Oracle Machine Learning

News

Machine Learning in Oracle Database

Mark Hornick | December 5, 2019
Senior Director, Data Science and Machine Learning

In keeping with Oracle's mission to help people see data in new ways, discover insights, unlock endless possibilities, customers wishing to use the Machine Learning, Spatial, and Graph features of Oracle Database are no longer required to purchase additional licenses.

Large and small businesses alike know the value of applying machine learning technology to solve key business problems and achieve competitive, if not leadership, position in their industries. To democratize machine learning across the enterprise and further the ability of businesses to leverage powerful machine learning tools, Oracle Database now includes all functionality from the Oracle Advanced Analytics option as part of its base license the in-database Oracle Machine Learning APIs and functionality for SQL, R and, *coming soon*, Python. This applies to Oracle Database versions currently supporting Oracle Machine Learning, with future support to include Oracle Database SE2. See our broader announcement as well as specific Spatial and Graph blogs.

Move the algorithms, not the data!

With Oracle Machine Learning, Oracle moves the algorithms to the data. Oracle runs machine learning within the database, where the data reside. This approach minimizes or eliminates data movement, achieves scalability, preserves data security, and accelerates time-to-model deployment. Oracle delivers parallelized in-database implementations of machine learning algorithms and integration with the leading open source environments R and Python. Oracle Machine Learning delivers the performance, scalability, and automation required by enterprise-scale data science projects – both on-premises and in the Cloud.

Oracle Database - the multi-model converged database

Users shouldn't have to create and manage multiple databases to access different analytical functionality, which adds complexity and cost. Instead, all such functionality should exist in a single converged, multi-model database, bringing together a broad set of algorithms that can operate on data with various data types and data models. This is a key differentiator for Oracle Database, and reinforces Oracle's goal to provide such advanced development tools to the widest range of developers.

Empower enterprise users with SQL, R, and Python

Different business problems and their underlying data require different analytical techniques and algorithms to be successful. Oracle Database includes over 30 such algorithms. Oracle Machine Learning provides natural interfaces for the key data science languages: SQL, R, and Python. With the SQL API, in-database models are *first class* database objects with many of the same data management features available for other database objects like tables and views.

By supporting R and Python, data scientists and other R and Python users access Oracle Database as a high performance compute engine for scalable data exploration and preparation with native R and Python functions and syntax. The in-database algorithms come with a natural R and Python interface as well.

Production deployment - the critical step

Some might refer to *production deployment* as the Achilles' heal of data science projects. When application developers or IT try to integrate machine learning models or open source R and Python scripts in production, they are faced with the realities of addressing backup, recovery, security, and scalability concerns. Providing an integrated machine learning platform, like Oracle Machine Learning, deployment is immediate – SQL-derived models exist in the database and can be invoked from SQL queries, R and Python scripts can be stored in Oracle Database and executed in engines spawned and controlled by Oracle Database. Oracle greatly simplifies production deployment by providing the "plumbing" so enterprise teams can focus on the machine learning solution.

