**OTT IMDB Rating prediction using IBM Watson studio**

**Submitted by**

**C.H Nikhil Teja**

**Rage Sai Kiran**

|  |  |  |
| --- | --- | --- |
| Index | NAME | PAGE NO |
| 1 | INTRODUCTION | 3 |
| 1.1 | Overview | 3 |
| 1.2 | Purpose | 3 |
| 2 | LITERATURE SURVEY | 3 |
| 2.1 | Existing problem | 3 |
| 2.2 | Proposed solution | 4 |
| 3 | THEORITICAL ANALYSIS | 5 |
| 3.1 | Block diagram | 5 |
| 3.2 | Hardware / Software designing | - |
| 4 | EXPERIMENTAL INVESTIGATIONS | 6 |
| 5 | FLOWCHART | 9 |
| 6 | RESULT | 10 |
| 7 | ADVANTAGES & DISADVANTAGES | 11 |
| 8 | APPLICATIONS | 11 |
| 9 | CONCLUSION | 12 |
| 10 | FUTURE SCOPE | 12 |

**INTRODUCTION**

**1.1 Overview**

As we all know that many people like watch movies in theater but due to this pandemic people were sitting at home not able to go to the theater. Days passed months passed then the OTT came to its place for producers it became GOD. After the movie released the only thing people care is about its rating ,Even if it is big hero’s movie also they need to think to watch the movie or not. After seeing rating from websites they will watch movie. Slowly many recent releases were released in those platforms. People changed from watching trail subscription to the paid subscription. There are many more movies are available in OTT. Not only movies many series were also available. AS per our available data set which was provided by the Smart internz platform we choose one data set and applied and with the help of the IBM Cloud software and services we choosed Watson studio. With the help of this Watson studio, we trained the dataset and deployed it and finally we were able to predict the IMDB rating of the particular movie. After predicting the rating with the help of NODE RED (which were provided by the IBM Cloud) we tried to create the interface.

**1.2 Purpose**

As we all know that the demand of OTT is increased very high. Many people were watching the OTT for movies and series. But many people don’t know what to watch so the first thing they get the mind is rating of the movie. It is like in you tube we all know that there are many videos but mostly we only watch which has more views and which has more likes. Based on that only we watch some particular video like that only people need to know which one need to watch after movie or series got released in the OTT. Not only that the movie success also depends on it because if some movie has more rating means many people watch that movie in the OTT platform. Many lives were depended on it.

**2.LITERATURE SURVEY**

**2.1 Existing problem**

* **Paper 1:**

Paper Title: Predicting IMDb Rating of Movies by Machine Learning Techniques

Journal / Conference Name & Year: Conference Paper,july 2019

Problem Statement: Film Industry is not only a industry or a centre of entertainment, rather it is now a centre of global business. All over the world is now excited about a movie’s box office success, popularity etc. A huge data is available online about these movies success or popularity. We have used hollywood movie list from Wikipedia and their rating from IMDb movie rating website to create our data set.

 Then machine learning classification algorithms are applied of the data set. Lastly an efficient model is developed to predict a movie’s IMDb rating. The model gives good classification measures with the data set. Index Terms—Movie Rating, Machine Learning, Prediction, Box Office, IMDb

Algorithms: Bagging ,Random Forest, J48 ,IBK ,Naive Bayes

Tools used: Weka 3.8.3 tool

Results attained: Among the five classifiers, random forest gives the highest classification accuracy. Though all the five classifiers give significantly good accuracy here, that is above 90%

* **Paper 2:**

Paper Title: Movie Rating Prediction using Ensemble Learning Algorithms

Journal / Conference Name & Year: International Journal of Advanced Computer Science and Applications, Vol. 11, No. 8, 2020

Problem Statement: Over the last few decades, social media platforms have gained a lot of popularity. People of all ages, gender, and areas of life have their presence on at least one of the social platforms. The data that is generated on these platforms has been and is being used for better recommendations, marketing activities, forecasting, and predictions. Considering predictions, the movie industry worldwide produces a large number of movies per year. The success of these movies depends on various factors like budget, director, actor, etc. However, it has become a trend to predict the rating of the movie based on the data collected from social media related to the movie. This will help a number of businesses relying on the movie industry in making promotional and marketing decisions. In this report, the aim is to collect movie data from IMDB and its social media data from YouTube and Wikipedia and compare the performance of two machine learning algorithms.

