.

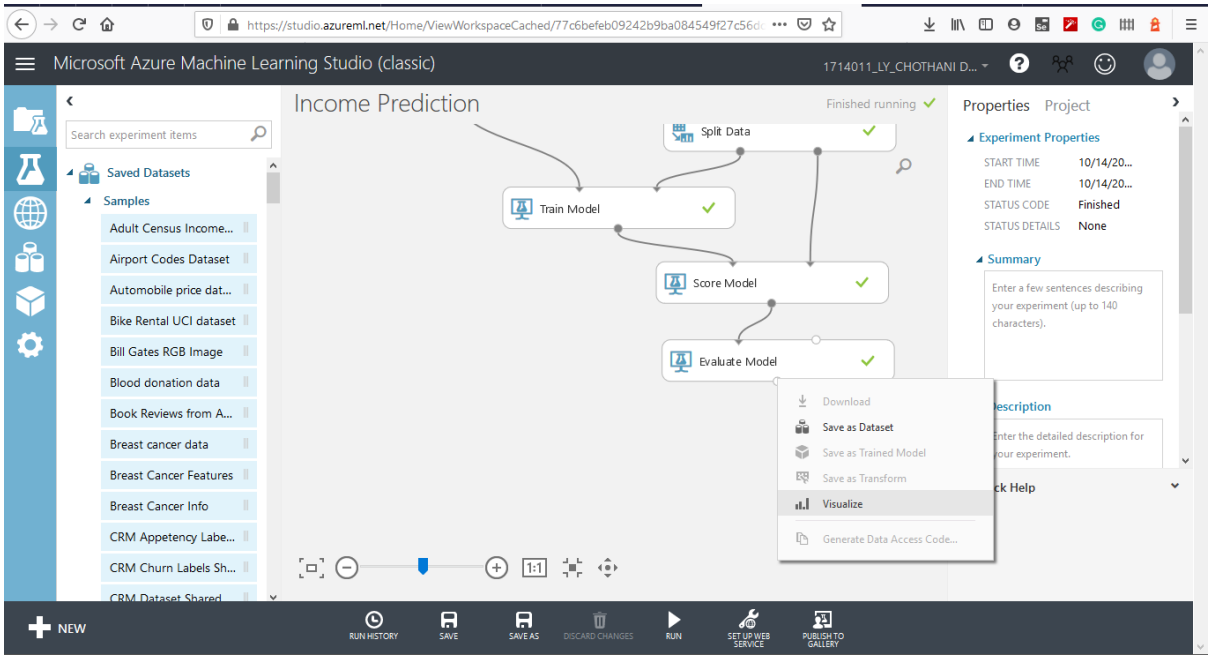




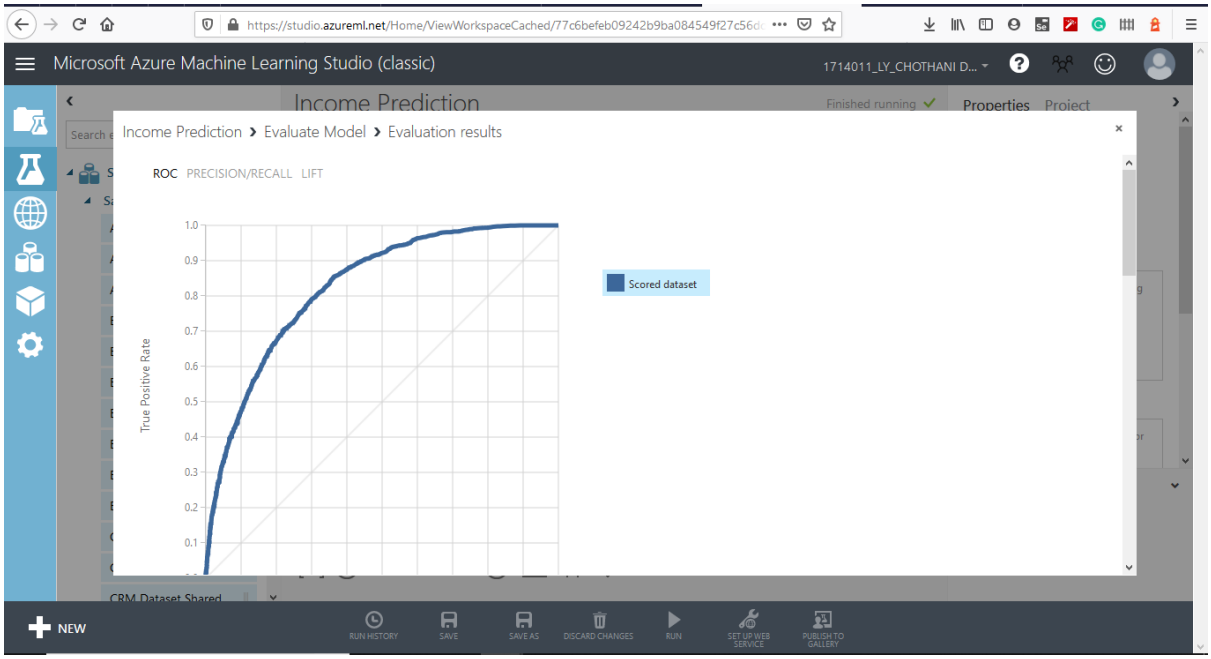
Finally, adding the model for evaluation.