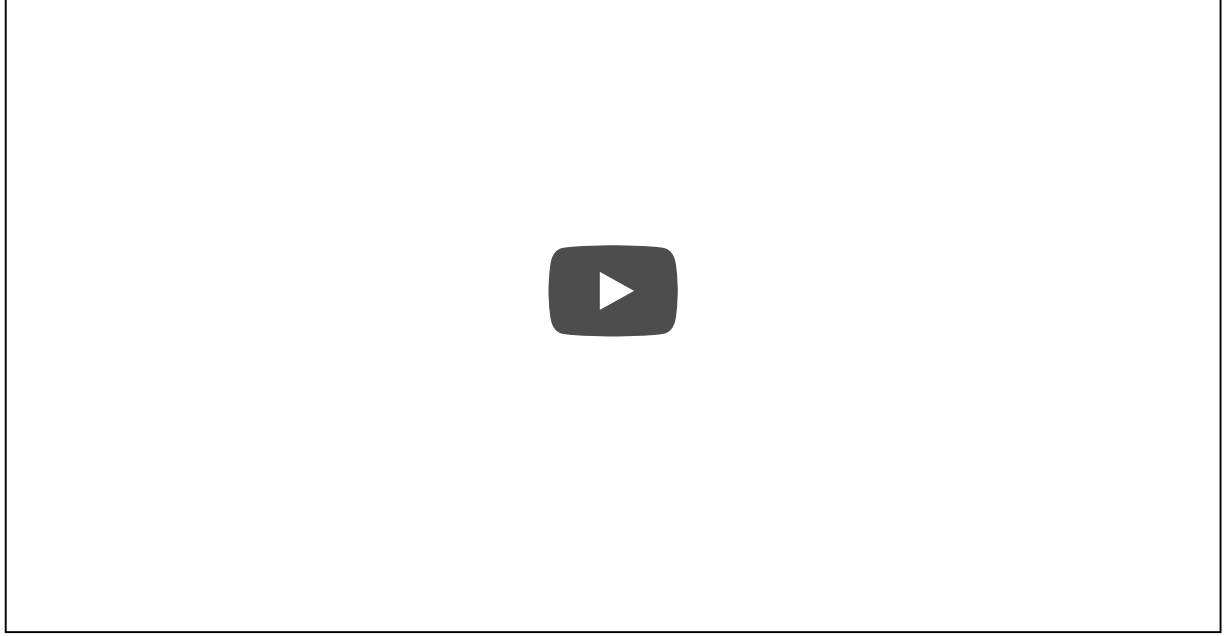
Machine learning can be applied to a wide range of business problems ranging from healthcare fraud detection to manufacturing root cause analysis, from retail product recommendation to equipment remaining useful life prediction. As enterprises amass greater data volumes and supplement corporate data with external data sources, the ability to integrate and prepare data at scale for machine learning is essential. Further, the ability to streamline the time from business problem definition to solution deployment is critical to deliver ROI on data science projects. Oracle Machine Learning customers have achieved impressive results, including:

StubHub, the world's largest ticket marketplace, uses Oracle Machine Learning in-database models and integrated R capabilities to run real-time fraud detection models in their database. With this solution, online fraud has been reduced by 90%. This not only saves money, but also significantly improves the customer experience. Of course, fraudsters notice when they get shut down and so change their way of operating, which means reacting quickly, or fraud goes up. And StubHub can do this very quickly. Updating their predictive model happens with the application and database still running. If they notice a problem in the morning, fix it in an hour, deploy immediately, and the new model can be operating before lunch with no downtime.

UK National Health Service (NHS), the system of public healthcare providers in the United Kingdom, were able to assemble and analyze billions of data points on prescriptions, medicines, medical exemptions, doctor relationships, and call center services from across the organization and use this to reveal potential new efficiencies and effective treatments to provide better outcomes. NHS achieved very fast ROI – 581million British pounds or $717million USD in the first 2 years – and now over 1.5billion British pounds or nearly $2billion USD within 5 years, and counting.

For more details about these and other customer stories see OML Customers.

Check out this video on Introduction to Oracle Machine Learning.



Features of Oracle Machine Learning

Here are a few highlights of Oracle Machine Learning functionality:

- Oracle integrates machine learning across the Oracle stack and the enterprise, fully leveraging Oracle Database and Oracle Autonomous Database
- Empowers data scientists, data analysts, developers, and DBAs/IT with machine learning
- Eliminates costly data movement and access latency
- Fast and scalable data exploration, data preparation, and machine learning algorithms
- Over 30 algorithms supporting: regression, classification, time series, clustering, feature extraction, anomaly detection
- R and Python integration directly supports data scientist role, leveraging open source ecosystems
- Ease of machine learning model and R/Python script deployment
- Automates key machine learning process steps, including AutoML for modeling
- Ease of machine learning model deployment through SQL and REST interfaces
- Leverages existing enterprise backup, recovery, and security mechanisms and protocols
- Brings the algorithms to the data

By including Oracle Machine Learning with Oracle Database on premises and in the Cloud, and with Oracle Autonomous Database, Oracle continues to support a next-generation converged data management and machine learning platform.

Mark Hornick

Senior Director, Data Science and Machine Learning

Mark Hornick is Senior Director of Product Management for the Oracle Machine Learning family of products. Mark has over 20 years of experience integrating and leveraging machine learning with Oracle technologies, working with internal and external customers in the application of Oracle's machine learning technologies for scalable and deployable data science projects. Mark is Oracle's representative on the R Consortium's Board of Directors and an Oracle advisor for the Analytics and Data Oracle User Community. He holds a bachelor's degree from Rutgers University and a master's degree from Brown University, both in computer science.

Previous Post

Configuring an ROracle Client with Autonomous Data Warehouse

Sherry LaMonica | 4 min read

Next Post

Join us at Analytics and Data Summit, February 25-27, 2020 in Santa Clara, CA

Charlie Berger | 2 min read