Algorithms: Random Forest and XGBoost

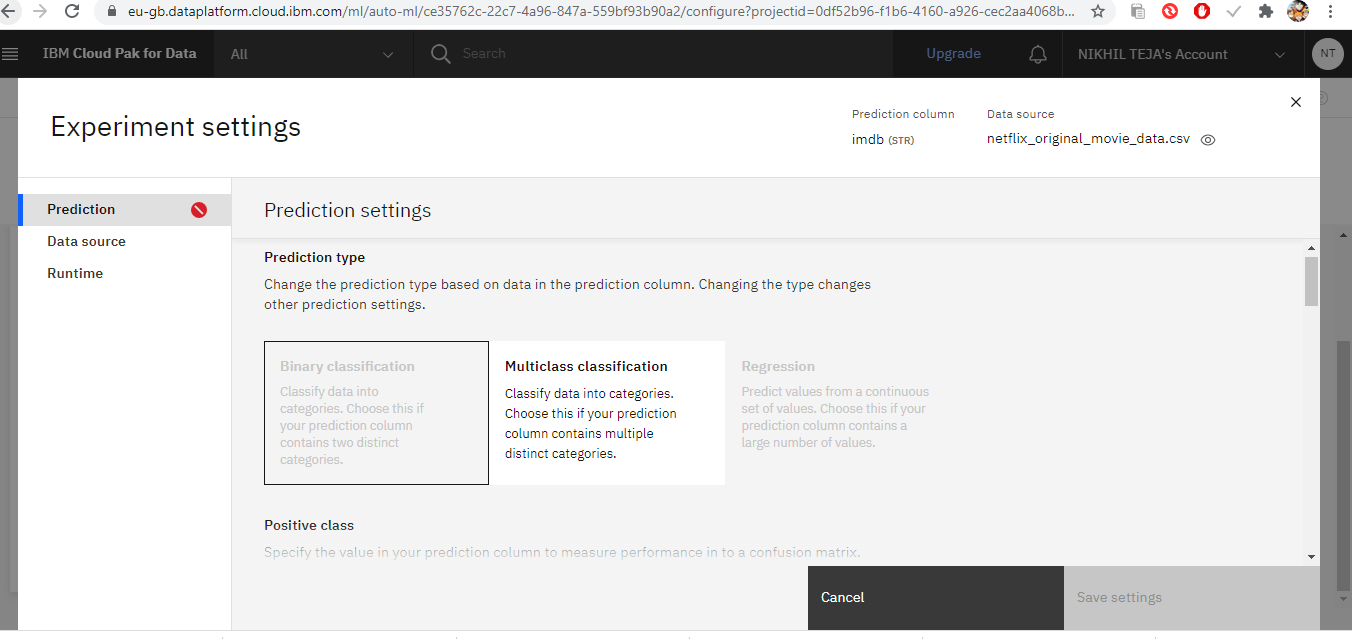
Tools used: Weka 3.8.3 tool

Results attained: The implementation results show that XGBoost is the better performer than the Random Forest algorithm for both datasets. And the data that contains both metadata and social media data performs better in predicting the ratings. Though, Random Forest in itself gives a good performance measure, it works well on training and test data.

**2.2 Proposed solution**

 What is the method or solution suggested by you?

 From the above mentioned both papers one paper used random forest and XG Boost and another paper used some Machine learning algorithms they used only limited algorithms.   
We need to add the data set into the Auto AI Experiment after that as shown in the below figure

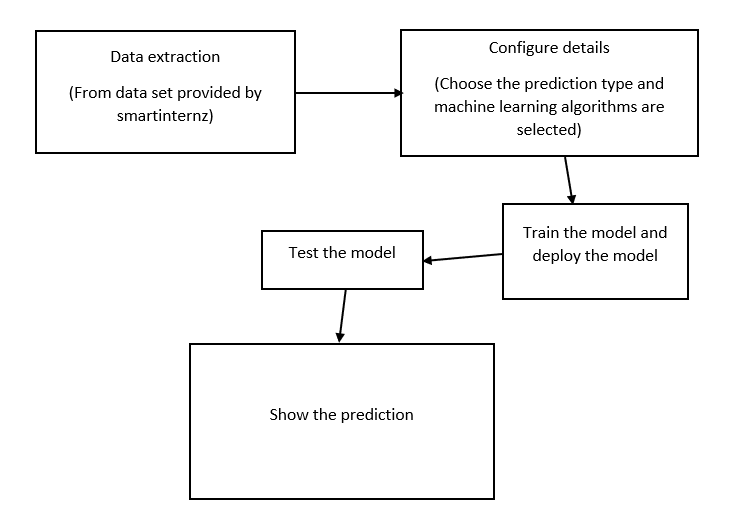


It will automatically choose prediction type according to the prediction type the algorithms for the particular data set will be shown.

For our data set it choosed regression and The algorithms are Ridge and XGB Regressor.

