Using BigQuery in ML

We have used machine learning algorithms to build, train and evaluate the model using BigQuery on Google Cloud Platform. We will be applying Linear regression on the chosen dataset (Weather dataset predicting rain fall in Australia). First, we create a separate project. Then we create a bucket and upload the dataset which we need to analyze using BigQuery. Then in BigQuery we create a dataset and further create tables inside the dataset. The project name which we will be using is kbsbigquery, dataset name is rainbigquery and the table name is rainpred. Moving on we will be writing linear regression codes to build, train and evaluate the model.

1. Code:
2. Code for building the model.

#standardSQL

CREATE MODEL `rainbigquery.PredictiingRain\_model`

OPTIONS

(model\_type='linear\_reg',

input\_label\_cols=['RainTomorrow']) AS

SELECT

Location, MinTemp, MaxTemp, Rainfall, Evaporation, Sunshine, WindGustDir, WindGustSpeed, WindDir9am, WindDir3pm, WindSpeed9am, WindSpeed3pm, Humidity9am, Humidity3pm, Pressure9am, Pressure3pm, Cloud9am, Cloud3pm, Temp9am, Temp3pm, RainToday, RISK\_MM, RainTomorrow

FROM

`kbsbigquery.rainbigquery.Prediction`

WHERE

RainTomorrow IS NOT NULL;

1. Code for evaluating the model using ML. EVALUATE:

#standardSQL

SELECT

\*

FROM

ML.EVALUATE(MODEL `rainbigquery.PredictiingRain\_model`,

(

SELECT

Location, MinTemp, MaxTemp, Rainfall, Evaporation, Sunshine, WindGustDir, WindGustSpeed, WindDir9am, WindDir3pm, WindSpeed9am, WindSpeed3pm, Humidity9am, Humidity3pm, Pressure9am, Pressure3pm, Cloud9am, Cloud3pm, Temp9am, Temp3pm, RainToday, RISK\_MM, RainTomorrow

FROM

`kbsbigquery.rainbigquery.Prediction`

WHERE

RainTomorrow IS NOT NULL

));

1. Code for predicting the model using ML. PREDICT:

#standardSQL

SELECT

predicted\_RainTomorrow

FROM

ML.PREDICT(MODEL `rainbigquery.PredictiingRain\_model`,

(

SELECT

Location, MinTemp, MaxTemp, Rainfall, Evaporation, Sunshine, WindGustDir, WindGustSpeed, WindDir9am, WindDir3pm, WindSpeed9am, WindSpeed3pm, Humidity9am, Humidity3pm, Pressure9am, Pressure3pm, Cloud9am, Cloud3pm, Temp9am, Temp3pm, RainToday, RISK\_MM, RainTomorrow

FROM

`kbsbigquery.rainbigquery.Prediction`

WHERE

Location = "Albury")

