

Hands-on Lab: Explore a Simple Generative Tool

Estimated time needed: 30 minutes

Overview

Generative AI models have revolutionized how you interact with technology, enabling you to create new content, generate realistic images, and translate languages with remarkable accuracy.

In this lab, you will gain hands-on experience with a simple generative AI tool, DataRobot, exploring its capabilities and applications.

Learning Objectives

After completing this lab, you will be able to:

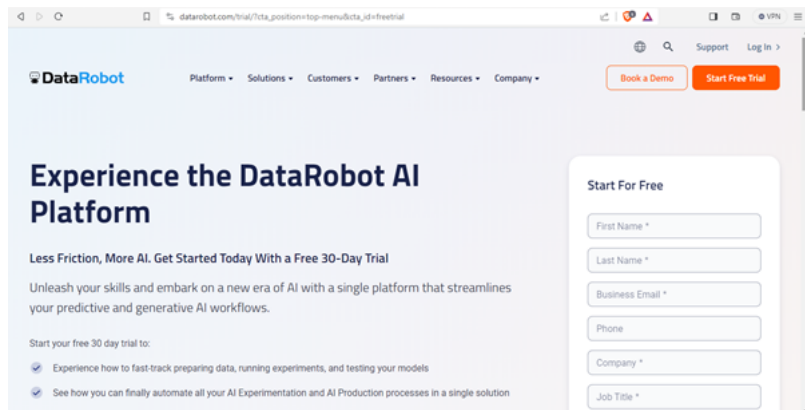
- Signup in DataRobot
- Add a data set to the use case
- Work on model building

Task 1: Sign-up in DataRobot

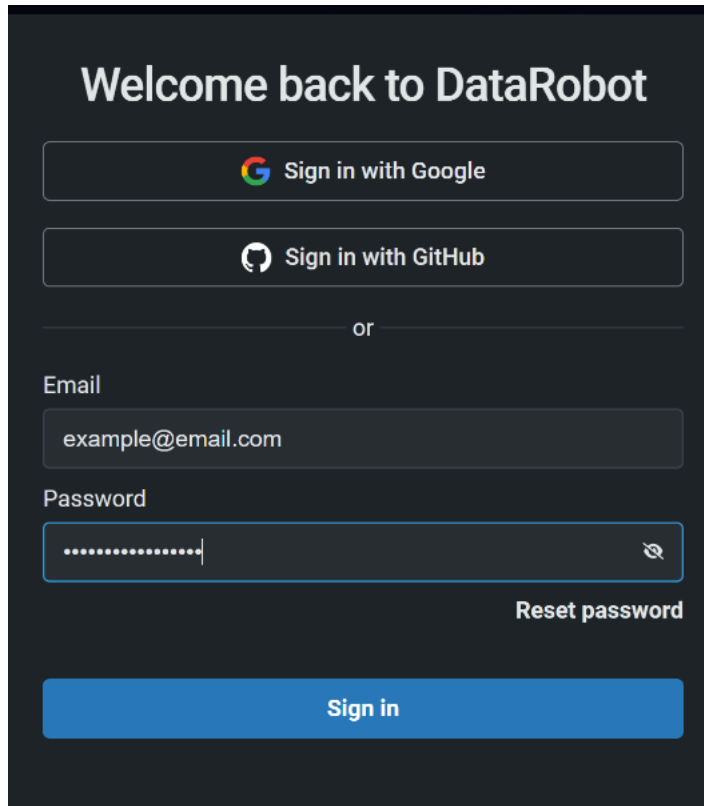
Step 1: Click www.datarobot.com

Step 2: Click Start Free Trial.

Step 3: Fill in the required information and create an account.

The screenshot shows the DataRobot website's landing page. The header includes the DataRobot logo and navigation links: Platform, Solutions, Customers, Partners, Resources, and Company. There are two orange buttons: 'Book a Demo' and 'Start Free Trial'. The main content area features the heading 'Experience the DataRobot AI Platform' and a subheading 'Less Friction, More AI. Get Started Today With a Free 30-Day Trial'. Below this, there is a paragraph about unleashing skills and a list of benefits for the free trial. On the right side, there is a 'Start For Free' form with input fields for First Name, Last Name, Business Email, Phone, Company, and Job Title, each followed by an asterisk indicating it is a required field.

Step 4: A new window will open; select the relevant option for signing up.