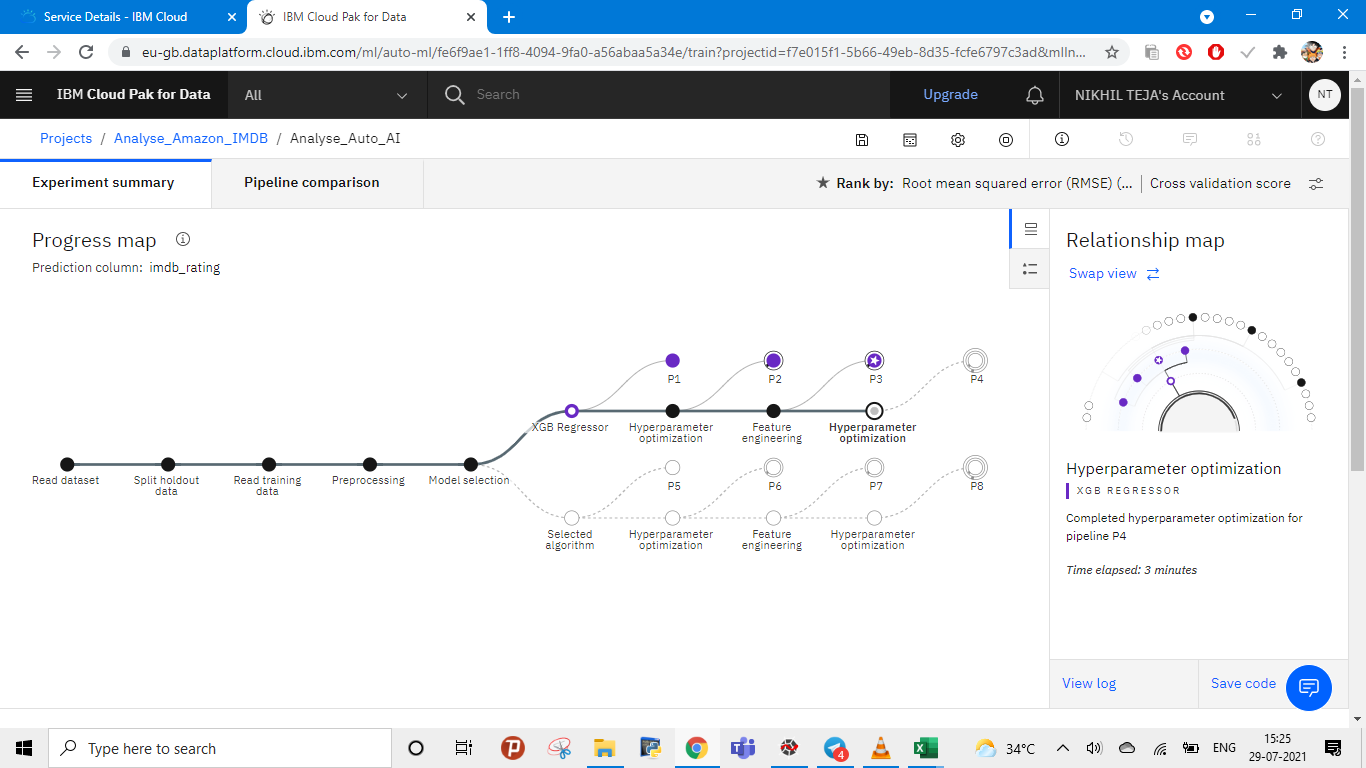
**3.THEORITICAL ANALYSIS**

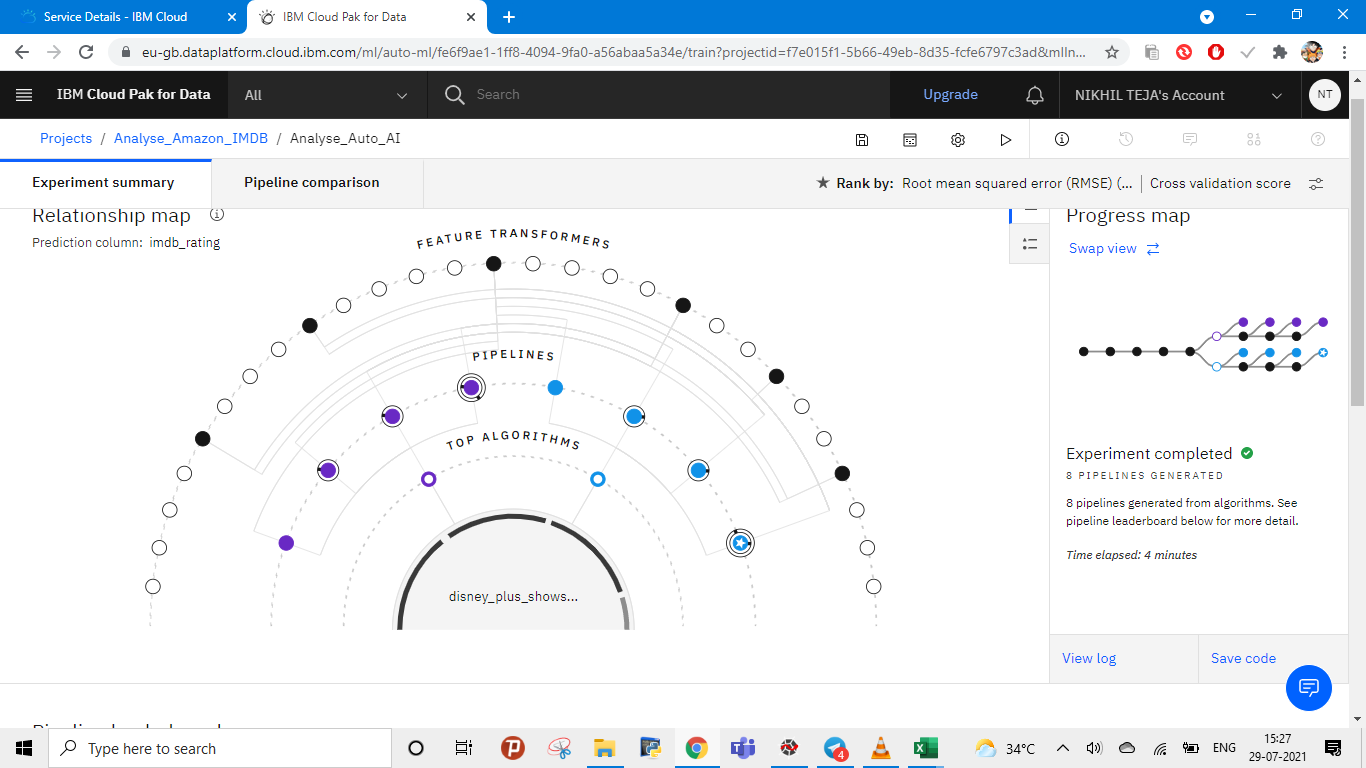
**3.1 Block Diagram Diagrammatic Overview Of The Project.**

