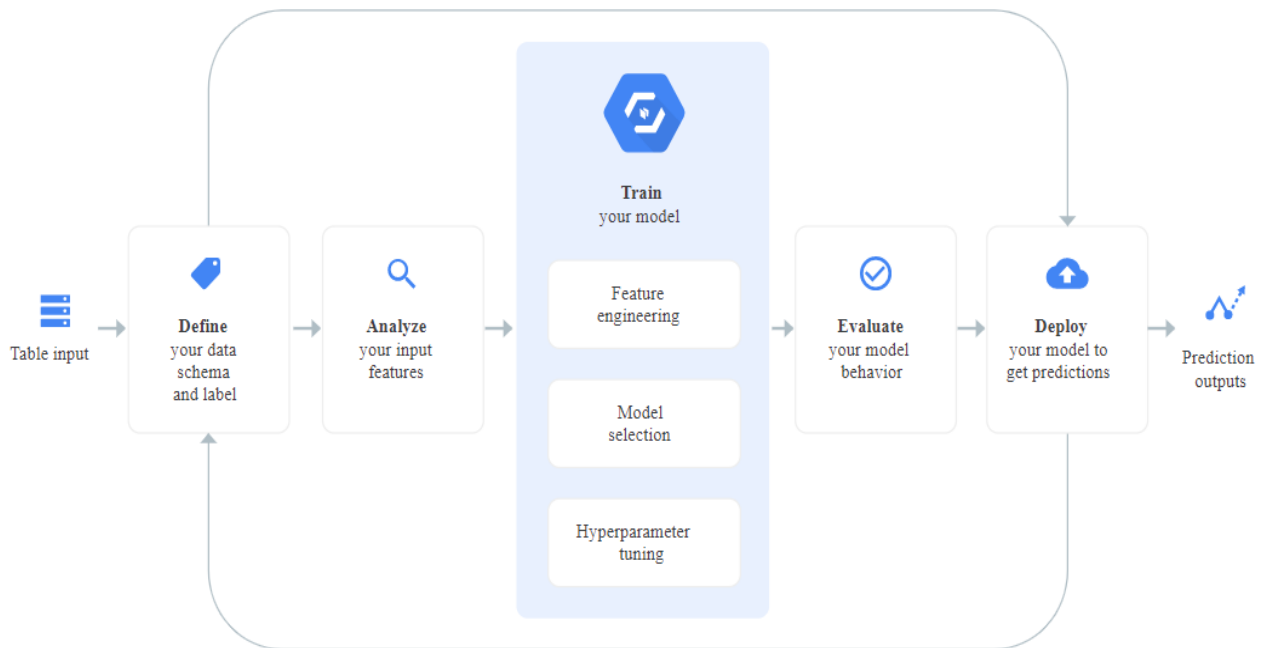
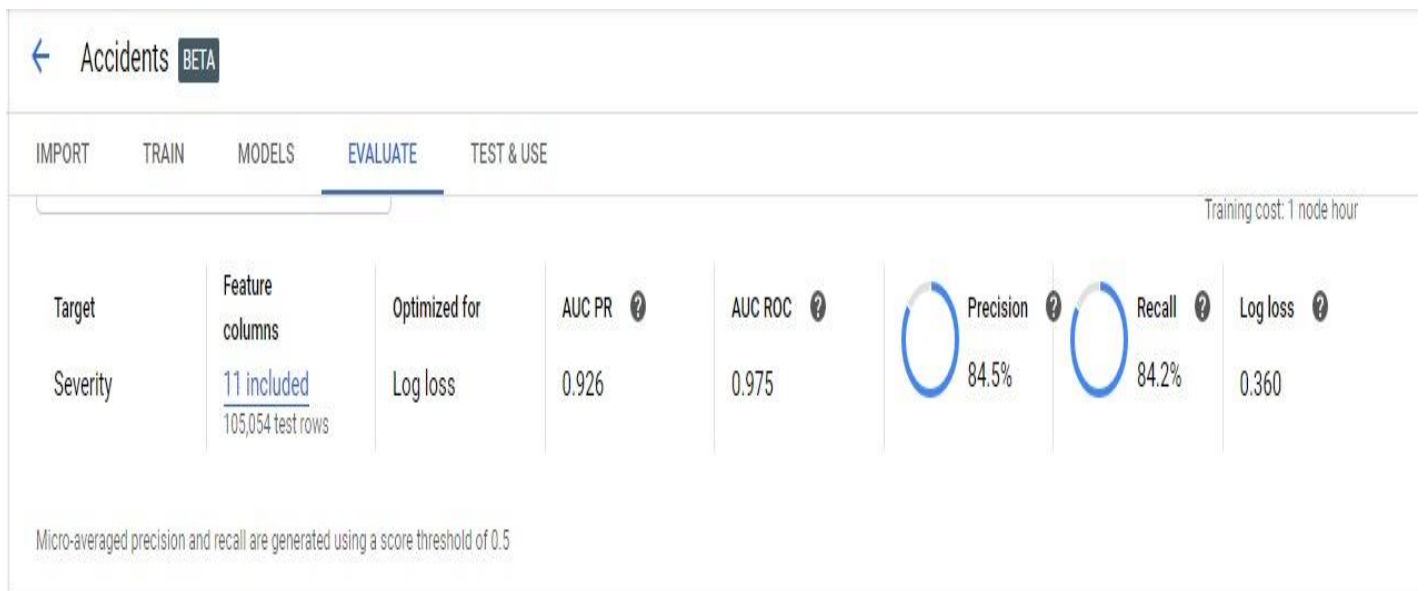