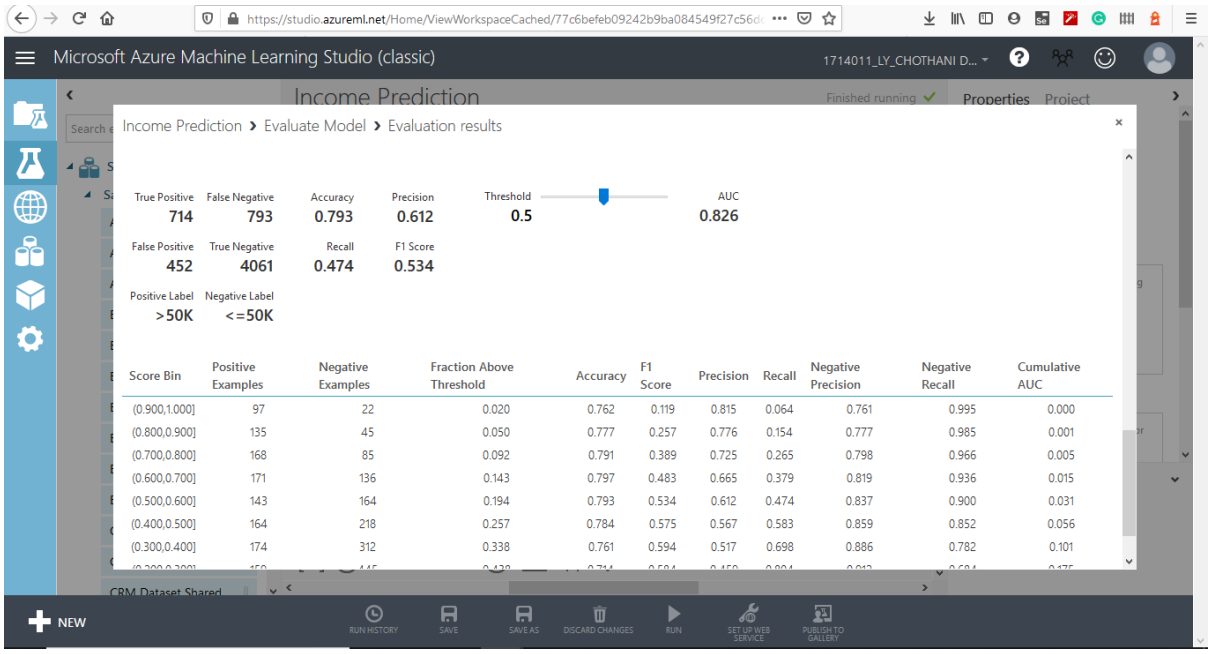


After running, we can see all the parts ran successfully.