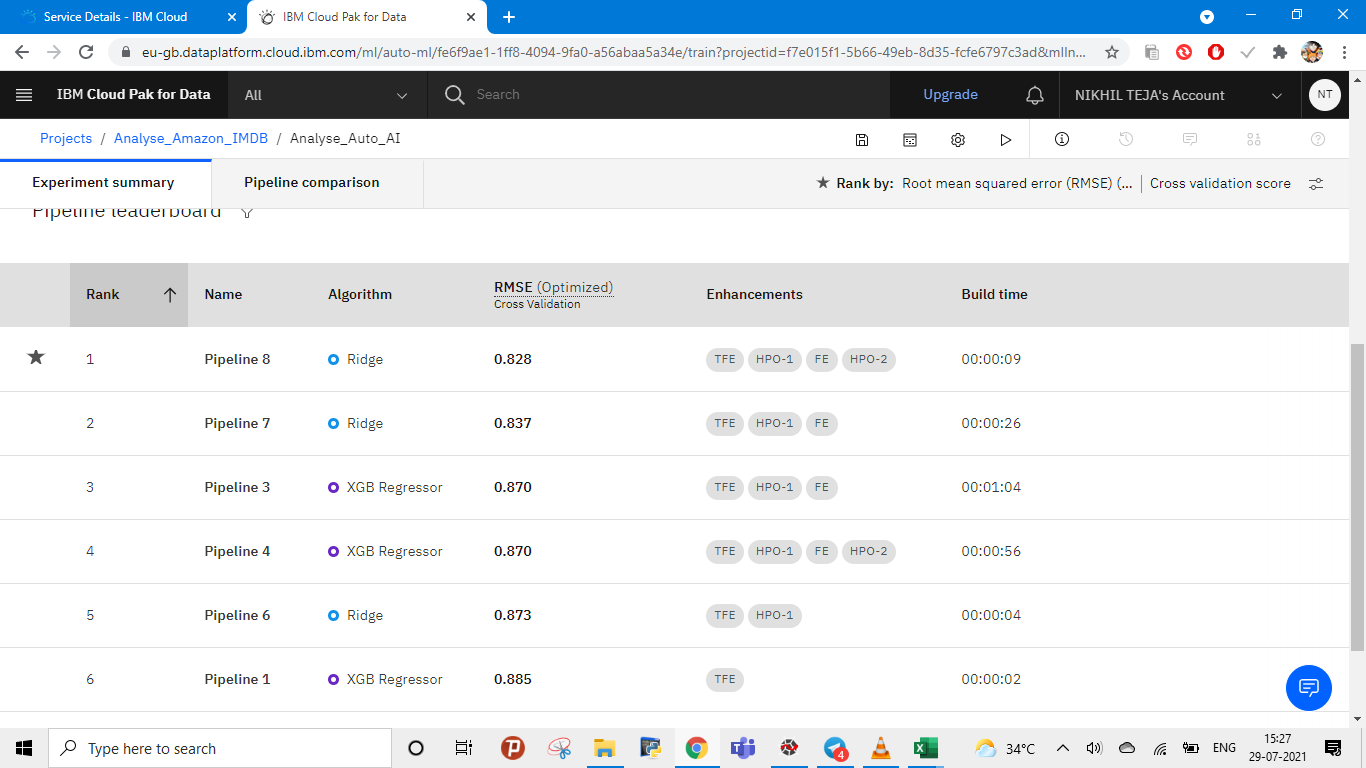
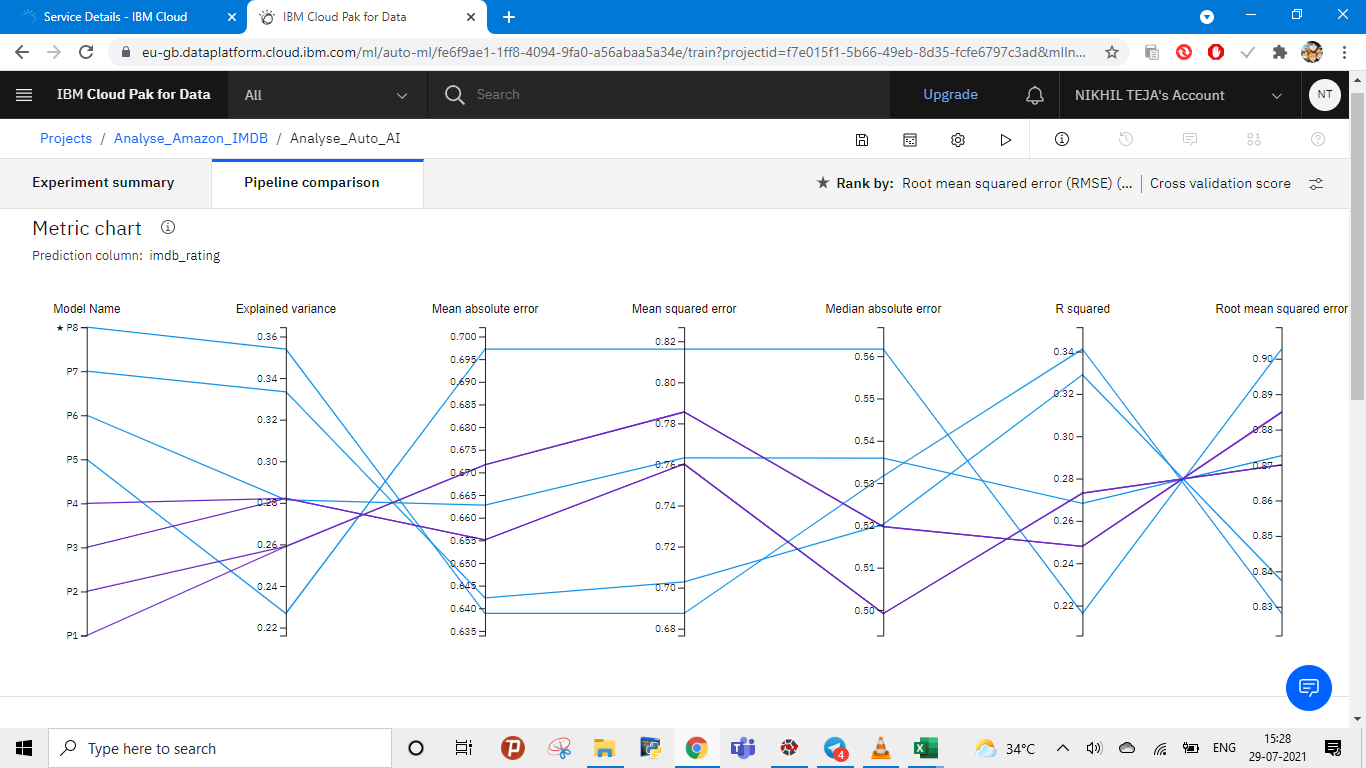