Welcome back to DataRobot

Sign in with Google

Sign in with GitHub

or

Email

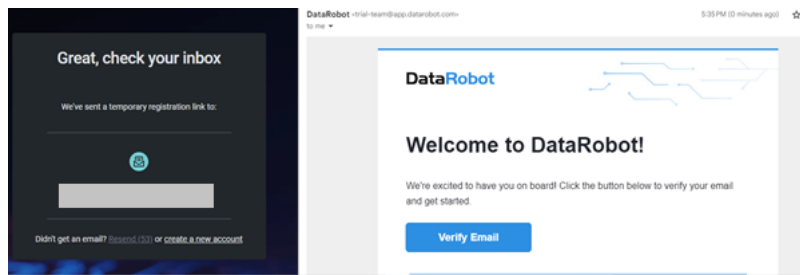
example@email.com

Password

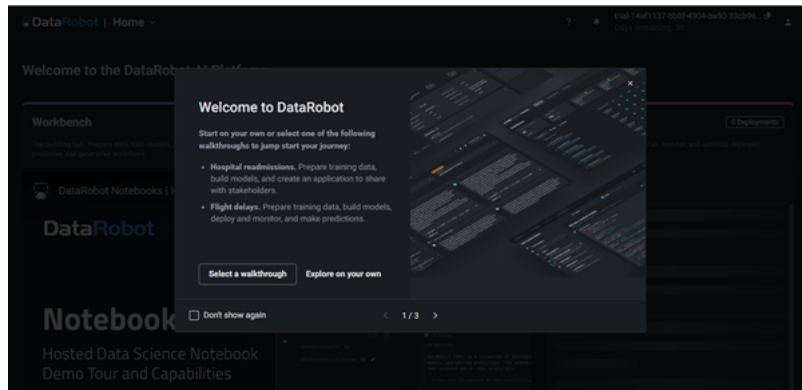
Reset password

Sign in

Step 5: Confirm your email by clicking **Verify Email** in your inbox.

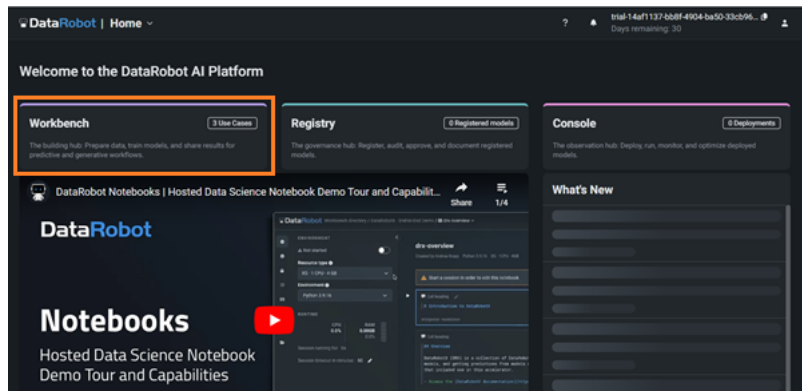


Step 6: Sign up and start your first experience of using the Generative AI tool.
The dashboard will look like the image below. You may like to familiarize yourself with the application by clicking **Select a walkthrough**.

