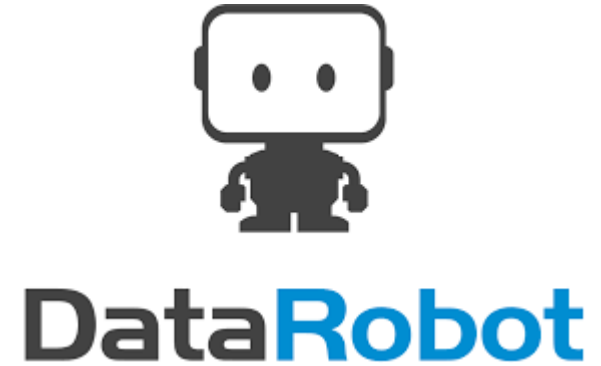


# Automatic Accurate Predictions with Data Robot

Jarlath Quinn – Analytics Consultant

[www.sv-europe.com](http://www.sv-europe.com)

A SELECT INTERNATIONAL COMPANY



Just waiting for all attendees to join...

# Automatic Accurate Predictions with Data Robot

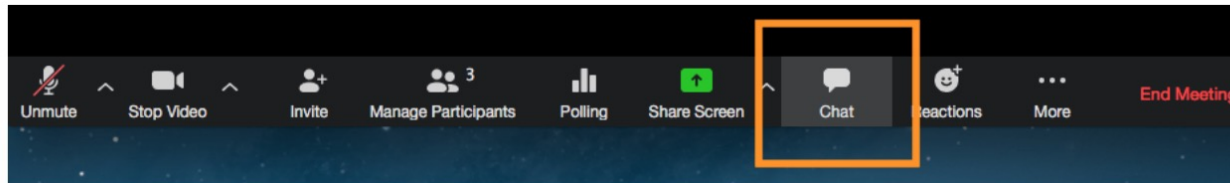
Jarlath Quinn – Analytics Consultant

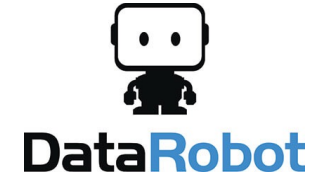
[www.sv-europe.com](http://www.sv-europe.com)

A SELECT INTERNATIONAL COMPANY

# FAQ's

- Is this session being recorded? Yes
- Can I get a copy of the slides? Yes, we'll email links to download materials after the session has ended.
- Can we arrange a re-run for colleagues? Yes, just ask us.
- How can I ask questions? All lines are muted so please use the chat panel – if we run out of time we will follow up with you.





- Gold accredited partner to IBM, Predictive Solutions and DataRobot specialising in advanced analytics & big data technologies
- Work with open source technologies (R, Python, Spark etc.)
- Team each has 15 to 30 years of experience working in the advanced and predictive analytics industry
- Deep experience of applied advanced analytics applications across sectors
  - Retail
  - Gaming
  - Utilities
  - Insurance
  - Telecommunications
  - Media
  - FMCG





- A mature, enterprise-ready **AI lifecycle** platform that supports:
  - Integration with Generative and Predictive AI
  - A wide portfolio of data formats and data preparation capabilities
  - Automatic feature-engineering
  - Automatic model-building using multiple DR and 3<sup>rd</sup> Party algorithms
  - Deep model evaluation functionality
  - Model transparency via blueprints and detailed documentation
  - Intelligent use of multimodal data e.g. text, image and geospatial

# Automatically detects and processes multi modal data



DataRobot Use Cases AI Catalog Data Models Deployments Model Registry Applications ending club 2012-2013 historic

Repository Leaderboard Insights Learning Curves Speed vs Accuracy Model Comparison

**Insight** Word Cloud Displays the most relevant words and short phrases in word cloud format.

**Model** Regularized Logistic Regression (L2) M59 BP62 | 64% Sample Size | Informative Features

21 models available

**USMC**

Coefficient: 0.2242

Appears in 15 rows

Filter stop words

[Export](#)

You can drag and zoom in on the chart to get a better look.