**4.EXPERIMENTAL INVESTIGATIONS**

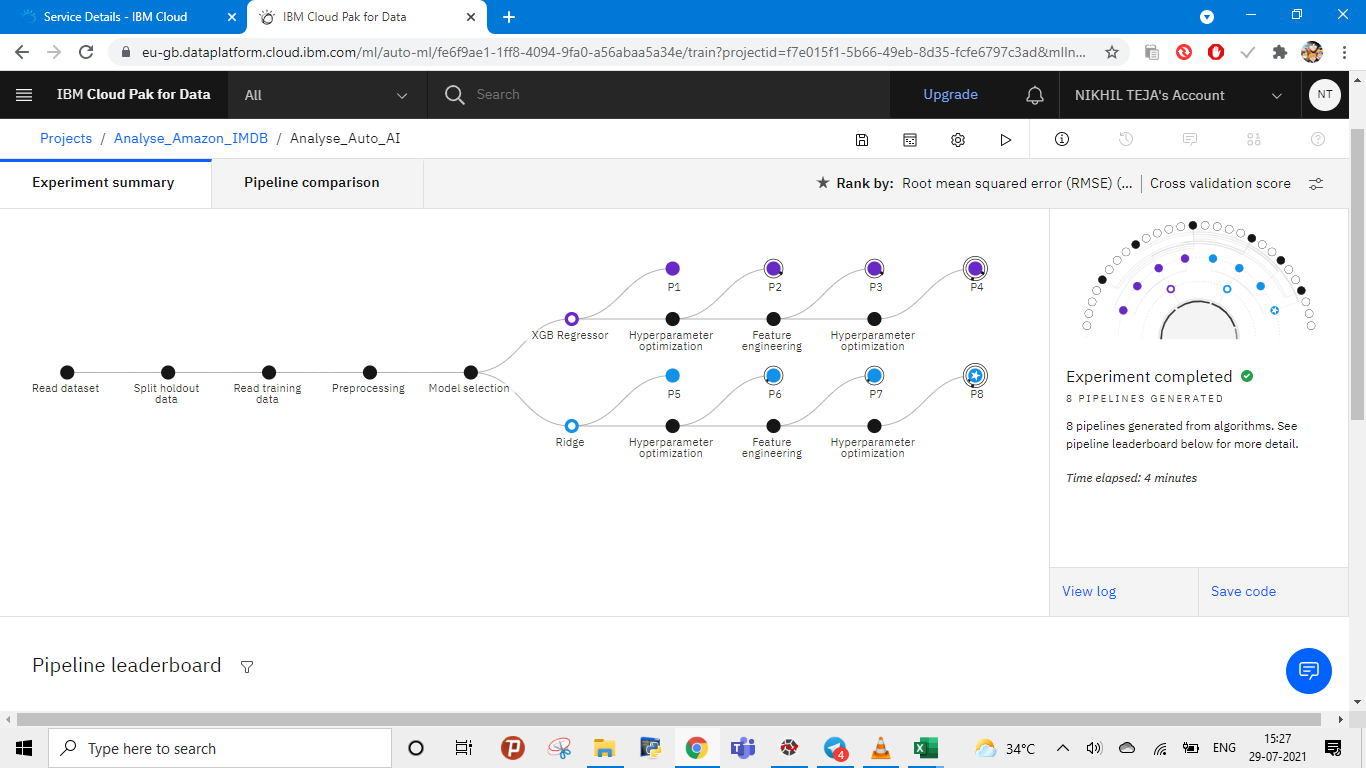
Analysis made while working on the solution