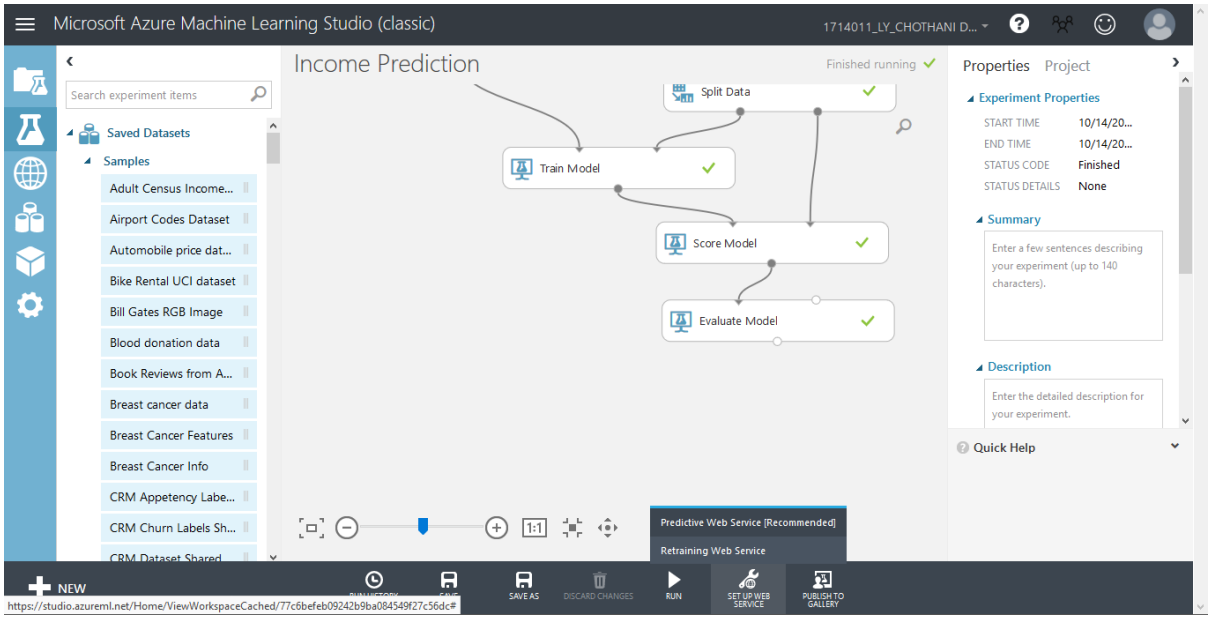


Now, time to visualize our data in order to see the precision and how the model is trained.

