Whitepaper

Enabling the AI-driven enterprise with DataRobot on Dell EMC infrastructure

On-premises data science platform reference architecture

Abstract

Dell Technologies, DataRobot® and Intel® deliver an on-premises enterprise artificial intelligence (AI) solution stack to enable organizations to implement AI transformation projects on trusted infrastructure.

February 2020

Contents

Executive summary 3
Solution overview 3
DataRobot machine learning platform software powered by Dell EMC infrastructure 3
Reference architecture and implementation 6
Value of DataRobot enterprise ML platform on Dell EMC infrastructure 14
Governance and transparency 14
Putting models into production 14
Summary 16
Learn more 16

Executive summary

Interest in machine learning (ML) and AI continues to rise, thanks to record-breaking algorithmic advances. As the desire to adopt AI grows within organizations and translates into actionable business value, it creates a shortage of data science talent. This leads to increased adoption of automated solutions like the DataRobot enterprise AI platform. An enterprise/business AI platform can augment existing domain experts in the company to create, deploy and manage AI projects within existing business processes and data pipelines.

This document provides an architecture and implementation of DataRobot on Dell EMC infrastructure that can be leveraged by organizations to implement an on-premises enterprise AI platform. This enterprise AI platform can integrate with existing big data and data lake platforms, or run independently as a standalone multi-user environment, and ingest data from multiple sources. This tighter integration with the organization's existing data sources enables faster ML implementations and efficient use within the organization by multiple teams.

Solution overview

datarobot.com

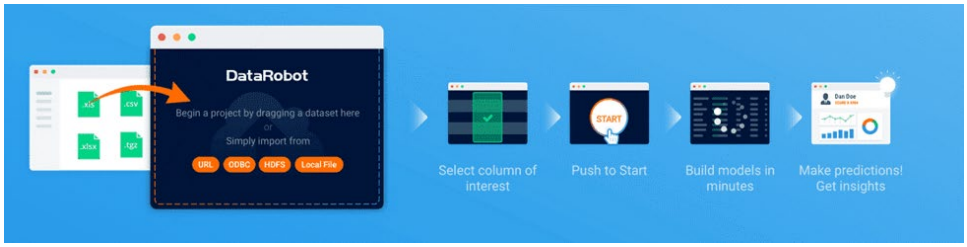
DataRobot is a leader in enterprise AI, delivering trusted AI software and AI enablement services to global enterprises competing in today's intelligence revolution. DataRobot enterprise AI software helps democratize data science with end-to-end automation for building, deploying and managing ML models. DataRobot, Dell Technologies and Intel have collaborated to design a scalable reference architecture to meet the needs of organizations that want to adopt or accelerate the use of AI to improve business value and processes. Armed with DataRobot on Dell EMC infrastructure, "citizen data scientists" gain the power to create advanced machine learning models without learning to code or understand when and how to apply certain algorithms. Data scientists are also more productive as repetitive steps in the model-building process are automated, allowing them to use their unique expertise for selecting and fine-tuning models.

DataRobot machine learning software powered by Dell EMC infrastructure

DataRobot software maximizes business value by delivering AI at scale and continuously optimizing performance over time. The company's proven combination of cutting-edge software and world class AI implementation, training and support services empowers any organization — regardless of size, industry or resources — to drive better business outcomes with AI.

Dell EMC PowerEdge servers are the bedrock of the data center. The PowerEdge server portfolio offers flexible designs for optimized application performance. The single-socket server portfolio provides balanced performance and storage capacity for future growth. The two-socket server portfolio brings a mix of features to maximize performance, scale for meet future demands, and adapt to virtually any workload with an optimum balance of compute and memory. The Dell EMC four-socket server portfolio fills the top end with the highest performance and extensive scalability for your applications from in-memory database workloads and high performance computing (HPC) to data analytics, AI and GPU database acceleration.

Intel® offers libraries and frameworks optimized for Intel Xeon Scalable processors. These include TensorFlow™, MXNet, PaddlePaddle, Caffe and PyTorch®. They enhance DL performance using software optimizations. The Intel distribution for Python, for example, accelerates AI-related Python libraries such as NumPy, SciPy, and scikit-learn with integrated Intel performance libraries such as Intel MKL to deliver faster AI training and inferencing. Data Robot integrates these libraries and frameworks, enabling enterprises to quickly deploy a reliable, powerful, comprehensive AI development software platform.



