Assignment 7. Reading Assignment for Machine Learning on Relational Database

Due date: 11:59pm Apr 3rd 2020

Files to submit:

- A report in PDF format. Select three papers and then in the report, write a 1-page review for each selected paper. Each review must include three parts:

  (1) Summary of the paper

  (2) Three strengths of the paper

  (3) Three weaknesses of the paper

Submission Website:

GradeScope (We will NOT accept submissions via email and Canvas).

If you have any questions regarding GradeScope, please contact TA.

Learning Goal:

1. Understand the motivating factor of running machine learning in RDBMS: declarative programming, automatic query optimization, normalized data model, and etc.;
2. Understand the challenges of running machine learning in RDBMS: recursive interface, how to represent tensors , and etc.;
3. Understand the current state of arts in running machine learning in RDBMS.

Tasks:

1. Select three papers from the list:

[1] Jankov, Dimitrije, et al. "Declarative recursive computation on an RDBMS: or, why you should use a database for distributed machine learning." *Proceedings of the VLDB Endowment* 12.7 (2019): 822-835. **Declarative recursive computation on an RDBMS or, why you should use a database for distributed machine learning.pdf**

[2] Luo, Shangyu, et al. "Scalable linear algebra on a relational database system." *IEEE Transactions on Knowledge and Data Engineering* 31.7 (2018): 1224-1238.

Scalable linear algebra on a relational database system.pdf

[3] Boehm, Matthias, et al. "Systemml: Declarative machine learning on spark." *Proceedings of the VLDB Endowment* 9.13 (2016): 1425-1436.

Systemml Declarative machine learning on spark.pdf

[4] Zhang, Ce, Arun Kumar, and Christopher Ré. "Materialization optimizations for feature selection workloads." *ACM Transactions on Database Systems (TODS)* 41.1 (2016): 1-32.

Materialization optimizations for feature selection workloads.pdf

[5] Kumar, Arun, Jeffrey Naughton, and Jignesh M. Patel. "Learning generalized linear models over normalized data." *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*. 2015.

Learning generalized linear models over normalized data.pdf

[6] Feng, Xixuan, et al. "Towards a unified architecture for in-RDBMS analytics." *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*. 2012.

Towards a unified architecture for in-RDBMS analytics.pdf

2. Write three 1-page reviews, one review for each paper:

- Each review must include three parts:

  (1) Summary of the paper

  (2) Three strengths of the paper

  (3) Three weaknesses of the paper

3. Submit one PDF file that includes the three surveys to GradeScope Assignment7.

Autograder Results

Results | Code

This assignment does not have an autograder configured.

STUDENT
Parth Rajendra Doshi

AUTOGRADER SCORE
0.0 / 0.0

QUESTION 2.1
summary **1.0** / 1.0 pts

QUESTION 2.2
3 strengths **3.0** / 3.0 pts

QUESTION 2.3
3 weaknesses **3.0** / 3.0 pts