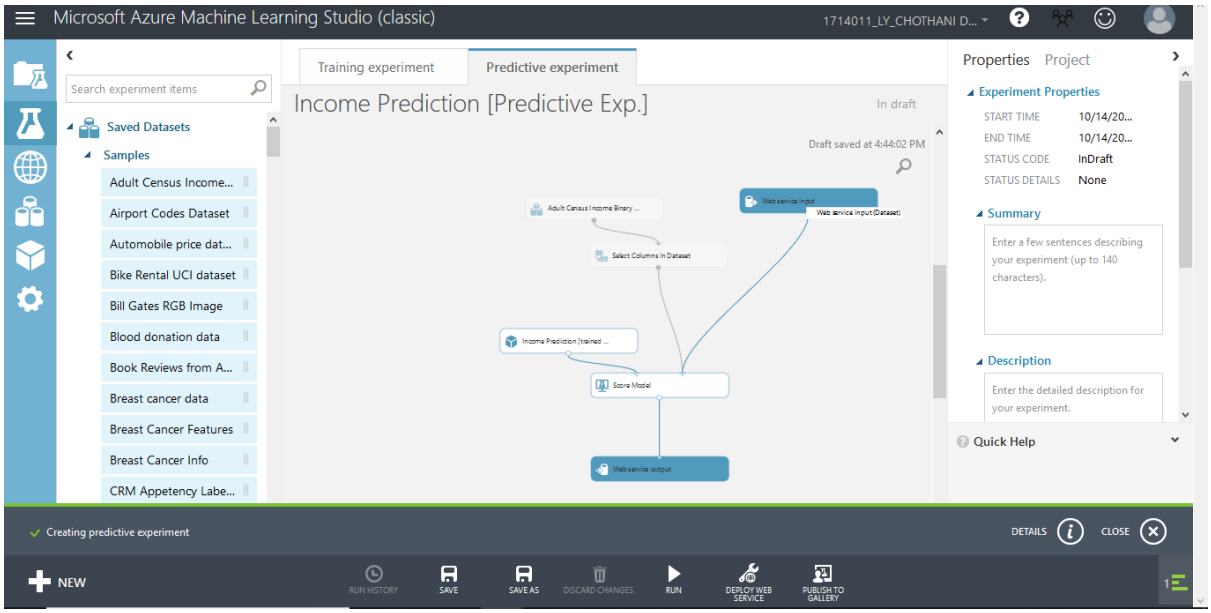




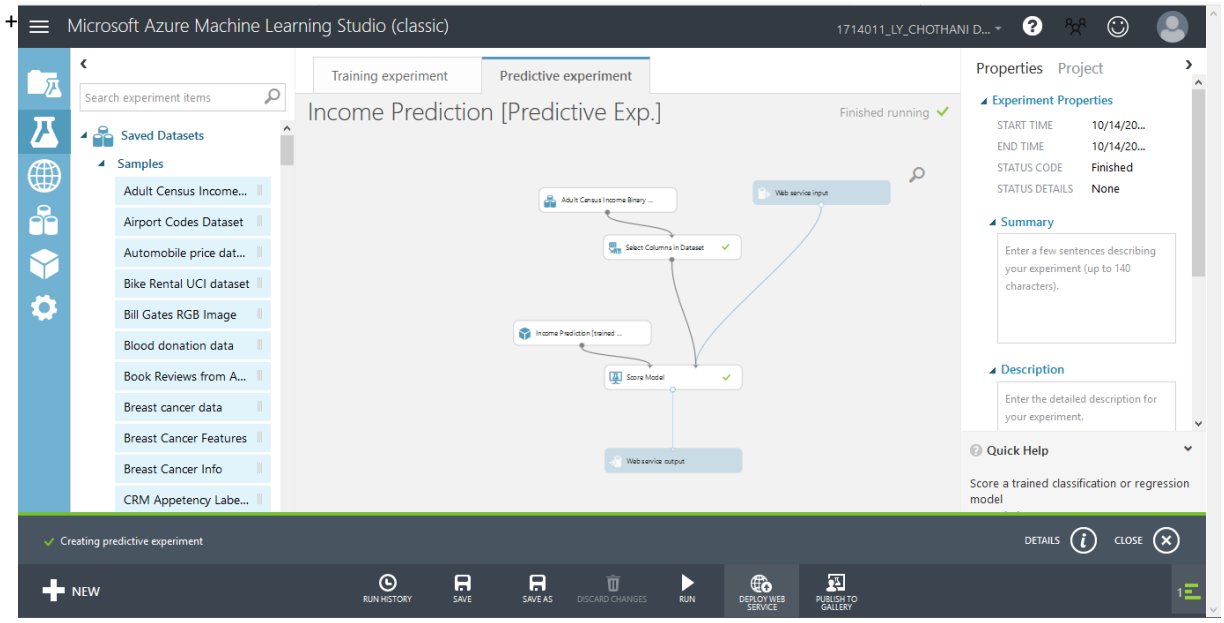
We can see that our model gives the accuracy of 79.3%



