The DataRobot automated machine learning platform captures the knowledge, experience, and best practices of the world’s leading data scientists to deliver unmatched levels of automation and ease-of-use for machine learning initiatives. DataRobot enables users of all skill levels, from business people to analysts to data scientists, to build and deploy highly-accurate predictive models in a fraction of the time of traditional modeling methods.

— Andrew Greenhut
Senior Data Scientist
LogMeIn

— Dave Truzinzki
Chief Digital Officer
Crest Financial

Human-centric design, ideal for business users
Domain expertise is the key to developing effective predictive models. DataRobot enables business users to generate accurate models quickly and perform sophisticated data science functions directly.

Built-in guardrails
With DataRobot, modeling projects follow a consistent methodology based on best practices so users can’t “forget” to perform a critical step, such as model validation.

Speeds model evaluation
DataRobot builds a leaderboard so you can see which models perform best with your data, and provides the tools you need to explore and compare individual models.

Advanced machine learning techniques
The DataRobot platform incorporates the techniques advanced data scientists use: boosting, bagging, random forests, kernel-based methods, generalized linear models, deep learning, and many others.

Unsupervised anomaly detection
Uncover anomalies in a dataset with DataRobot’s unsupervised ensemble blend model, which can offer new insights even in familiar datasets.

Builds the workflow for you
DataRobot creates the predictive modeling workflow for you. It knows what to do at each step of the process, and does it automatically, without prior programming or manual input from users.

Automates feature engineering
DataRobot prepares data automatically, performing operations like one-hot encoding, missing imputation, text mining, and standardization to transform features for optimal results.

Leverages innovative open source engines
To harness the most advanced techniques, DataRobot uses open source machine learning libraries like R, scikit-learn, TensorFlow, Vowpal Wabbit, Spark ML, and XGBoost.

Supports advanced tuning
DataRobot automates model tuning, but also supports manual tuning so data scientists can tune and adjust machine learning algorithms for even better results.

Multiclass model support
DataRobot allows for classification on targets with up to 10 distinct values, offering both real-time and batch support for uncovering the predictive class and showing its probability across all classes.
Deployment Options

DataRobot Cloud
• SaaS offering powered by Amazon Web Services (AWS)

DataRobot Enterprise Hadoop
• Cloudera
• Hortonworks

DataRobot Enterprise Linux
• Virtual Machines
• Virtual Private Cloud – AWS, Google Cloud, Microsoft Azure, and more

Built for the Enterprise

Operating at enterprise scale requires blazing performance, strict adherence to controls, and relentless focus on data protection. DataRobot is an enterprise-ready platform, delivering the governance, training, and world-class support your organization needs to quickly get up and running.

Use on-premise or in the cloud
On-premise: You can deploy DataRobot on-premise on standalone servers, an existing Hadoop infrastructure, or in a Virtual Private Cloud (VPC).
Cloud: The DataRobot Cloud is hosted on Amazon Web Services (AWS), delivering the flexibility and speed necessary for any enterprise.

Leverages distributed processing
DataRobot leverages modern distributed processing, running experiments in parallel to radically reduce the time it takes to run a complete data science project.

Enables rapid collaboration
With DataRobot, business users, data scientists, and stakeholders work together on machine learning projects to deliver better results with less wasted effort.

Eliminates model deployment bottlenecks
There are multiple options for deploying your finished models with DataRobot, including native scoring, exportable prediction code, and prediction APIs for real-time and batch scoring.

Resource monitoring and reporting
DataRobot’s Resource Monitor feature provides a view of platform usage across the organization, including which workers and models are taking up runtime, enabling effective resource planning.

Integrates with Hadoop
DataRobot uses your Hadoop distribution’s application management services to distribute runtime libraries to Hadoop Data nodes. Working directly with HDFS, running predictions in DataRobot does not require a proprietary storage layer or the movement of data to an edge node. The DataRobot workload runs in YARN containers, so you do not need to partition your cluster to prevent resource conflicts.

Works with enterprise data
No matter where your data resides – relational databases, Hadoop clusters, text files, or other sources – DataRobot quickly and easily connects to your data source.

Explainable models
Users can download DataRobot’s diagnostic charts, data, and documentation to share them with executives, stakeholders, and regulators.

Supports advanced security
DataRobot offers native security for fine-grained role-based authorization and supports Kerberos and LDAP protocols. In Hadoop, it works with your existing encryption policies.

Editable rating tables
Customizable rating tables allow users to edit and manipulate the tables according to their unique business rules, allowing for an optimal blend of machine learning and human experience.

Sample Industries

DataRobot has customers and use cases in a myriad of industries, including:

- Marketing
- Retail
- Manufacturing
- Banking
- Fintech
- Insurance
- Public Sector
- Healthcare
- Non-profit
- Internet of Things (IoT)
- Sports
- Transportation
- Hospitality
- Energy

Contact Us

DataRobot | One International Place, 5th Floor | Boston, MA 02110
www.datarobot.com | info@datarobot.com

© 2017 DataRobot, Inc. All rights reserved. DataRobot and the DataRobot logo are trademarks of DataRobot, Inc. All other marks are trademarks or registered trademarks of their respective holders.