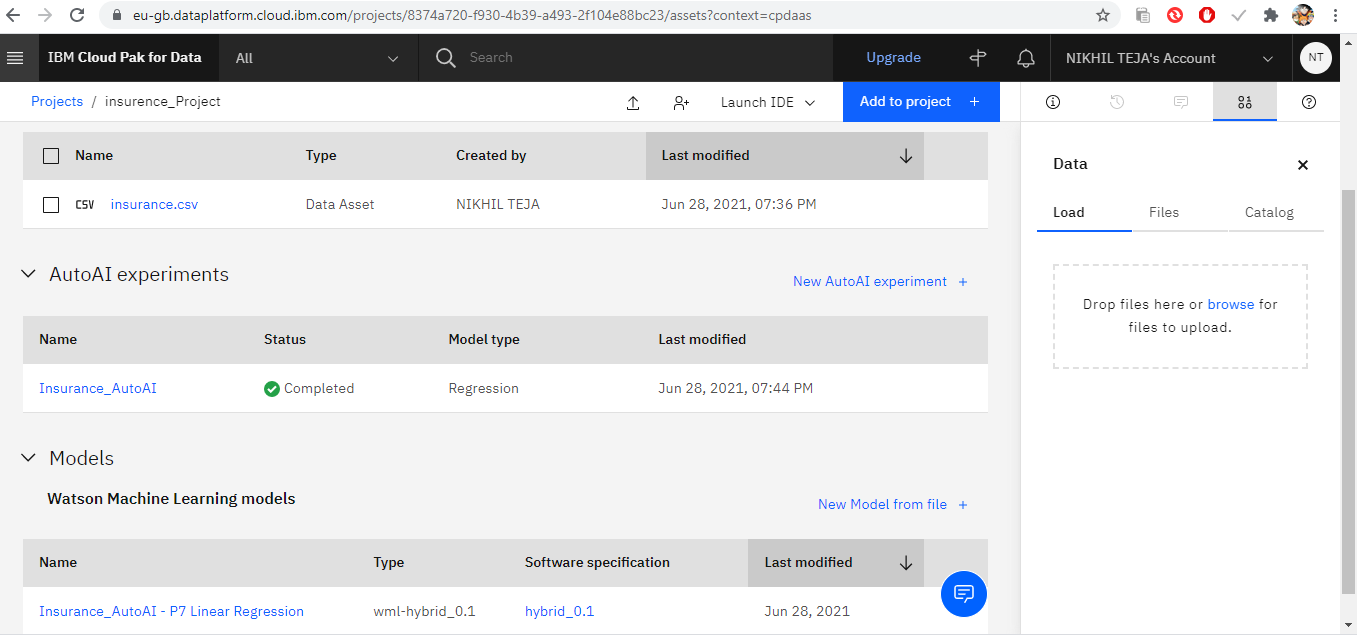




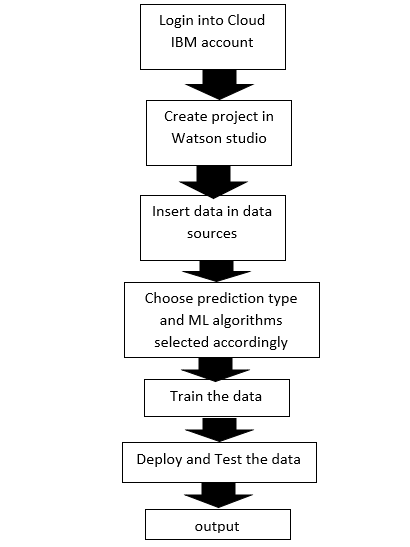


The star represents the best algorithm suitable for the project among the remaining algorithms