DataRobot Use Cases Data Models Deployments Model Registry Insights Repository Applications AI Catalog Ch...e

Menu Search Feature List: All Features View Raw Data

Feature Name Data Quality Index

price (Numeric)	i	61
-----------------	---	----

Histogram Frequent Values Table Geospatial Map

Map Legend

- Hexagon color by Avg
  - 308 to 266
  - 266 to 225
  - 225 to 184
  - 184 to 143
  - 143 to 102
  - 102 to 61
  - 61 to 20
- Hexagon height by Count

**Light Gradient Boosted Trees Regressor with Early Stopping**

Optimal encoding of categorical variables | Missing Values Imputed | Single Column Converter | SparseNet Image Pretrained Featurizer | Ridge Regressor | Light Gradient Boosted Trees Regressor with Early Stopping

M64 BP52 64.6% **RECOMMENDED FOR DEPLOYMENT**

Evaluate Understand Describe Predict Compliance

Feature Impact Feature Effects Prediction Explanations Word Cloud Image Embeddings Activation Maps

Filter by predicted All values Filter by actual All values Show color overlay

Insight is based on 30 images from validator set



- A mature, enterprise-ready **AI lifecycle** platform that supports:
  - Code-first approaches via Jupyter-compatible DataRobot Notebooks
  - Single point for model deployment in multiple modes
  - Champion-challenger and re-training policies
  - Model governance options including monitoring, service help, accuracy, drift as well as fairness/bias
  - Work-bench interface where different personas can collaborate



# Two modes to get started: Workbench

## Use Case directory


+ Create Use Case

LAST MODIFIED USE CASES



### Lending Club





Jarlath Quinn  
Modified 48 min ago

 1  1  0  0

ALL USE CASES 1

Search

Settings

Name ↑	Datasets	Experiments	Apps	Notebooks	Created by	Last modified	Members	
 Lending Club	1	1	0	0	 Jarlath Quinn	48m		





# Two modes to get started: Classic

DataRobot Value Tracker AI Catalog Notebooks Data Models 28 Deployments DataRobot NextGen OLD Lending Club (Deployed) ? 2

Begin a project by dragging a dataset here

or

simply import from:

Data Source URL Local file

or Browse

AI Catalog

Accepted formats: .csv, .tsv, .dsv, .xls, .xlsx, .sas7bdat, .geojson, .gz, .bz2, .tar, .tgz, .zip



# Scenario: Predicting Loan Default



# LENDING CLUB

DataRobot



Let's take a look...

# Deployed Models



DataRobot Deployments Model Registry Applications AI Catalog

Deployments List Prediction Jobs

### Active Deployments

9

0 Inactive

### Predictions

1.44M (+6.6%)