1. Results and Analysis:

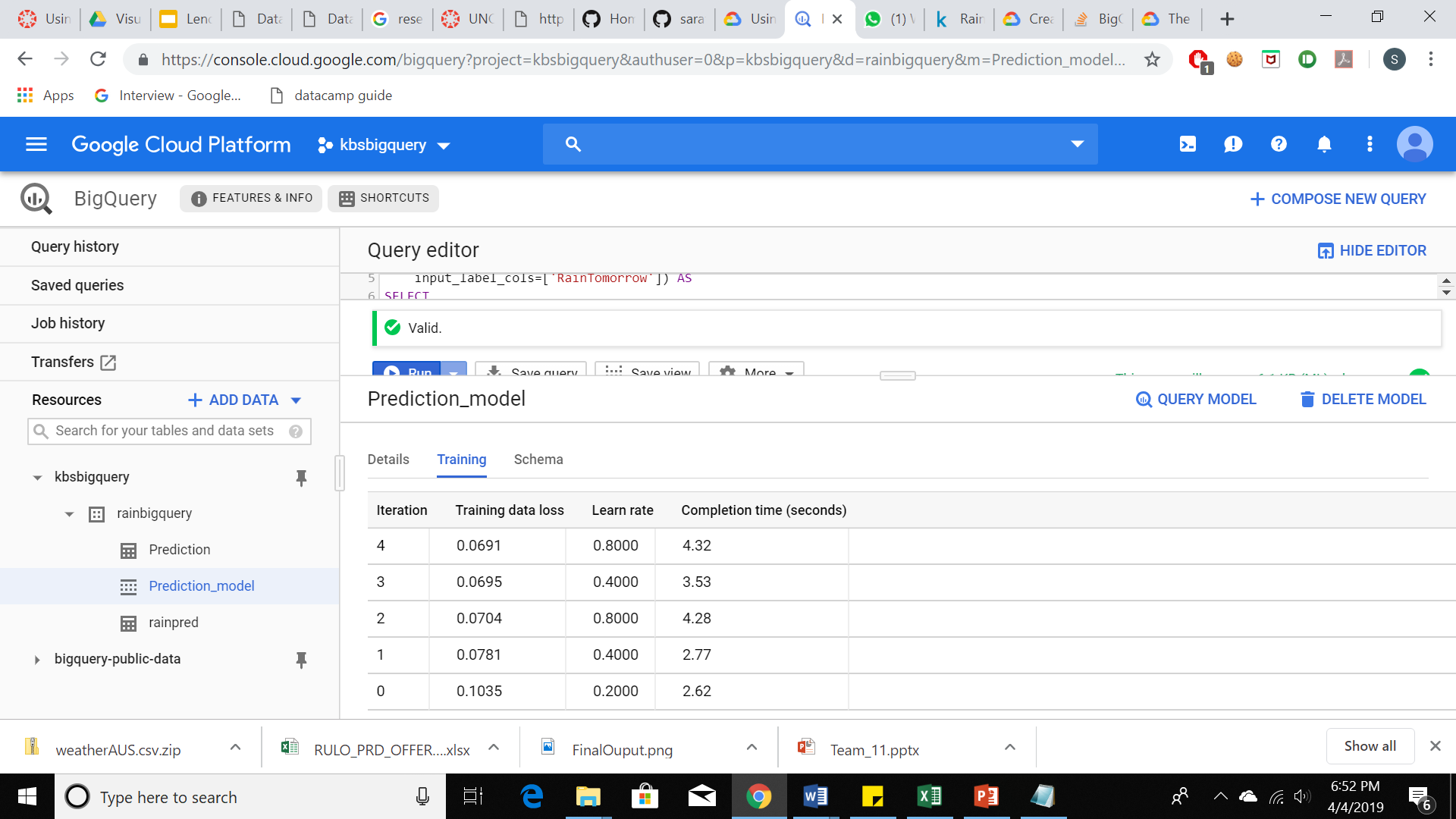


Fig 1. Building the prediction model

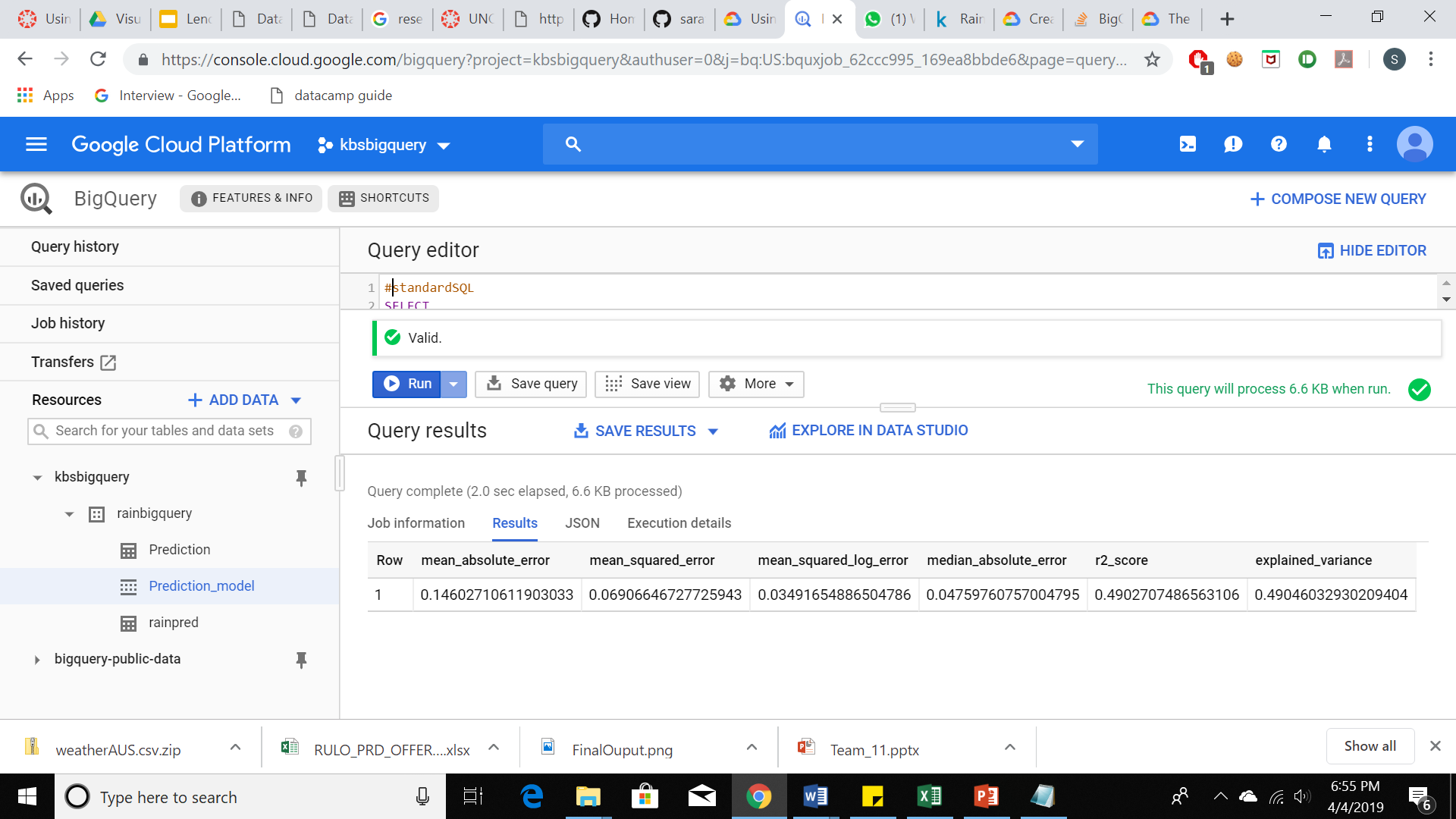


Fig 2. Evaluating the model built

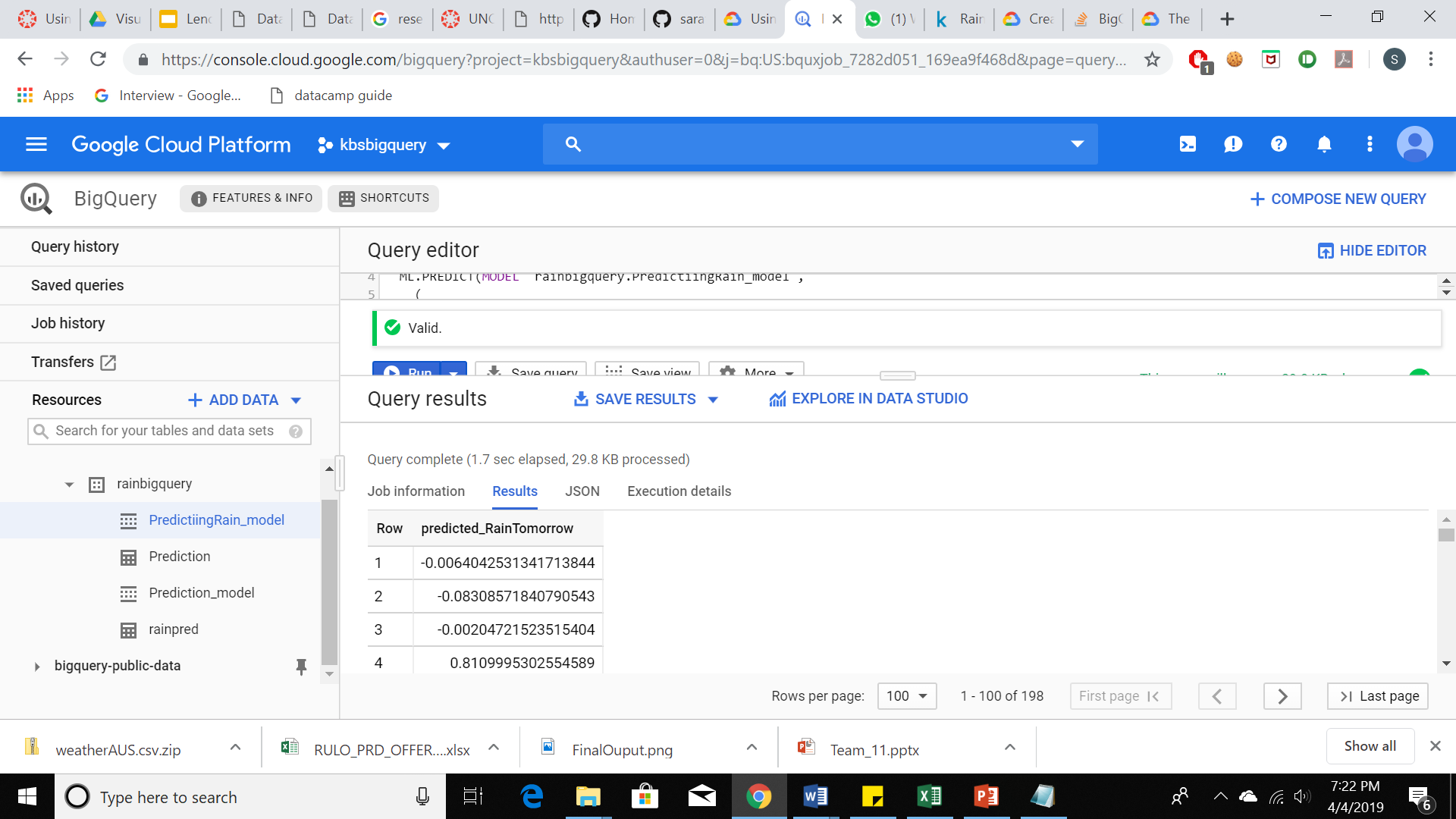


Fig 3. Predicting the evaluated model.

Figure 1 as shown above displays the model created and the details of training dataset where Training data loss with learn rate is displayed and the time taken for each iteration to complete. We can also see the schema of the built model. In Figure which depicts evaluating the model, we can find information about mean absolute error, mean squared error, log error, absolute error and variance score. These scores and values give useful information about our model created after evaluation. Figure 3 as shown above gives information about the rain fall predicted tomorrow. The scores indicate if a rain fall will occur tomorrow or not ranging between 0 and 1. So we have taken the dataset, selected the required or desired features and built, trained and predicted the model using linear regression and BigQuery accurately. In this manner, the model can be expanded to any weather datasets all over the world.