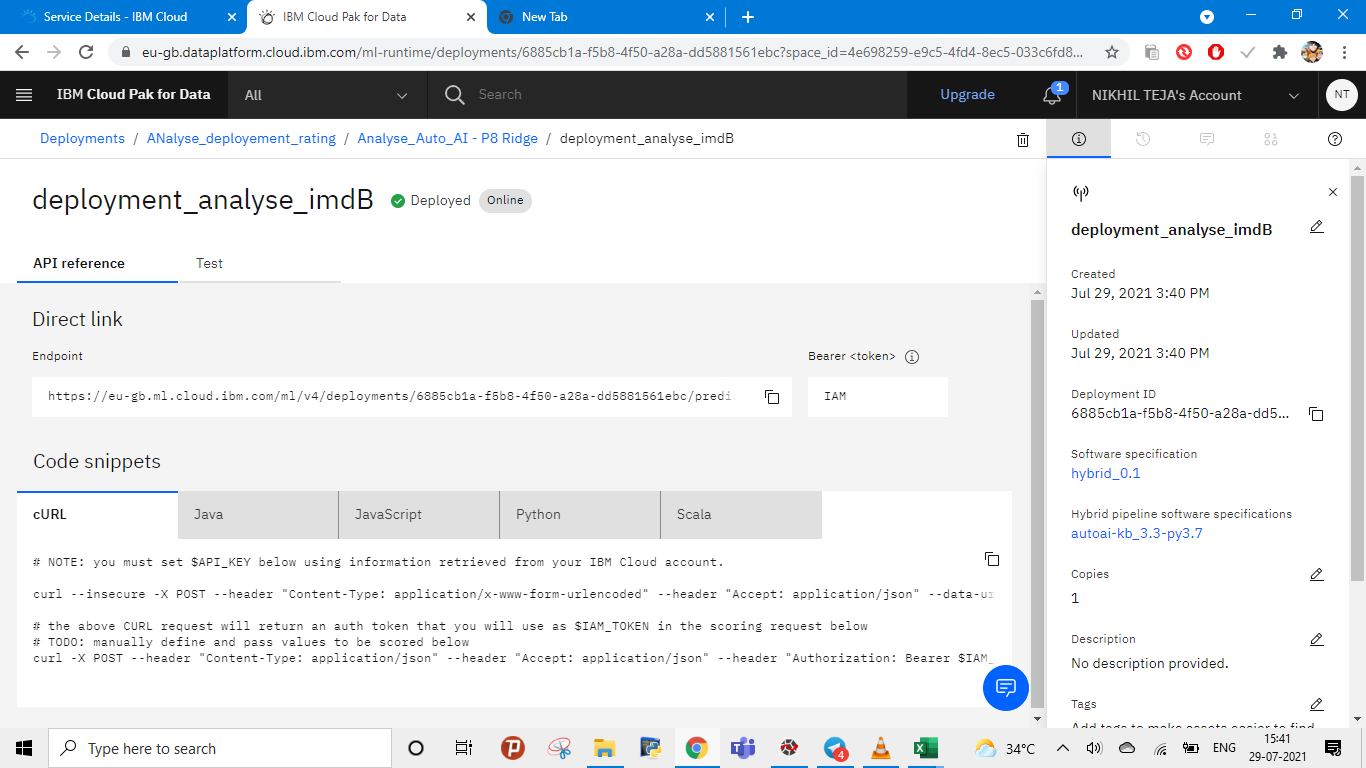


**5.FLOWCHART**



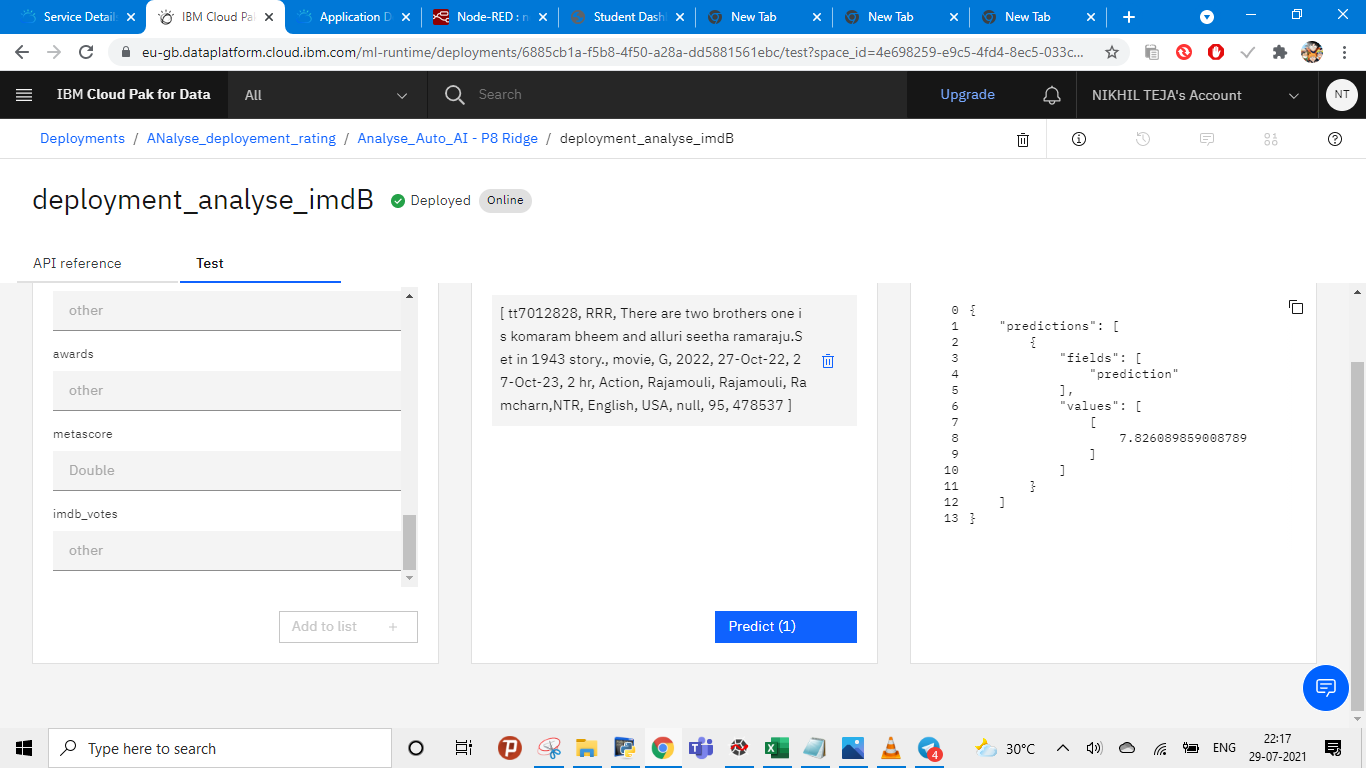
**6.RESULT**

After we deployed it is shown as below figure



This is my API key

After giving the inputs it will the output as shown below



Movie rating is shown here

Entered inputs are shown here

Here we need to type inputs

**ADVANTAGES**

There are numerous beneficial things about movie ratings.

Some people watch film only after knowing the rating, some may only watch the film if only it has more rating.

Director and artists future depend on the rating because if only the rating is good, they get some more chances in the industry.

Additionally, this gives guardians more control on the thing their kid is watching or review. Evaluations additionally assist with holding youngsters and youthful grown-up back from watching and bareness, illicit drug use, and solid brutality. This can assist with holding a kid back from watching something they didn't wish to see. In case guardians are reconsidering about a film they can generally check the audits, rating, or even proceed to watch the film all alone.