Last 7 Days

### Service Health Summary



9 Passing  
0 At Risk  
0 Failing

### Data Drift Summary



9 Passing  
0 At Risk  
0 Failing

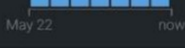
### Accuracy Summary



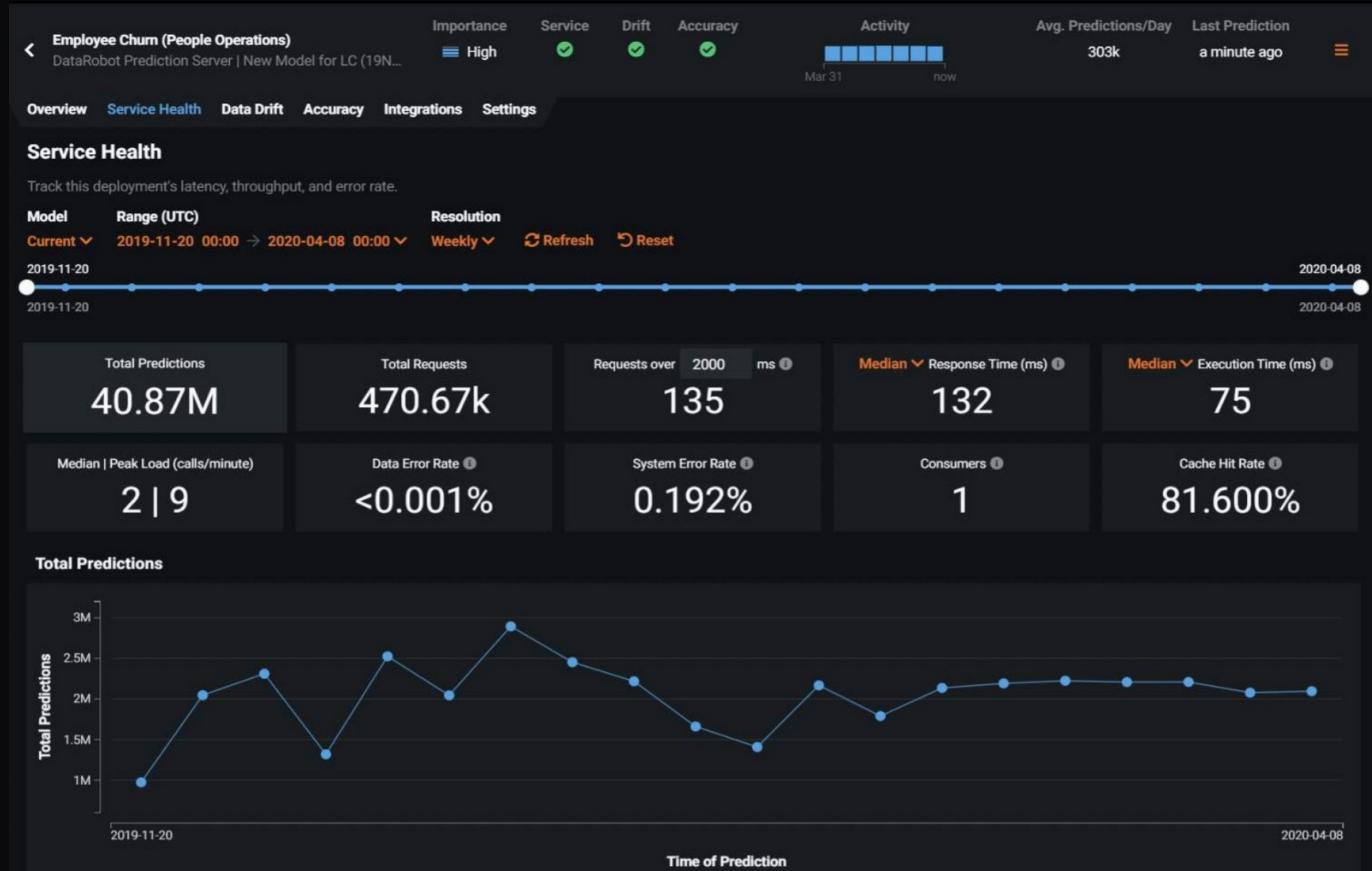
9 Passing  
0 At Risk  
0 Failing

Search + Add Deployment Filters (0)

Columns Prediction Health Rows 1-9 of 9

Deployment Name	Service	Drift	Accuracy ↑	Activity	Avg. Predictions/Day	Last Prediction
<b>MMM BostonMon Twitter messages polarity monitoring in Bosto</b> DataRobot Prediction Server   newdata.csv <b>NEEDS APPROVAL</b>	✓	✓	✓	 May 22 now	28k	a few seconds ago
<b>Demo Probability of Default (Risk Management)</b> DataRobot Prediction Server   New Model for LC (19NOV) Project	✓	✓	✓	 May 22 now	22k	a minute ago
<b>Employee Candidate Score (People Operations)</b> DataRobot Prediction Server   New Model for LC (19NOV) Project	✓	✓	✓	 May 22 now	22k	a few seconds ago
<b>Interest Score (Marketing)</b> DataRobot Prediction Server   New Model for LC (19NOV) Project	✓	✓	✓	 May 22 now	22k	a few seconds ago
<b>Compute Optimization (CloudOps)</b> DataRobot Prediction Server   New Model for LC (19NOV) Project	✓	✓	✓	 May 22 now	22k	a few seconds ago

# Deployed Models: Service Health



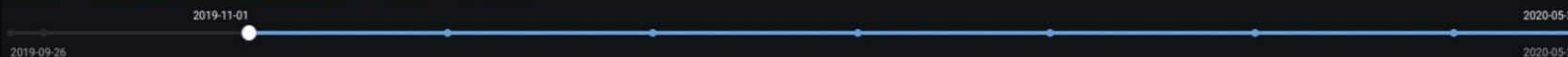
# Deployed Models: Accuracy



## Accuracy

Assess the accuracy of this deployment's predictions over time.

Model: Previous | Range (UTC): 2019-09-26 00:00 → 2020-05-20 00:00 | Resolution: Monthly | Refresh