Auto ML

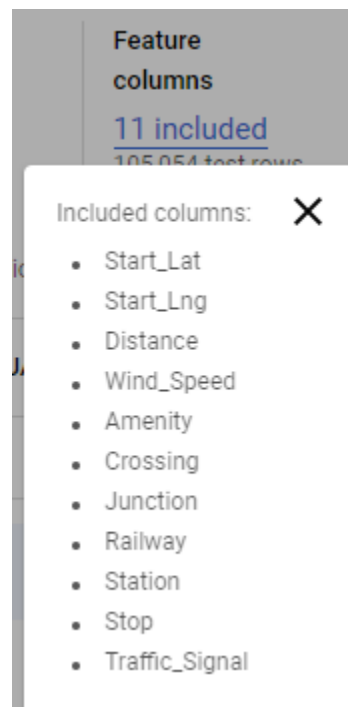
How AUTO ML WORKS:



MODELLING RESULTS:



Below features are included for training our model, these features are chosen as per their correlation with the target variable “**Severity**”



MODEL USED IN AUTO ML: GRADIENT BOOSTING DECISION TREE ALGORITHM

Gradient boosting algorithm sequentially blends weak learners so that each new learner matches the residuals of the previous step in order to improve the model. The final model aggregates the outcomes from each move and reaches a strong learner. Gradient-boosted algorithm for decision trees uses decision trees as weak-learners. The residuals are identified via a loss function. For example, *for a regression task, mean squared error (MSE) can be used, and logarithmic loss (log loss) can be used for classification tasks.* It should be noted that the current trees in the model do not change with the introduction of a new tree

Hyperparameters are key components of learning algorithms that affect the performance and accuracy of the model. Learning rate and n estimators are two important hyperparameters for gradient boosting decision trees in gradients. Learning rate, called α , simply means how quickly the model learns. Every added tree modifies the model overall. Learning rate regulates the extent of the modification.

LOGS VIEWER:

```
{
  "insertId": "w72sc0g4m7lvs5",
  "jsonPayload": {
    "@type": "type.googleapis.com/google.cloud.automl.master.TablesModelStructure",
    "modelParameters": [
      {
        "hyperparameters": {
          "Max tree depth": 9,
          "Tree L2 regularization": 3,
          "Tree L1 regularization": 3,
          "Center Bias": "False",
          "Tree complexity": 1,
          "Model type": "GBDT",
          "Number of trees": 150
        }
      },
      {
        "hyperparameters": {
          "Number of trees": 150,
          "Max tree depth": 9,
          "Center Bias": "False",
          "Tree L2 regularization": 3,
          "Tree L1 regularization": 3,
          "Tree complexity": 1,
          "Model type": "GBDT"
        }
      },
      {
        "hyperparameters": {
          "Number of trees": 300,
          "Max tree depth": 12,
          "Tree L2 regularization": 3,
          "Tree L1 regularization": 1,
          "Center Bias": "False",
          "Tree complexity": 3,
          "Model type": "GBDT"
        }
      }
    ]
  }
}
```

