



Internship – Machine Learning for Optimizing Oracle Database Performance

Oracle

Oracle, a global provider of enterprise cloud computing, is empowering businesses of all sizes on their journey of digital transformation. Oracle Cloud provides leading-edge capabilities in software as a service, platform as a service, infrastructure as a service, and data as a service.

Oracle's application suites, platforms, and infrastructure leverage both the latest technologies and emerging ones—including artificial intelligence, machine learning, blockchain, and Internet of Things—in ways that create business differentiation and advantage for customers. Continued technological advances are always on the horizon.

Oracle Labs

Oracle Labs is the advanced research and development arm of Oracle. We focus on the development of technologies that keep Oracle at the forefront of the computer industry. Oracle Labs researchers look for novel approaches and methodologies, often taking on projects with high risk or uncertainty, or that are difficult to tackle within a product-development organization. Oracle Labs research is focused on real-world outcomes: our researchers aim to develop technologies that will someday play a significant role in the evolution of technology and society. For example, chip multithreading and the Java programming language grew out of work done in Oracle Labs.

Internship Details

Machine learning (ML) is at the forefront of the rising popularity of data-driven software applications. Recent studies show that learning techniques can be used not only to boost performance of pure data-driven application (such as recommender systems, image recognition, etc) but also to learn whole components/parts of a processing engine such as a database. For example, SageDB [1] is a database where core components are replaced with learned models that can provide superior performance. Other examples, such as Bao [2], show how tree neural networks and advanced sampling methods can be leveraged in order to learn hints that can boost the performance of the query optimizer.

The internship will consist in exploring how similar learning techniques can be used to boost the performance of various components inside the Oracle Database.

Qualifications

The successful candidate is expected to complete the internship using a wide and diverse set of skills.

Required Skills

- Thorough understanding of CS fundamentals including data structures, algorithms, **machine learning techniques, AI** and complexity analysis
- Good problem-solving skills
- **Good C / Python programming**
- Familiarity with **neural networks and sampling algorithms**
- Notions of database architecture and experience with system development is a big plus
- Familiarity with SQL is a big plus

For more information about the internship, please contact Vlad Haprian (vlad.haprian@oracle.com).

References

[1] http://www.alexbeutel.com/papers/CIDR2019_SageDB.pdf

[2] <https://dl.acm.org/doi/pdf/10.1145/3448016.3452838>

“Working at Oracle Labs is a great experience that I recommend to any student in computer science. It has been one of the best parts of my studies, as I was able to research, design and implement new features in a state-of-the-art distributed graph processing engine. Furthermore, the diversity of the people working at Oracle Labs, as well as its great working atmosphere, made my internship even more enjoyable.”

<Arnaud Delamare >

EPFL student, now employed by Oracle Labs, 6-months intern

Related Topics

Our group at Oracle Labs further offers various internship topics in the following areas:

- Automated Machine Learning with Explainability (AutoMLx)
- Automating OCA Verification of GitHub Pull Requests
- BPF Linux Schedulers
- Extending a Distributed Graph Engine (Oracle Labs PGX)
- Extending a Web-Based Enterprise Data Science Platform
- Graph Machine Learning at Oracle
- Graph Support in the Oracle Database
- Machine Learning and Data Analysis Techniques for Domain Global Graphs
- Machine Learning Processing in DB Systems
- Oracle Database Multilingual Engine - Modern Programming Languages in the Database

If you are interested in more than one of these areas, it is sufficient to apply once. In our interview process, we are going to take all your areas of interest into account.