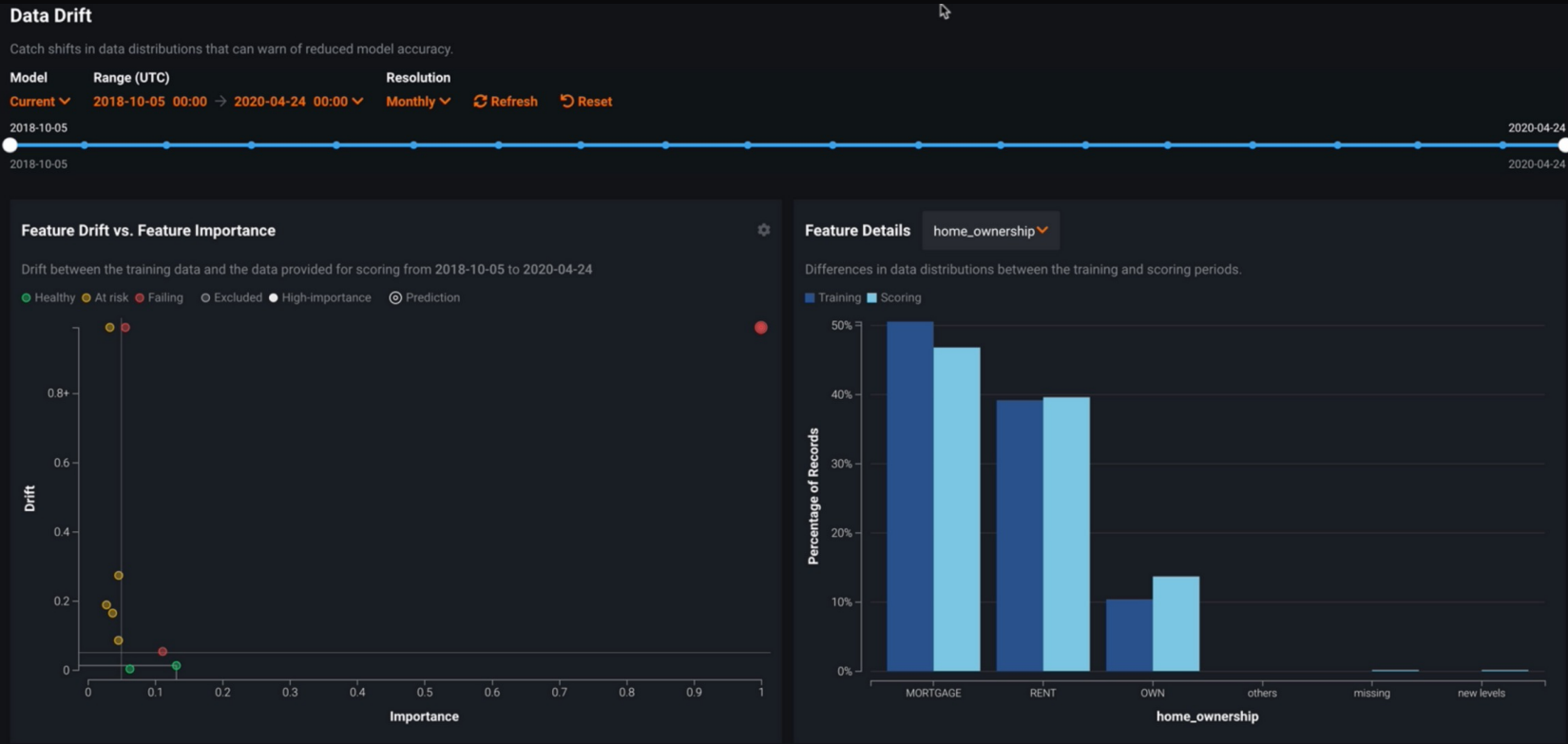
## Accuracy over Time



## Predicted & Actual



# Deployed Models: Data Drift





# Working with Smart Vision Europe

- **Consulting Services**

## **Project Support**

Purchase 1-2 days of consultancy time to have an expert work alongside you on your own project

## **Analytics Advice**

Give us 3-5 days to investigate your data & analytical strategy and we'll present our recommendations re: improvements & alternatives

## **Analytical Deep-Dive**

Let us explore your data landscape to test hypotheses, identify problem areas, find key outcome drivers or develop new applications

# Working with Smart Vision Europe Ltd.

- **Sourcing Software**
  - You can buy your analytical software from us often with discounts
  - Assist with selection, pilot, implementation & support of analytical tools
  - <http://www.sv-europe.com/buy-spss-online/>
- **Training and Consulting Services**
  - Guided consulting & training to develop in house skills
  - Delivery of classroom training courses / side by side training support
  - Identification & recruitment of analytical skills into your organisation
- **Advice and Support**
  - offer 'no strings attached' technical and business advice relating to analytical activities
  - Technical support services



Contact us:

+44 (0)207 786 3568

[info@sv-europe.com](mailto:info@sv-europe.com)

Twitter: @sveurope



[Follow us on Linked In](#)



[Sign up for our Newsletter](#)

Thank you



Contact us:

+44 (0)207 786 3568

[info@sv-europe.com](mailto:info@sv-europe.com)

Twitter: @sveurope

[Follow us on Linked In](#)



Thank you