CREATE OR REPLACE TABLE `project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2`

AS

SELECT \*,

 CASE

  WHEN split\_field < 0.8 THEN 'training'

  WHEN split\_field >= 0.8 THEN 'evaluation'

 END AS split\_dataframe

FROM (

  SELECT \* , ROUND(ABS(RAND()),1) AS split\_field

  FROM `project-tuananh-2023.Tuananh\_ML.Pizza\_new`

);

SELECT split\_dataframe, count(split\_field) as number\_row

FROM  `project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2`

GROUP BY split\_dataframe ;

-- Tạo mô hình LOGISTIC REGRESSION

CREATE OR REPLACE MODEL `project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2\_logistic\_model`

OPTIONS(model\_type='LOGISTIC\_REG',

        auto\_class\_weights = TRUE,

        early\_stop = TRUE,

        input\_label\_cols = ['brand'])

        AS

SELECT

\* EXCEPT (int64\_field\_0,id, split\_dataframe, split\_field)

FROM

`project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2`

WHERE split\_dataframe LIKE 'training';

--  Đánh giá mô hình

SELECT \*

FROM ML.EVALUATE(MODEL `project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2\_logistic\_model`,(

SELECT

\* except(int64\_field\_0,id, split\_dataframe, split\_field )

FROM

`project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2`

WHERE split\_dataframe like 'evaluation'));

SELECT \*

FROM ML.CONFUSION\_MATRIX(MODEL `project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2\_logistic\_model`,(

SELECT

\* except(int64\_field\_0,id, split\_dataframe, split\_field )

FROM

`project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2`));

WHERE split\_dataframe LIKE 'training';

---- Tạo mô hình  BOOSTED TREE

CREATE OR REPLACE MODEL `project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2\_boostedtree\_classifier\_model`

OPTIONS(model\_type='BOOSTED\_TREE\_CLASSIFIER',

        NUM\_PARALLEL\_TREE = 50,

        BOOSTER\_TYPE = 'GBTREE',

        EARLY\_STOP = TRUE,

        INPUT\_LABEL\_COLS = ['brand']

  ) AS

SELECT \* EXCEPT(int64\_field\_0,id, split\_dataframe, split\_field)

FROM `project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2`

WHERE split\_dataframe like 'training';

--  Đánh giá mô hình

SELECT \*

FROM ML.EVALUATE(MODEL  `project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2\_boostedtree\_classifier\_model`,(

SELECT

\* except(int64\_field\_0,id, split\_dataframe, split\_field)

FROM

`project-tuananh-2023.Tuananh\_ML.Pizza\_new\_2`

WHERE split\_dataframe like 'evaluation'));