**DISADVANTAGES**

There are a couple of thing that should be remembered about Imdb movie ratings. Each evaluating a film has doesn't imply that all the substance that make up the specific rating will be in the film. Additionally, numerous individuals watch films without parental direction and wind up seeing things that their folks don't endorse. Moreover, some film evaluations are over looked by guardians and the watcher might wind up seeing things the person in question shouldn't. Additionally, film rating can likewise be miss driving and cause certain individuals not to watch the film that is the reason is imperative to know what the film contains or to go see it before you permit your youngster to watch it first. A few group, despite the fact that under matured, do sneak into motion pictures that perhaps excessively extreme for them to watch. These are a few imperfections in the film rating framework.

**APPLICATIONS**

* GroupLens:
  + Helped in developing initial recommender systems by pioneering collaborative filtering model
  + It also provided many data-sets to train models including MovieLens and BookLens
* Amazon:
  + Implemented commercial recommender systems
  + They also implemented a lot of computational improvements
* Netflix Prize:
  + Pioneered Latent Factor/ Matrix Factorization models
* Google-Youtube:
  + Hybrid Recommendation Systems
  + Deep Learning based systems
  + Social Network Recommendations

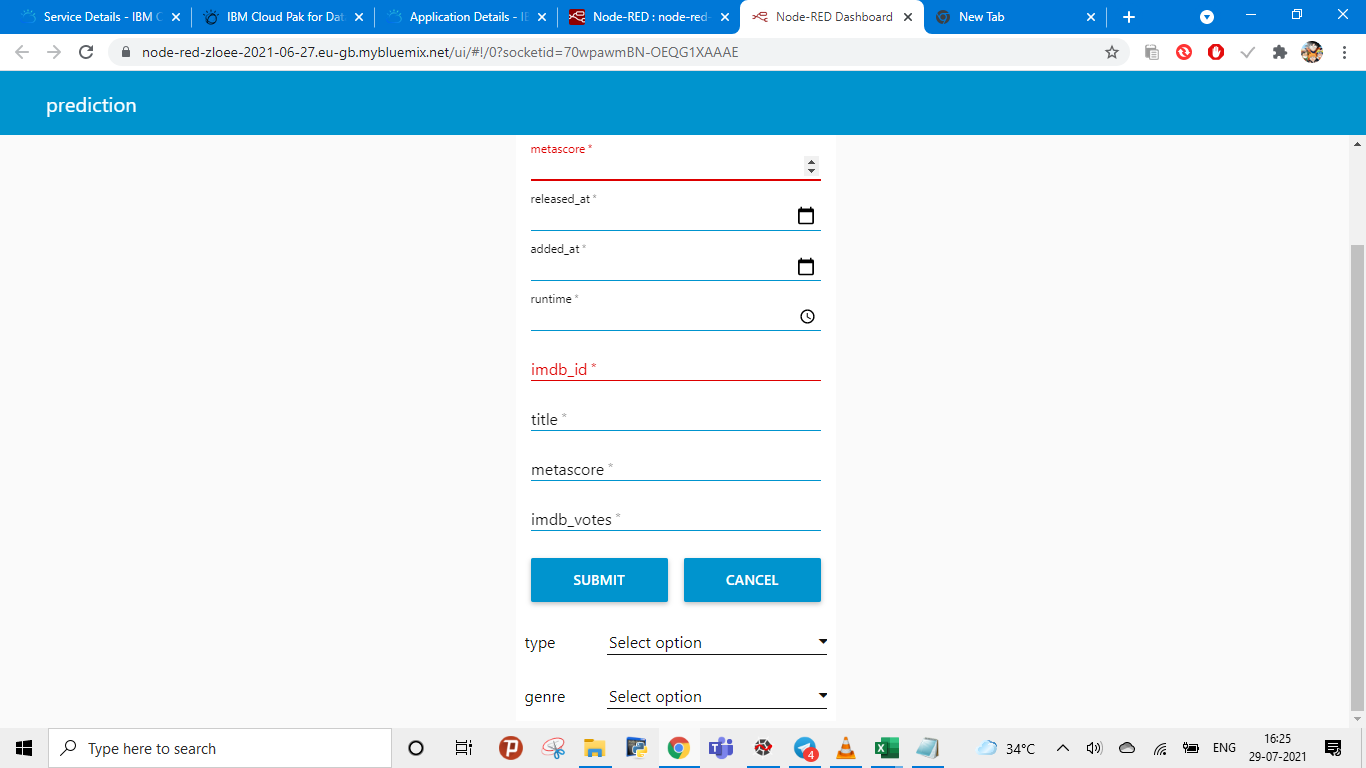
**9.CONCLUSION**

* Log into the cloud account-Completed
* Create IBM Watson Studio Service-Completed
* Create IBM Watson Studio Project-Completed
* Associate a cloud object storage-Completed
* Assosiate auto AI Experment-Completed
* Train the models-Completed
* Create the deployment service-Completed
* Test the model-Completed
* Show the output-Completed

**10.FUTURE SCOPE**

* Create the UI-Completed
* Grab the values-Completed
* Set global variables to the UI—Completed
* Get the access token with the help of http request-Completed
* Send the values to scoring end point along with the access token using http request node-Pending

This is the interface created using node red.



This is the backend created up to access token in node red.