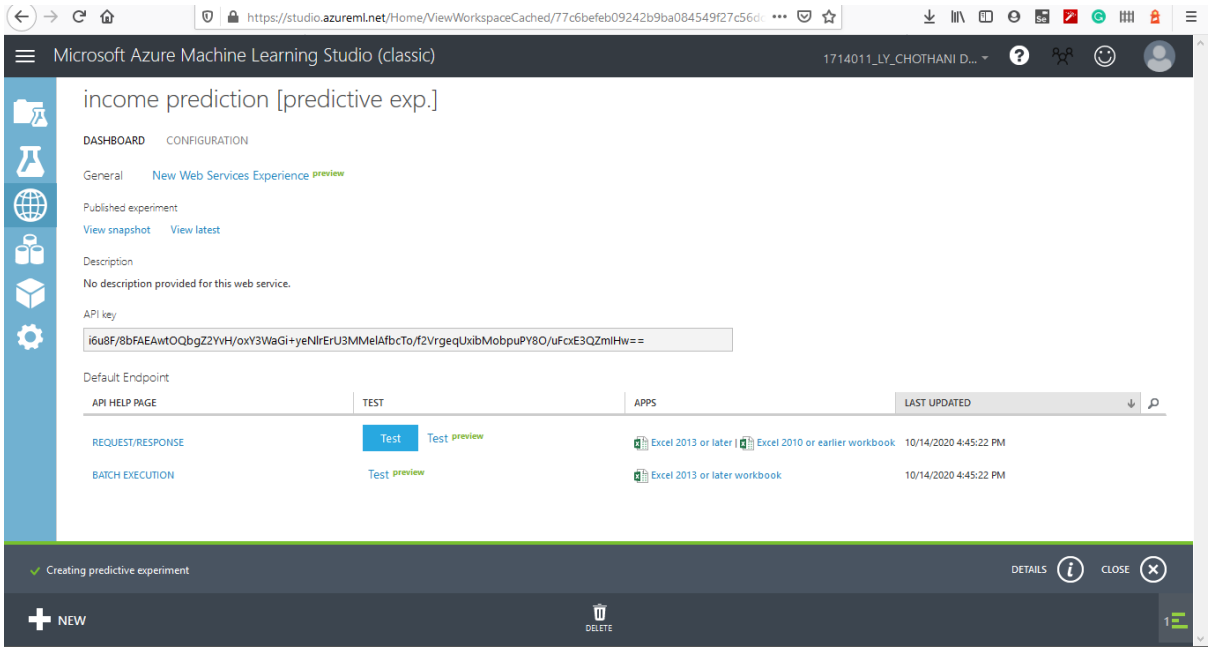
Now, making it available as a web service.



So, this will be the final flowchart where web-based input & output is also present.

