**README:**

Steps to execute the Code:

Executing train.py code:

* First, train.json is preprocessed using the python code.
* Copy train.py into a python editor.
* Make sure that the train.json is copied into project’s directory.
* Run the train.py file.
* Generate a csv file and load this file into a Databricks table.
* Note the link for table (TrainTable).

Executing test.py code:

* First, test.json is preprocessed using the python code.
* Copy test.py into a python editor.
* Make sure that the test.json is copied into project’s directory.
* Run the test.py file.
* Generate a csv file and load this file into a Databricks table.
* Note the link for table (TestTable).

Executing code on databricks:

* The testfile variable is loaded with TestTable link and datafile variable is loaded with TrainTable link.
* Run all the queries.
* The comments in the code mention the functionality of the block.
* The results are seen after each block.

The published databricks link is

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/1663046746195488/4164515032482306/863366668872171/latest.html>