RESULTS IN BIGQUERY

BigQuery

FEATURES & INFO

SHORTCUT

Query history

Saved queries

Job history

Transfers

Scheduled queries

Reservations

BI Engine

Resources

Unsaved query Edited

+ COMPOSE NEW QUERY

HIDE EDITOR

FULL SCREEN

1 SELECT * FROM `group-6-kbs.export_evaluated_examples_Accidents_20200505113401_2020_05_05T22_17_00_406Z.evaluated_examples` Where Severity = '2' LIMIT 1000

Valid.

Run Save query Save view Schedule query More

This query will process 10.4 MB when run.

Query results

SAVE RESULTS EXPLORE DATA

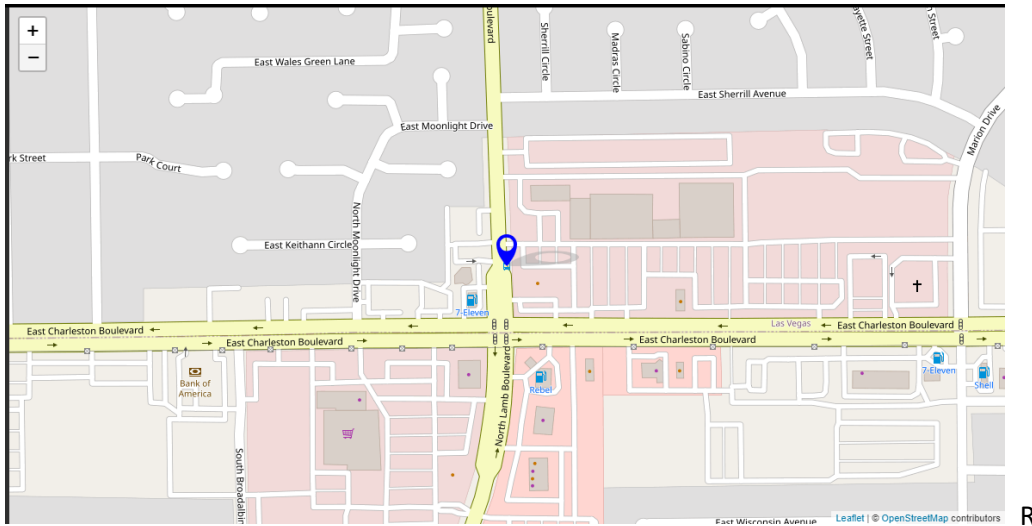
Query complete (0.4 sec elapsed, 10.4 MB processed)

Job information Results JSON Execution details

Severity	Start_Lat	Start_Lng	Distance	Wind_Speed	Amenity	Crossing	Junction	Railway	Station	Stop	Traffic_Signal	predicted_Severity.tables.score	predicted_Se
3	33.814602	-118.176018	0.4399999998	0.0	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	0.4576586186885834	2
												0.5100741386413574	3
												0.031102098524570465	4
												0.0011650783708319068	1
3	39.623886	-104.90094	1.6799999948	0.0	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	0.6518879532814026	2
												0.33827394247055054	3
												0.0060858007306657	4

ONLINE PREDICTION:

We Used **Charleston Boulevard and Lamb Boulevard**(**36.159629° -115.080135°**) in las vegas for online prediction, given that there is a stop, junction and presence of traffic signals, distance affected 10 mph, we obtained a **severity level prediction of 2** which is justified. As there is a chance traffic can be affected for several distance we can also see that Severity of 3 has 41% probability.



R

Predict label

Severity

Prediction result

Baseline prediction value: 0.716

2



3



4



1