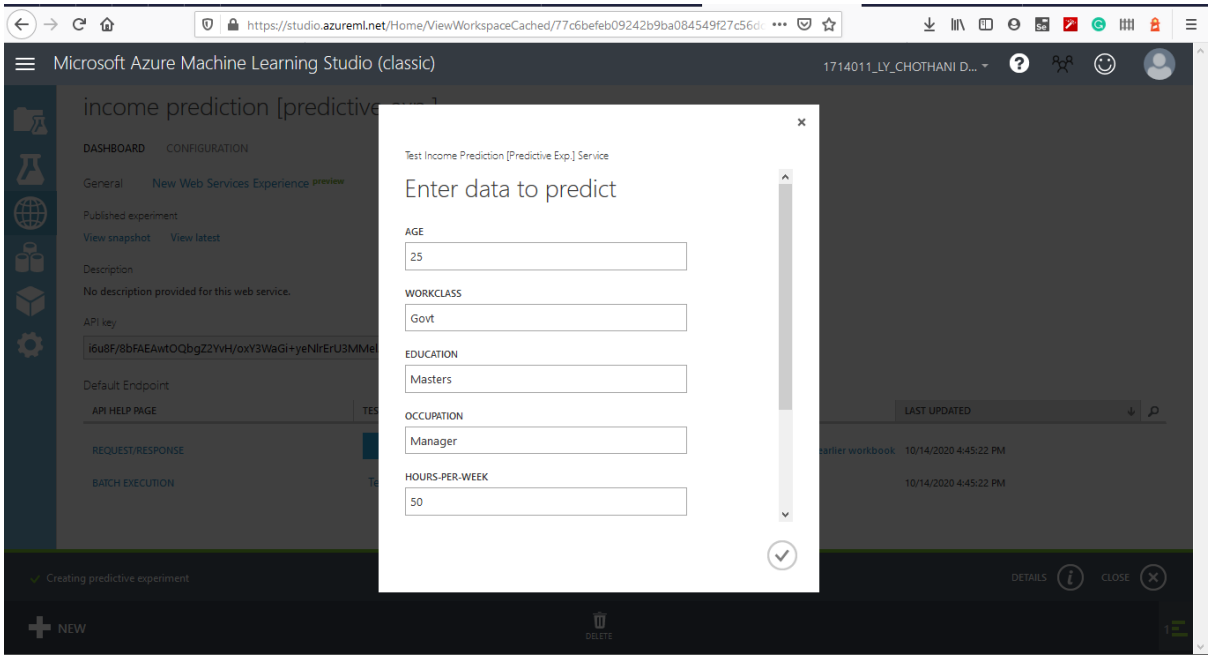


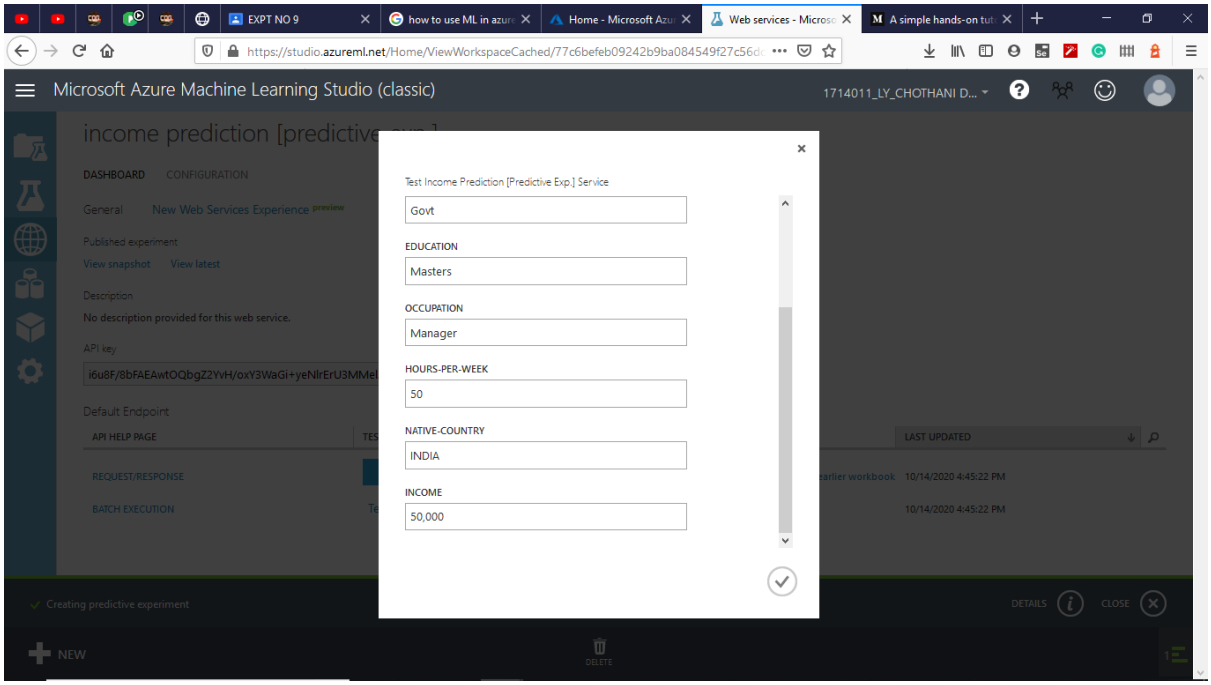
Deploy web service.