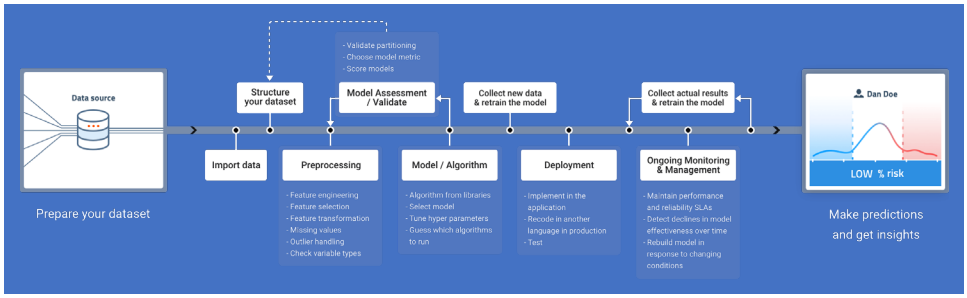
Source: datarobot.com

As a comprehensive solution, DataRobot adds value throughout the critical phases of developing and deploying ML models.

- **Identify data.** The AI catalog inside DataRobot serves as a centralized source of truth for data engineers, data stewards, data scientists and analysts to gain self-service access to AI assets they can trust.
- **Ingest data.** DataRobot transforms structured and unstructured data into the specific format each algorithm needs for optimal performance and follows best practices for data partitioning.
- **Engineer features.** DataRobot engineers new features from existing numeric, categorical and text features. It knows which algorithms benefit from extra feature engineering, and which don't, and only generates features that make sense given the data characteristics.
- **Explore algorithms.** No single algorithm can solve every business problem or data set. DataRobot provides access to hundreds of diverse algorithms and the appropriate pre-processing for users to test against their data to find the best one for their AI challenge.
- **Select algorithms.** DataRobot helps users select the algorithms that make sense for their data.
- **Train and tune the ML model.** DataRobot trains models on the user's data and uses smart hyper-parameter tuning to tune the most important hyper-parameters for each algorithm.
- **Find optimal algorithm combinations.** Ensemble or blender models typically outperform individual algorithms. DataRobot finds the optimal algorithms to blend, and tunes the weighting of the algorithms within each ensemble model.
- **Compare models head-to-head.** DataRobot builds and trains dozens of models, compares the results and ranks the models by accuracy, speed and the most efficient combination of the two—regardless of which programming language or machine learning library the models came from. Users can explore them with the intuitive graphical user interface (GUI) and choose which ones to move forward with.
- **Build trust.** To ensure transparency, DataRobot explains model decisions in a human-interpretable manner, showing which features have the greatest impact on the accuracy of each model and the patterns fitted for each feature. DataRobot also provides prediction explanations to illustrate the key reasons why a specific prediction was made.
- **Deploy production-ready models**. DataRobot produces production-ready models that users can operationalize rapidly and integrate with enterprise applications with just a few lines of code, whether for real-time predictions, batch deployments, scoring on Apache® Hadoop®, or other methods. Users can also develop their own models using R, Python®, Apache Spark®, MLlib, H2O®, and other tools and call the DataRobot library to activate them.

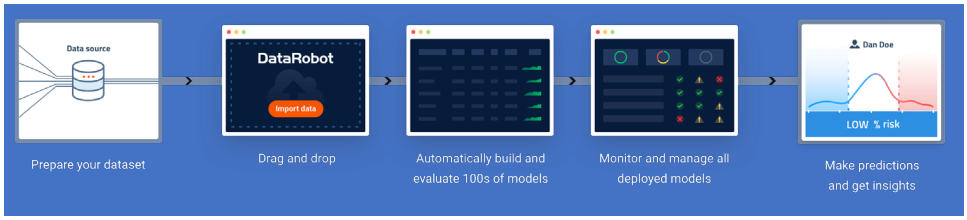
- **Monitor and manage.** Post-deployment, DataRobot makes it easy to compare predictions to actual results and train a new model on the latest data. DataRobot proactively highlights if a model's performance is deteriorating over time.

Before DataRobot, expert data scientists needed to perform many time-consuming and tedious data science processes:



Source: datarobot.com

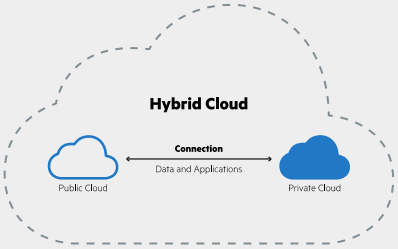
With DataRobot, guardrails and other automated best practices for data science allow citizen data scientists to build AI they can trust, and expert data scientists to increase their productivity.



Source: datarobot.com

DataRobot integrates easily within the ecosystem of technologies already in the enterprise. These include security and data privacy technologies, data integration and visualization tools, and infrastructure platforms, such as Hadoop and SQL databases. Structured and unstructured data can be ingested from data lakes, tables and other enterprise sources. Users interact with the system via graphical or programmatic interfaces.

AI workloads rely on data to operate and need lots of it from different sources both within and outside the organization. It usually makes more sense to keep the application co-resident with the data sets.



2nd-generation Intel Xeon Scalable processors

The 2nd-generation Intel Xeon Scalable processors are optimized for demanding data center workloads. This processor family features higher frequencies than previous generation Intel Xeon Scalable processors, along with architecture improvements and AI and deep learning (DL) inference workload enhancements.

The 2nd-generation Intel Xeon Scalable processors take AI performance to the next level with Intel DL Boost, which extends the Intel Advanced Vector Extensions 512 (Intel AVX-512) instruction set with Vector Neural Network Instructions (VNNI). Intel DL Boost significantly accelerates inference performance for DL workloads optimized to use VNNI—sometimes by as much as 30X, compared to a previous-generation Intel Xeon Scalable processor.

DataRobot software includes three independent but fully integrated products:

- **Automated Machine Learning** incorporates a variety of regression techniques, from simple linear regression to statistical classic regression models to more complex techniques, such as gradient boosting and neural networks. The software also solves simple binary classification problems and complex, multiclass problems with up to 100 categories.
- **Automated Time Series** automates the complex process of developing sophisticated models to predict the future values of a data series based on its history and trend. The platform integrates time series feature engineering to discover predictive signals. It uses both basic and advanced time series models to optimize forecasting accuracy, and offers many ways to visualize insights over time, and to deploy models to production.
- **MLOps** provides a centralized hub to deploy, manage and govern machine learning models in production to maximize investments in data science teams and to manage risk and regulatory compliance. This includes not only models built using the automated machine learning capabilities of DataRobot, but also models built using teams of expert data scientists with platforms and frameworks such as Python and R.

Reference architecture and implementation

Many organizations need their AI and analytics infrastructure on-premises to access pre-existing data sets. These data sets are too large and spread out within multiple organizations to be conveniently or cost-effectively transferred to and from the cloud. Additionally, some organizations have the capability to acquire and operate compute resources at much cheaper rates than cloud providers offer.

A recent Moor Insights publication reports on the benefits and limits of cloud hosting for AI and HPC. It states that "starting with the cloud may make a lot of sense if an organization wants to experiment with AI" due to the services that are available and the ease of acquiring the required compute resources.[1]

It continues, "However, many organizations will eventually need significant computing infrastructure for AI and HPC as their applications begin to run at scale. This, along with data transfer and throughput fees, begins to tip the cost balance in favor of building on-premises infrastructure as the organization matures in AI."

In addition to cost, other important factors that organizations have to consider are "data gravity" and security/privacy concerns. AI workloads rely on data to operate and need substantial amounts of data from different sources both within and outside the organization. It usually makes more sense to keep the application co-resident with the data sets. When it comes to security, AI projects that involve massive proprietary data sets and security concerns can override the potential ease of use, especially in financial and healthcare markets.

In addition to a managed offering on different public cloud providers, the DataRobot software platform supports multiple on-premises environments, from single node Linux® systems to large Hadoop environments. Within these environments, DataRobot services are deployed as Docker® containers.

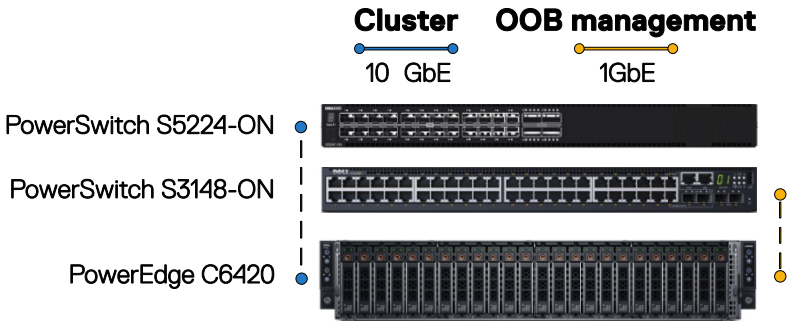
We tested multiple installations ranging from a single Linux server running all the DataRobot services to a resilient multi-node virtualized environment that provides flexibility to carve out resource requirements custom to each service for optimal performance and predictable response times.

[1] Moor Insights and Strategy, "AI and HPC: Cloud or On-Premises Hosting." February 2019.

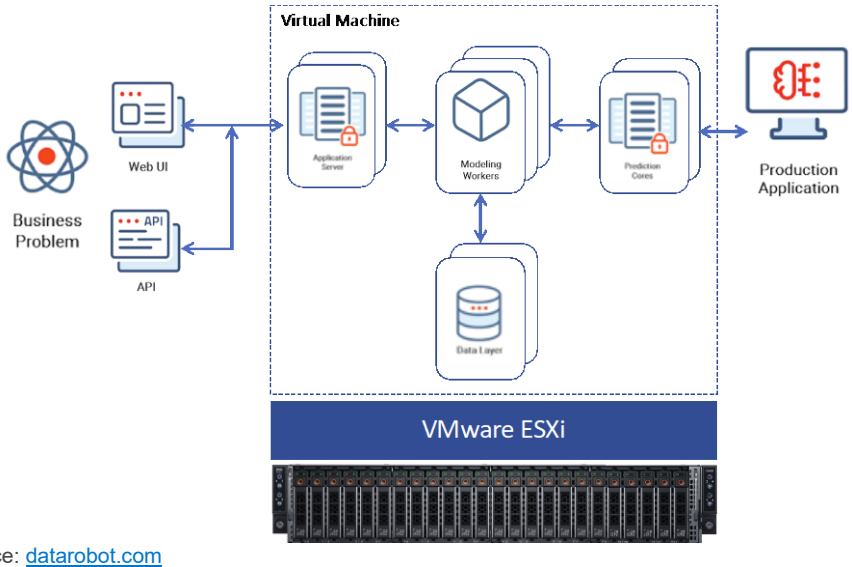
In another instance, the computational workload and object storage services can be distributed across a Hadoop cluster. Hadoop deployments give organizations the flexibility of installing DataRobot in an already provisioned Hadoop cluster, allowing hardware costs savings and a direct path to their data.

The reference architecture below shows the hardware infrastructure to support multiple concurrent users using a standalone Linux deployment. We deployed Linux virtual machines (VMs) in a VMware® ESX® environment and deployed each individual DataRobot service in its own VM.

To support concurrent users, we require a total of four Dell EMC PowerEdge servers to meet the application, compute and storage needs of the DataRobot software platform. Although multiple servers from the Dell Technologies portfolio are well suited for this workload, we tested and highlight the Dell EMC PowerEdge C6420 Server in this paper, with four 2-socket servers in a 2U chassis. Other two-socket servers, such as PowerEdge R640 and R740XD, may be substituted, keeping the CPU core counts and memory capacity consistent with the recommendations provided in this document.



The key services in a DataRobot environment are the application server, data layer, modeling workers and prediction servers. The following diagram shows a high-level flow.



Source: datarobot.com

The **application server** houses all of the main administrative components. It handles authentication, project management and user administration and provides an endpoint for the DataRobot public API. It also manages the queue of modeling requests made by various projects, which are executed by **modeling workers** running on modeling nodes. A single modeling node can host multiple modeling workers (portions of processing power for performing tasks), and a single DataRobot system can have multiple modeling nodes.

When you purchase DataRobot, you get access to a set number of concurrent modeling workers. In general, one modeling worker can train one model or generate one additional insight option (feature impact and prediction explanations, for example).

Modeling workers operate in parallel, allowing users to simultaneously create multiple models and generate additional insights. This helps to minimize the time required to train all the models associated with a specific project.

DataRobot software is architected to ensure the amount of data a customer can possibly use to train a model is limited only by the physical hardware resources allocated. Modeling workers require varying amounts of machine resources so they can operate and complete their tasks successfully. **Base modeling workers** are assigned a specific CPU and memory limit (typically 4 CPU cores and 30Gb RAM) that allows them to build models with training data sets up to 1.5Gb in size (uncompressed on disk), and ten classes for multi-classification applications.

Flexible modeling workers, in contrast, can scale from four CPU cores up to 20, and utilize much larger memory allocations, to process significantly larger training data sets and up to 100 classes for multi-classification applications. DataRobot dynamically adds computing power and memory as data size or problem complexity increases. The capacity that a flexible modeling worker can scale to is limited by the amount of available resources on the modeling node.

Modeling workers are also stateless, so they can be configured to join and leave the environment on demand. This can save on hardware costs when configured with a virtual private cloud (VPC). Within a Hadoop cluster, these workers are YARN containers. Trained models are written back into the **data layer**, and their accuracy is reflected on the model leaderboard through the application server.

Predictions can be generated in batch by utilizing the **web UI** to upload data, or via an **API** endpoint. In both cases, a modeling worker is temporarily used to generate predictions. For low-latency and high-capacity prediction environments, **prediction cores** are recommended. Prediction cores are reserved for making predictions to avoid contention with modeling activities.

This system of reserving prediction cores effectively ensures that prediction resources are available when a particular user wants to generate predictions and that modeling users don’t have to wait while modeling resources are consumed for predictions. Furthermore, key statistics about those predictions and the data provided are returned back to the application server and displayed to users for monitoring the health of the models, including an analysis of data drift and accuracy.

Intel DC solid-state drives (SSDs)

Intel DC SSDs are optimized for Intel Xeon Scalable processors and are available in a variety of media types and capacities. For example, Intel 3D NAND SSDs (such as the Intel SSD DC P4500, P4501, and P4600 Series) are optimized for cloud infrastructures, offering outstanding quality, reliability, advanced manageability, and serviceability to minimize service disruptions.

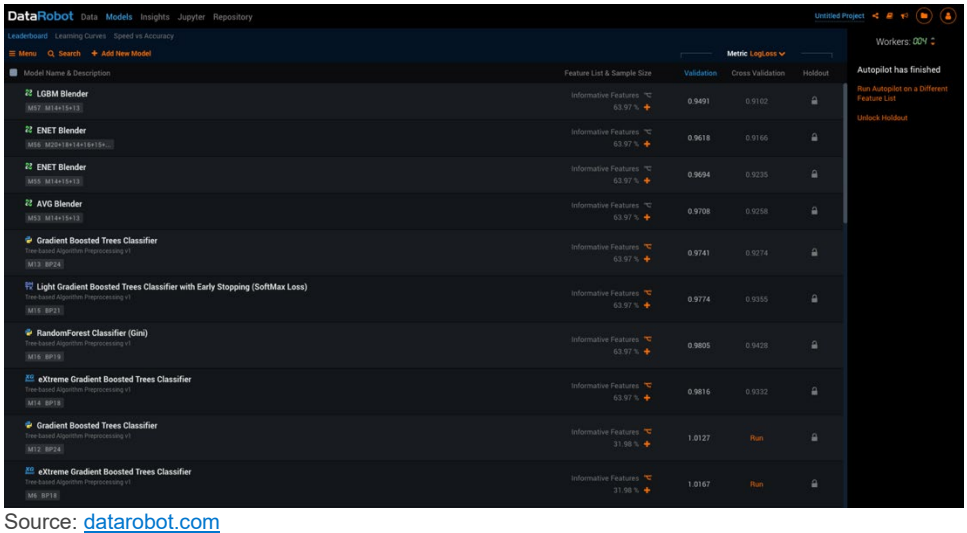
The Intel Optane™ SSD DC P4800X accelerates applications for fast caching and fast storage to increase scale per server and reduce transaction costs for latency-sensitive workloads. In addition, the Intel Optane SSD DC P4800X enables data centers to deploy larger and more affordable data sets to gain new insights from large memory pools.

Prediction cores can also be deployed in an environment disconnected from DataRobot, allowing enterprises to deploy models in segregated networks. The recommended configuration to support 30, 20 and 12 modeling workers are shown in the table below:

	30 modeling workers (large)	20 modeling workers (medium)	12 modeling workers (base)
Servers	4x PowerEdge C6420	3x PowerEdge C6420	2x PowerEdge C6420
Processor	Intel® Xeon® Gold 6248		
Memory	384GB DDR-2933		
Storage	2x 480GB BOSS (boot), 6 x Dell Express Flash NVMe P4610 1.6TB SFF		
Networking	Intel Ethernet 10G 4P X710 SFP+ rNDC		

The reference architecture was setup in the Dell EMC PowerEdge test data center, and multiple use cases were tested to determine infrastructure requirements when running a diverse workload. We used the automated machine learning features and capability of the DataRobot platform to execute the different use cases. We selected five public data sets as our use case examples and DataRobot was leveraged to ingest data, engineer features, explore and select algorithms for the selected data, and tune and train multiple ML models. As part of the model training phase, DataRobot uses smart hyper-parameter tuning to tune the most important hyper parameters for each algorithm. These phases are computationally intensive and use a scale-out technique leveraging multiple modeling workers in parallel and intelligent selection of algorithms to minimize the training times.

DataRobot presents a leaderboard that shows the top-performing models, and they can be explored to compare the accuracy scores, validation accuracy and additional model details, such as input variables used. The most apt model for production may not be the most accurate model as there are other considerations — such as prediction speed, and interpretability — that determine suitability for deployment. DataRobot provides tools, reports and plots to compare models for evaluation that make it easy for a user to determine which model to pick from the leaderboard.



Source: datarobot.com

Use cases evaluated on DataRobot

Fraud detection. Consumer fraud (such as credit card fraud, loan fraud and bank fail) can be difficult to perceive due to outdated rules or systems. DataRobot can be used to detect and prevent fraud by, in the case of credit card fraud, modeling previous credit card transactions. This model identifies whether a recent transaction is fraudulent. The goal is to detect 100% of the fraudulent transactions and to reduce the classifications of incorrect fraud.

Data set: Predicting Fraud in Financial Payment Services

Customer churn prediction. Customer churn is a major impediment to organizational growth. DataRobot can develop predictive models to predict customer churn and create actionable alerts that can be fed to customer retention (banking sector) and marketing departments.

Data set: Bank Customer Churn Prediction

Risk rating in insurance companies. Multiple risks are associated with insurance companies, such as mortality rates, morbidity rates and catastrophic risks. With the help of DataRobot, these risks are evaluated and rated. This platform helps the insurance companies measure risk to calculate a recommend premium.

Data set: Prudential Life Insurance Assessment

Higgs boson detection. The boson particle was only found due to modern improvements to machine learning and is currently a common benchmarking data set for determining how well a system can handle extreme numbers of parameters in the data set.

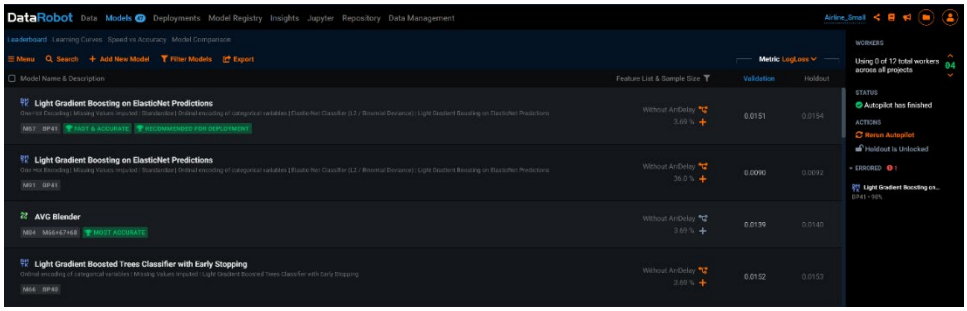
Data set: HIGGS Data Set

Airline delay prediction. Predicting airline delay is an advantage for travelers and operators. DataRobot empowers customers to choose alternate flight routes or at least choose the option of cancelling or not buying airline ticket in the first place if there is high risk delay. Operators can analyze their organization's performance and their competitors' performance when they face a delay.

Data set URL: Year 2008

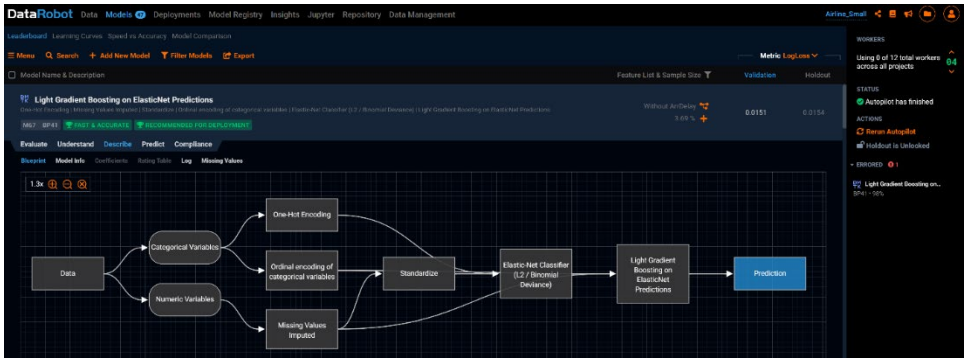
Source: progressive.in

The leaderboard view showing multiple models trained to predict the airline delays is shown below.



Source: datarobot.com

Learn more about the model recommended for deployment, in this case, the LGB model, by viewing the blueprint.



Source: datarobot.com

Integration with big data Hadoop environments

DataRobot makes a powerful addition to a Hadoop-based data analytics environment. When deploying to a Cloudera® Hadoop environment, software integration built into DataRobot simplifies deployment, monitoring and management. Benefits include:[2]

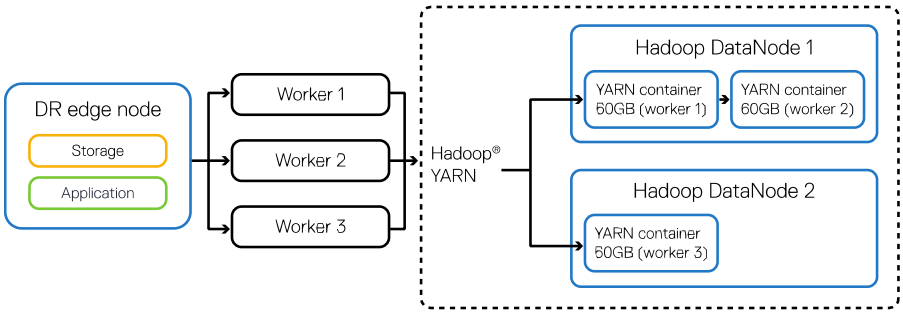
- **Installs via Parcels.** Allows DataRobot to deploy new and updated releases using Cloudera Manager with no downtime
- **CSD integrated.** Allows Cloudera Manager to monitor resources used by DataRobot
- **Kerberos.** LDAP/AD support (authentication)
- **Sentry enabled.** Authorization/Governance/Compliance
- **YARN enabled.** Resource manager for multi-tenant environments
- **Uses Spark/MLlib.** Flexible, in-memory data processing for modeling tasks

Through these integrations, DataRobot users can leverage existing data lakes that represent years of historical data stored in Hadoop Distributed File System (HDFS) and make them available for automated model training. The Apache Spark integrations for model training and inference allow scaling of DataRobot-based analytics workloads in the same Apache Spark and Hadoop environment data operations teams have been managing for years.

[2] DataRobot, DataRobot Brings Automated Data Science to Hadoop. March 2016.

DataRobot integrates with Hadoop as an edge node to the cluster. In this configuration, YARN is responsible for resource management and Apache Spark executors are leveraged for modeling tasks. When a worker node receives a model training operation from DataRobot, the required memory is allocated by YARN on a Hadoop DataNode and model training begins.

DataRobot modeling on Hadoop

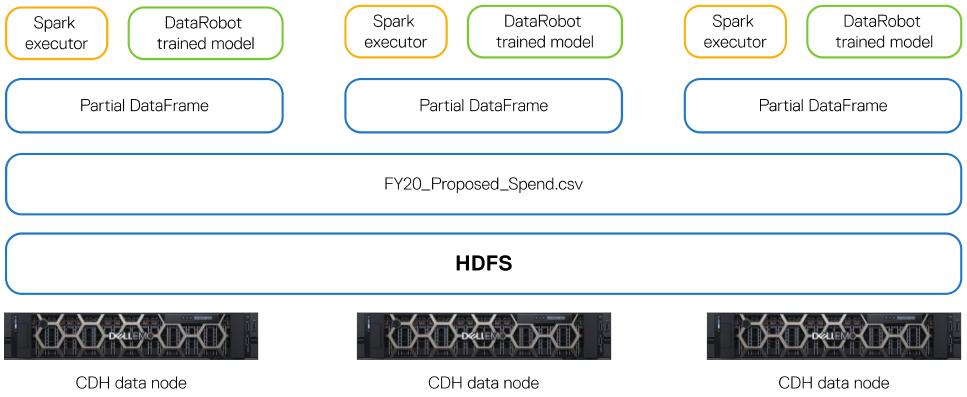


- Yarn allocates memory on a data node when a worker wants to train a model.
- Each model is trained in memory on an available data node

Source: Transforming Insurance Analytics with Big Data and Automated Machine Learning

In this architecture, each YARN container is responsible for training a single model candidate and DataRobot, acting as the Spark Driver, orchestrates the selection process. This allows DataRobot to use algorithms and machine learning libraries that are not designed for distributed training or available in a Spark compatible format. Each Spark Executor executes against the local node resources for computation and does not need to share data with other executors like in a distributed training operation of a single model.[3]

When using Apache Spark for modeling tasks, the inference or model scoring operations behave differently. Unlike the model candidate selection process done during training, the inferencing operation benefits significantly from the in-memory distributed Apache Spark data structures. During inference, each worker uses an identical copy of the trained model. Each scores a partition of the DataFrame, which is a distributed representation of the data set. The results for each partition are sent to DataRobot for reporting.



[3] Dell Technologies, Training an AI Radiologist with Distributed Deep Learning. August 2018.

The testing of DataRobot used the following configuration of Dell EMC PowerEdge C6420 Servers following the design and best practices outlined in the architecture guide for the Dell EMC Ready Solution for Cloudera Hadoop.[4]

System Element	DataRobot Application Node	Cloudera Worker Nodes	Cloudera Management Node
Server Model	1x Dell EMC PowerEdge C6420	2x Dell EMC PowerEdge C6420	1x Dell EMC PowerEdge C6420
CPU	Intel Xeon Gold 6230 CPU	Intel Xeon Gold 6240 CPU	Intel Xeon Gold 6230 CPU
DRAM Memory	192GB	384GB	192GB
Capacity Storage	3.8TB SATA SSD	3.8TB SATA SSD	3.8TB SATA SSD
Network	1 — Embedded Intel Gigabit I350-t LOM 2 — Intel Ethernet 10G 2P X520 Adapter	1 — Embedded Intel Gigabit I350-t LOM 2 — Intel Ethernet 10G 2P X520 Adapter	1 — Embedded Intel Gigabit I350-t LOM 2 — Intel Ethernet 10G 2P X520 Adapter
Software (Pre-requisite)	CentOS® Linux 7.6.1810 Docker services	Cloudera Hadoop 5.16.1 HDFS DataNode YARN Node Manager Spark Gateway Hive Gateway	Cloudera Manager 5.16.1 YARN Resource Manager HDFS NameNode Zookeeper Spark History Server
DataRobot Software	DataRobot 5.2.2-RELEASE Application		DataRobot 5.2.2-RELEASE Parcel and services Master service ETL controller ETL default and ETL quick worker services

Role
DataRobot Application Node — This node will be used to install DataRobot application, and maintain the core Docker services, such as Nginx, Mongo, Redis, RabbitMQ and Hadoop config sync.
Cloudera Management Node — The management node is be used to install the Cloudera manager, HDFS Name Node, YARN Resource Manager.
Cloudera Worker Node — The worker node will be used to offload the modeling tasks managed by YARN scheduler.

[4] Dell Ready Bundle for Cloudera Hadoop Architecture Guide version 5.10

Value of DataRobot enterprise ML platform on Dell EMC infrastructure

DataRobot enterprise AI software democratizes data science and automates the end-to-end process for building, deploying and maintaining AI at scale. Powered by the latest open-source algorithms and available in the cloud, on-premises or as a fully managed AI service, DataRobot gives users the power of AI to drive better business outcomes.

- **AI with ROI.** DataRobot is developed with a relentless focus on delivering value and success for organizations that want to be AI-driven. It accelerates AI use case throughput by increasing the productivity of data scientists and empowering non–data scientists to build, deploy and maintain AI without having to learn traditional data science methods. Instead of spending weeks or months developing and testing a few hand-coded machine learning models, existing teams (regardless of their data science expertise) can build hundreds of models and deploy the best performing model in hours.
- **AI you can trust.** DataRobot is committed to delivering reliable AI that always works as planned, so users have the confidence to make significant business decisions. Dedicated development teams are focused on providing an intuitive AI experience, so users can easily understand DataRobot predictions and forecasts and explain them to others regardless of their level of data science skills. Built-in guardrails, automated model documentation and other capabilities ensure human-centric AI that consistently reflects an organization's unique values and ethics.
- **AI you own.** AI infrastructure has the potential to be your most strategic asset, and you should own every aspect of it. DataRobot gives organizations the flexibility to use the platform how and where they want it, with multiple options for building and managing their AI. Your AI intellectual property remains yours without vendor lock-in.

Governance and transparency

Many approaches to AI and machine learning are done in a "black box," with limited details about how a model was trained and why it made the predictions it did. DataRobot is different — there's an intense focus on transparency throughout the model building, deployment and management process.

AI creators can view details about every step in the model building process, including how data was processed, what features were engineered, and what algorithms were used. AI operators can automatically generate model documentation to understand how a model was trained, as well as easily identify the data that has the greatest impact on model results. This process also allows them to apply business rules and logic to identify any potential for bias or other issues that can compromise the integrity of a model. And, AI consumers get human-friendly explanations for every prediction made by the model, including the top factors that contributed to the decision that was made.

Putting models into production

AI and machine learning projects should be driving the future of your business. But many companies can't take full advantage of data science investments. These organizations lack the skills and resources to deploy, manage and govern machine learning in production. Moreover, the results of attempting to use AI models without proper monitoring and controls can be catastrophic, including lost revenues and the trust of business executives, investors and customers.

Every model built in DataRobot is immediately ready for deployment.

- **Upload** a new data set to DataRobot to be scored in batch and downloaded.
- **Create** a REST API endpoint to score data directly from applications. An independent prediction server is available to support low-latency, high-throughput prediction requirements.
- **Export** the model for in-place scoring in Hadoop.
- **Download** scoring code, either as editable source code or self-contained executables, to embed directly in applications to speed up computationally intensive operations.

Once deployed, DataRobot MLOps ensures that the machine learning models driving your business are accurate and consistent throughout changing market conditions. At a glance you can view a summary of metrics from all models in production, including the number of requests (predictions) and key health statistics:

- **Service health** looks at core performance metrics from an operations or engineering perspective: latency, throughput, errors and usage.
- **Data drift** proactively looks for changes in the data characteristics over time to let you know if there are trends that could impact model reliability.
- **Accuracy** compares actual values (or ground truth) corresponding to our predictions so you can assess model performance using standard machine learning metrics.

DataRobot provides the ability to frequently update models; test new, competitive models; and change applications on the fly while continuing to serve business applications. Enforce governance policies related to ML deployment and capture the data that is required for strong governance practices, including who is publishing models, why changes are being made, and what models were in place over time.

Benefits of Dell Technologies, DataRobot and Intel Collaboration

- Access and test-drive the DataRobot AI platform in Dell Customer Solution Centers and consult with solution experts who can provide guidance on deployment or integration into existing IT environments.
- Take advantage of a simplified ordering process for DataRobot on Dell EMC infrastructure.
- Receive collaborative support capability through Dell and DataRobot.

Summary

We presented an on-premises implementation of the DataRobot platform on Dell EMC infrastructure that can be leveraged by organizations to implement an enterprise AI platform in their data centers. This enterprise AI platform can integrate with existing big data and data lake platforms or run independently as a standalone multi-user environment.

Three different standalone configurations were validated to accommodate the different modeling and prediction capabilities required by your data science teams. The integration of DataRobot with Hadoop enables leveraging of Hadoop compute resources for DataRobot machine learning and the use of existing data stored in HDFS.

This white paper and reference architecture is a collaborative effort between DataRobot, Intel and Dell Technologies to enable organizations to:

- Stand up an on-premises solution faster than if they try to do it on their own (determine optimal configs, sizing, etc.).
- Deploy with confidence as engineers have tested stability and interoperability.
- Realize cost savings as hardware infrastructure is right-sized for different use cases and training/inference capability.

Organizations that are using Dell EMC Ready Architectures for Hadoop can integrate DataRobot and extend their big data analytical capabilities.

Learn more

- DataRobot
- Delltechnologies.com/ai
- Delltechnologies.com /referencearchitectures
- Delltechnologies.com/hpc
- Hpcatdell.com
- Delltechnologies.com/servers
- @dellemcservers
- Intel and AI
- Intel SSD Data Center Family
- Intel Deep Learning Boost
- Intel Framework Optimizations
- Intel Deep Learning Reference Stack
- Intel Builders
- #IntelBuilders

DataRobot

DELL Technologies

intel