The power of prediction

The mission of the data scientist in the federal government is among the most critical; without the ability to make use of agencies' massive amounts of data, timely and accurate decisions are much harder to make. DataRobot allows you to focus on doing more by helping you solve your most pressing problems in less time.

Validate accuracy
Run datasets simultaneously through thousands of proven models in seconds to improve the accuracy of your outcomes .

Reduce model delivery time
Make numerous date-driven decisions and implement your models in days.

Clearly communicate predictions and outcomes
Demonstrate outcomes to leadership and key stakeholders, even showing different results as variables change.

Increase cross-team collaboration
Collaborate across teams to build and test custom algorithms and workflows.

DataRobot can support your agency's top challenges, including Fraud Detection and Prevention, Insider Threat, Scientific Research, Emergency and Disaster Response, Predictive Maintenance

How DataRobot can help federal data scientist

Expand your toolkit
Keeping up with the growing ecosystem of machine learning algorithms has never been this easy. DataRobot's broad set of diverse, best-in-class algorithms from R, Python, H20, Spark, TensorFlow, and other sources is continually expanding and at your disposal. With one line of code or a single click, DataRobot trains, tests, and compares hundreds of different models, including techniques you may never have used.

Massively parallel
DataRobot uses hundreds, or even thousands, of servers — as well as multiple cores within each server — to parallelize data exploration, model building, and hyper-parameter tuning. Large datasets? No problem. DataRobot uses distributed algorithms from Spark and H20 to spread the computation across your servers.

Model optimization
Within every project, DataRobot automatically identifies the best pre-processing and feature engineering for each modeling technique. Employing text mining, variable type detection, encoding, imputation, scaling, transformation, and automated feature engineering (to name a few), DataRobot scientifically selects the hyper-parameters and options that optimize out-of- sample performance (for example, AUC, LogLoss, Gini, etc.) for your models. Of course, DataRobot also provides you with options to select your tuning parameters and do custom feature engineering.

Easy model deployment
With DataRobot's API, you can operationalize your models with just a few lines of code, regardless of whether you need real-time predictions, batch scoring, or scoring on Hadoop. Update your models with no downtime and no need to write new scoring code...ever.

Secure data science for even the most sensitive projects

DataRobot understands that data-driven challenges in the government involve sensitive data at all classification levels. Our solution was built from the ground up with security requirements like FedRamp and FISMA in mind. We offer two secure deployment models:

On-site deployment

DataRobot is deployable on premises in a secure data center in freestanding clusters or in Hadoop. DataRobot integrates natively with Cloudera, leveraging Cloudera Manager for easy deployment and low maintenance as well as a partnership with HortonWorks.

Secure cloud deployment through GovCloud and C2S

DataRobot meets the requirements of the CIA's C2S cloud. DataRobot is also available on AWS GovCloud, which has currently received a Provisional Authority to Operate (P-ATO) under the FedRAMP High baseline.

Improving missions outcome with DataRobot

DataRobot distills the knowledge of the world's top data scientists into its automated machine learning software platform, empowering users of all skill levels to make predictions with greater accuracy, in a fraction of the time traditionally needed. Whether your agency has a team of data scientists or just a few analysts, DataRobot can help your data-driven programs tackle more challenges and deliver greater impact.

End-to-end data science

By automating the machine learning process from data to deployment, DataRobot allows you to deploy and scale your organization's data science efforts quickly. DataRobot builds on algorithms from R, Python, H2O, Apache Spark and XGBoost and leverages open source data processing, storage, and resource management frameworks such as Apache Hadoop.

DataRobot automated...

DataRobot reads and profiles data from a wide variety of sources, automatically correcting issues along the way.

Learning
DataRobot learns from your data using different techtiques, validates what it learns with experimental data, and displays the results so you can choose the best predictive model for your data.

Data preparation
DataRobot uses information about your data and your business problem to select precisely the right mix of machine learning algorithms to test.

Feature Engineering
DataRobot automatically preprocesses data for best results with the technique to be tested.

Deployment
DataRobot offers a real-time prediction engine and a batch prediction engine for Hadoop. Prediction engines run independently of the learning engine, with self-service API's, so your data scientists can continue working on current projects as new predictions are needed.

DataRobot can support you today!

We can leverage your datasets to run a comprehensive proof of concept and provide you with the tools you need, starting day one. To see a video of DataRobot in action, visit www.datarobot.com

DataRobot
for Federal Data Scientists

1775 Tysons Boulevard Tysons, VA 22102 | public-sector@datarobot.com