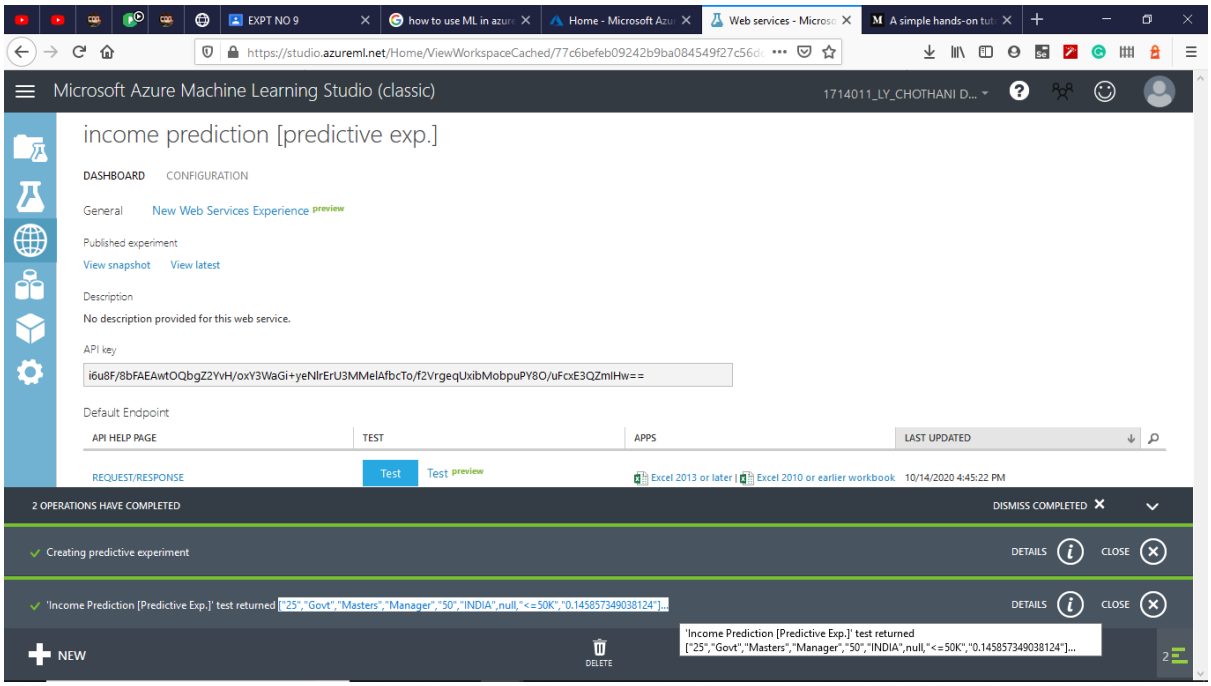
And we can see, it is successfully deployed.

Time to test:





Entering all the inputs required.



And we can see the output i.e. the estimated income of that particular employee.

Questions:

Discuss the tangible and intangible benefits the business has observed after the implementation.
Answer:

- The deserving got the right salary.
- Manpower got saved due to automation.
- This model can be used now anywhere in the enterprise.
- Lower expenditure.

Outcomes: Realize adequate perspectives of big data analytics in various applications.

Conclusion: I was able to train a model which predicts the employees salary based on their age, experience, and qualifications using Azure ML Studio and then was able to publish it on GitHub: https://github.com/deepc594/INCOME-PREDICTION-AZURE-ML

Grade: AA / AB / BB / BC / CC / CD /DD

Signature of faculty in-charge with date

References:

Books/ Journals/ Websites: