

Intern with Oracle Labs AutoMLx!

The Automated Machine Learning with Explainability (AutoMLx) group at Oracle Labs has open internship positions available on Machine Learning development topics.

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. The resulting rapid proliferation of ML technology, explosive data growth, and shortage of data science expertise have caused the industry to face increasingly challenging demands to keep up with fast-paced develop-and-deploy model lifecycles. Recent academic and industrial research efforts have started to address this problem through automated machine learning (AutoML) pipelines and have focused on model performance as the first-order design objective. Oracle AutoML [1] is a novel iteration-free AutoML pipeline designed to not only provide accurate models, but also in a shorter runtime. These objectives are achieved by eliminating the need to continuously iterate over various pipeline configurations. Our approach to AutoML has been shown to achieve better scores at a fraction of the time compared to state-of-the-art, making it a prime candidate for addressing industry challenges.

Moreover, as there is an ever-increasing number of deployed models, automated explainability techniques are required to provide insight into their complex behavior and the data they work with. They help humans understand the reasons behind model predictions as well as aid in debugging, improve performance, identify bias and unfairness in the models, and potentially support compliance with new regulations such as "right to explanation". Our group is focused on building a novel framework that gives novice and advanced data scientists the power of AutoML and explainability in a single, easy-to-use, python package.

Featured internship topics: AutoML and/or explainability for classification, regression, anomaly detection, and forecasting tasks; Explore support for federated learning; Explore techniques to reduce model bias while tuning;

Extend dataset support for unstructured (e.g., NLP) and semi-structured (e.g., video/audio/graph) data; Generic model support including GNNs, DNNs and/or RNNs.

Required Skills

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

- Understanding of ML, including traditional and DL algorithms
- Experience with python, scikit-learn, Tensorflow and/or PyTorch
- Familiarity with anomaly detection or forecasting algorithms is a plus
- Excellent problem-solving and analytical skills
- A grade of at least 5.0 in the master/bachelor studies and at least 5.0 in a machine learning course is required.

This project is suitable for regular internship or M.Sc. Thesis

For more information, contact hesam.fathi.moghadam@oracle.com.

References

[1] Yakovlev, Anatoly, et al. "Oracle automl: a fast and predictive automl pipeline." Proceedings of the VLDB Endowment 13.12 (2020): 3166-3180.

Related Topics

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

- 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 for Optimizing Oracle Database Performance
- 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 will take all your areas of interest into account.

“Oracle Labs is a great place to perform meaningful research which can directly impact the industry. Projects are well supervised, and the people are always there to help you in order to end up with a successful project.”

Mikael Morales
Former EPFL student,
6-month intern