Introducing Oracle AutoML

It should come as no surprise then to see Oracle begin building advanced analytics capabilities into Oracle Database in support of data science workloads. For almost two decades the company has been working to build intelligence into its flagship database. In January 2021, Oracle took another sizable step in this direction with the introduction of Oracle Database 21c. Running on Oracle Cloud (and available on the company's free tier of Oracle Autonomous Database), this new release features a number of data-centric enhancements including blockchain tables, in-database JavaScript execution, in-memory graph analysis capabilities, and native binary JSON document support.

It also introduced what at first blush may appear as a strange addition to those not familiar with the evolution of Oracle's database, the inclusion of automated ML (AutoML). In fact, however, this addition directly builds on Oracle's existing efforts to support in-database ML algorithms, work that have been ongoing since the introduction of Oracle Database since 9i in 2001. With this release, for example, Oracle incorporated data mining and advanced analytics capabilities, featuring more than 50 statistical functions (regression, classification, even decision trees).

Popularized by Google in 2018 as a means of speeding up the task of selecting the most appropriate neural network to complete a given deep learning (DL) task, AutoML has grown into a more widely applicable means of automating a wide array of ML tasks including data preparation, model selection, feature selection and engineering, as well as hyperparameter tuning. Touted as a means of democratizing data science, AutoML allows non-technical domain experts to participate in the data science process. And true to its roots, it still allows data scientists to speed up and standardize repetitive tasks. These tools can be found within broader MLOps platforms as with Amazon SageMaker AutoPilot, as a stand-alone tool like H2O.ai AutoML, or as a pure Python framework like Auto Sklearn.

With Oracle Database 21c, Oracle enters this space with an AutoML solution that seeks to support both data scientists and non-technical domain experts. Available as a feature of OML4Py, which is itself a component included with Oracle Database and Oracle Autonomous Database, the company's new AutoML feature focuses on three main tasks: algorithm selection, feature selection, and hyperparameter tuning. And like most AutoML tools, it endeavors to both improve data scientist productivity and enable non-experts to participate in data science.