Evaluating Machine Learning Algorithms on Google BigQuery ML

Ayush Sharma (0774583)
Jasmeet Singh (0758054)

TRENT UNIVERSITY

INTRODUCTION

- Title: Evaluating Machine Learning Algorithms on Google BigQuery ML
- Aim: Assess the capabilities of BigQuery ML using Logistic Regression & KNN.
- Dataset: Human resources dataset (10,000 and 1 million entries) for classification tasks.
- Objective: Understand BigQuery ML's efficacy in diverse machine learning scenarios.

TRENT UNIVERSITY

EXPERIMENT & RESULTS

- Tools Utilized:
  - BigQuery Studio, Looker Studio, Jupyter Notebook.
  - Logistic Regression (precision, recall, accuracy) & K-means (DBI, MSD).
- Experiment Design:
  - Comprehensive evaluation of algorithms on varied dataset scales.
  - Integrated functionalities for analysis and documentation.
- Results:
  - Distinctive performance of Logistic Regression and KNN across datasets.
  - Insights into algorithm performance variability concerning dataset sizes and variables.

TRENT UNIVERSITY

CONCLUSION & RECOMMENDATIONS

- Conclusion:
  - BigQuery ML's strengths: Integration and handling large-scale datasets.
  - Considerations: Trade-offs between resources, execution times, and expenses.
  - Algorithm Performance: Variable precision in Logistic Regression and consistency in K-means.
- Recommendations:
  - Adoption based on reliance on Google Cloud, data scale, and specific ML tasks.

TRENT UNIVERSITY

THANK YOU!

TRENT UNIVERSITY