DataRobot helps insurers and other industries leverage the power of machine learning, providing a platform for users of all skill levels to produce accurate predictions in a fraction of the time required by conventional tools and methods. Leveraging the knowledge, experience, and best practices of the world’s leading data scientists, DataRobot automates the machine learning process, from data to deployment.

“DataRobot solves predictive modeling challenges across all core functions in an insurance company - marketing, underwriting, pricing, claims, audit, and operations. That by itself is extremely powerful, but what really differentiates DataRobot is its user-centric design. DataRobot was founded and built by insurance industry veterans, and we know that successful adoption of machine learning means engaging the users, regulators, and consumers.”

– Satadru Sengupta
DataRobot GM Insurance

**Insurance Use Cases**
For more information about DataRobot use cases in insurance, visit our Insurance Solutions page.

- Claim Development Modeling
- Claim Payment Automation Modeling
- Conversion Modeling
- Direct Marketing
- Fraudulent Claim Modeling
- Insurance Pricing
- Life Insurance Underwriting for Impaired Life Customers

**Ideal for business users**
Insurance domain expertise is the key to developing effective predictive models. DataRobot enables business users to generate accurate models quickly and efficiently, performing 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 advanced data science techniques: bagging, boosting, deep learning, frequency severity methods, generalized additive models, generalized linear models, kernel-based methods, random forests, and many others.

**Builds the workflow for you**
DataRobot creates the predictive modelling 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**
Data is prepared automatically, with DataRobot 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 from R, scikit-learn, TensorFlow, Vowpal Wabbit, and XGBoost.

**Supports advanced tuning**
DataRobot automates model tuning, but supports manual tuning so data scientists can tune and adjust machine learning algorithms for even better results.
YARN container size is 4 cores and 60GB.

The following table shows the recommended minimum configuration for installing DataRobot on a Linux server.

<table>
<thead>
<tr>
<th>System</th>
<th>Processor</th>
<th>Desktop/Server Intel Xeon E5-2699 v3</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU Speed (GHz)</td>
<td>Cores per CPU</td>
<td>2.3</td>
</tr>
<tr>
<td>Total Cores</td>
<td>Total RAM (GB)</td>
<td>18</td>
</tr>
<tr>
<td>Storage</td>
<td>256</td>
<td></td>
</tr>
<tr>
<td>4x 1TB SSD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The following table shows an equivalent minimum configuration for installing DataRobot using DataRobot. The minimum configuration required for a Hadoop services, e.g. Microsoft Azure, can be used.

<table>
<thead>
<tr>
<th>Instance Count</th>
<th>Instance Type</th>
<th>Total Cores</th>
<th>Total Memory (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>r3.large</td>
<td>4</td>
<td>30.5</td>
</tr>
<tr>
<td>1</td>
<td>c3.large</td>
<td>4</td>
<td>7.5</td>
</tr>
<tr>
<td>3</td>
<td>r3.2xlarge</td>
<td>24</td>
<td>183</td>
</tr>
<tr>
<td>1</td>
<td>r3.xlarge</td>
<td>4</td>
<td>30.5</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>36</td>
<td>251.5</td>
</tr>
</tbody>
</table>

The following table shows the minimum configuration required for a Hadoop deployment using DataRobot. The minimum YARN container size is 4 cores and 60GB. An additional 20GB on each data node is required by Cloudera for each data node.

<table>
<thead>
<tr>
<th>Count</th>
<th>Edge/Data Node</th>
<th>Min Cores</th>
<th>Min Memory (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Edge</td>
<td>8</td>
<td>128</td>
</tr>
<tr>
<td>7</td>
<td>Data Node</td>
<td>28</td>
<td>560</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>36</td>
<td>688</td>
</tr>
</tbody>
</table>

The minimum configuration for installing DataRobot, including native scoring, exportable prediction code, and prediction APIs for real-time and batch scoring.

**Focus on Actuaries and Regulation-Friendly Design**

Actuaries have their own requirements for predictive analytics and machine learning, and DataRobot incorporates the insurance industry domain expertise of some of its top data scientists into the platform to deliver a number of valuable features for actuaries.

**Advanced insurance modeling**

Frequency-Severity Modeling is the expected insurance pricing method used by European actuaries, and Generalized Additive Models (GAM) and Generalized Additive² Models (GAM²) provide the intelligibility and transparency needed by actuaries.

**Ideal for actuaries**

DataRobot supports advanced machine learning techniques for actuaries, including Offsets and Exposure to support operational constraints and practices, and Tweedie Loss metrics to increase robustness and accuracy.

**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 get up-and-running quickly.

**Use on-premise or in the cloud**

**On-premise:** You can deploy DataRobot on-premise on standalone servers, existing Hadoop infrastructure, or in a Virtual Private Cloud (VPC).

**Cloud:** DataRobot offers a variety of cloud services to meet your needs, including the DataRobot Cloud or installation on the cloud vendor of your choice.

**Leverages distributed processing**

DataRobot leverages modern distributed processing, running experiments in parallel to radically reduce the time needed 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.

**Validated rating tables**

Validated rating tables are available for download from GAM and GA2M models to validate, obtain insight, or generate predictions, including the strength of signal captured by main effects and interactions and pairwise interactions found by the model.

**Insights for actuaries**

Model X-Ray visualization improvements include adjusted charts to incorporate the use of exposures and step functions that result from the Generalized Additive Models to provide more insightful information for actuaries.

**Integrates with Hadoop**

DataRobot uses your Hadoop distribution’s application management services to distribute runtime libraries to Hadoop Data nodes. Working directly with HDFS, DataRobot does not require a proprietary storage layer or moving 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. DataRobot meets the requirements of the CIA’s C2S Cloud.

**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.