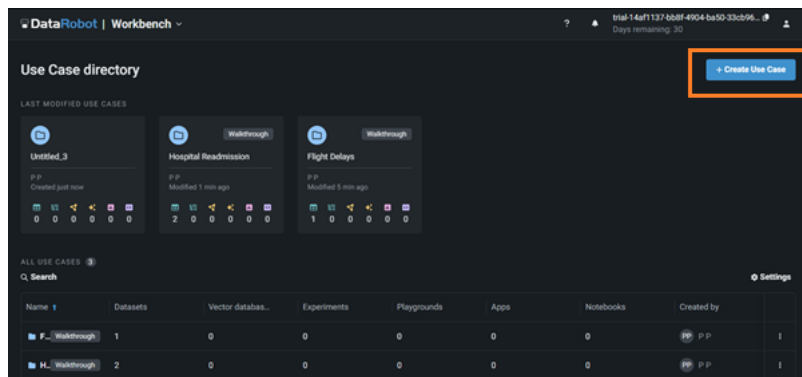


Task 2: Add a data set

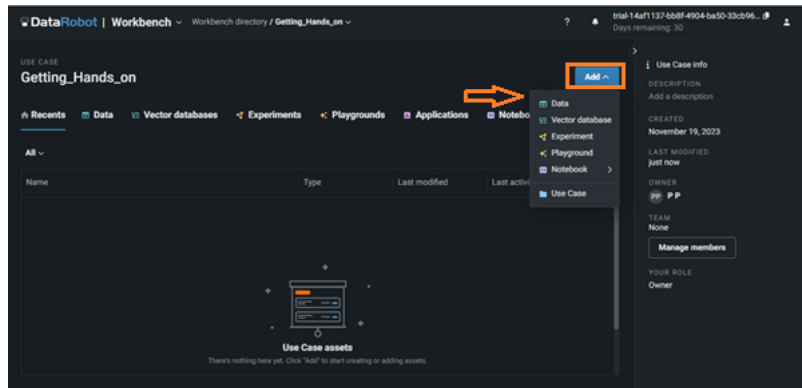
Step 7: The dashboard will appear shortly, and your screen will look as shown below. Click **Workbench**.



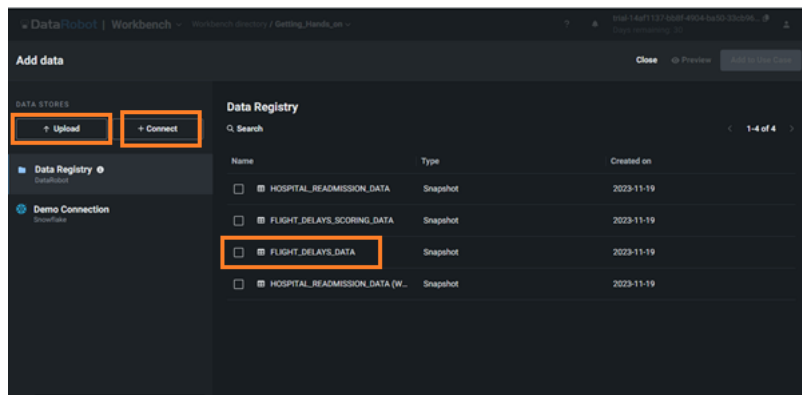
Step 8: Click **Create Use Case**.



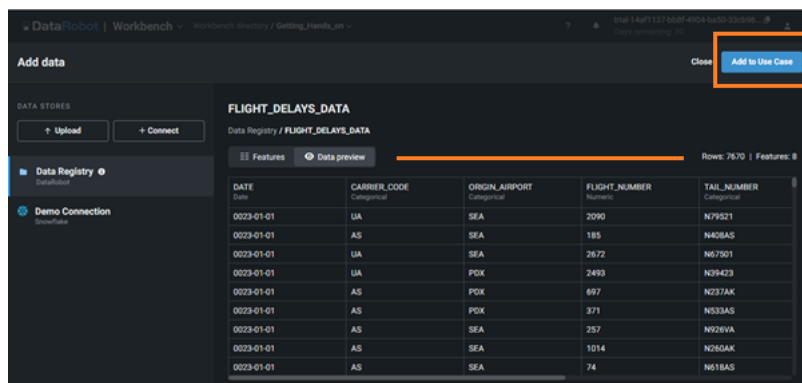
Step 9: Click **Add** and **Data** to include the data set in your use case.



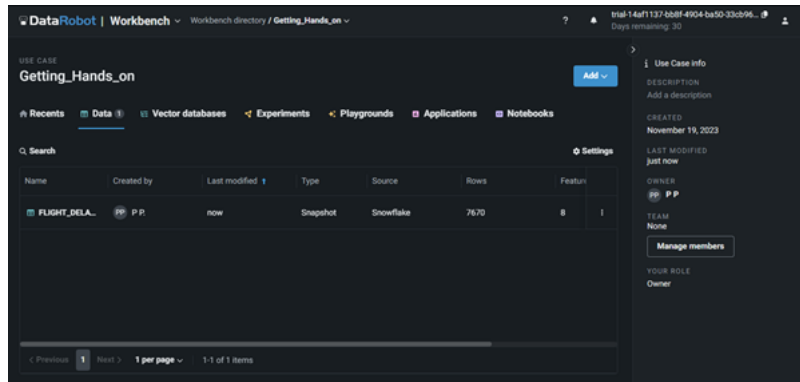
Step 10: **Upload** your data set or **Connect** to the data source; however, for this lab, you can select an in-built sample data set *FLIGHT_DELAYS_DATA*.



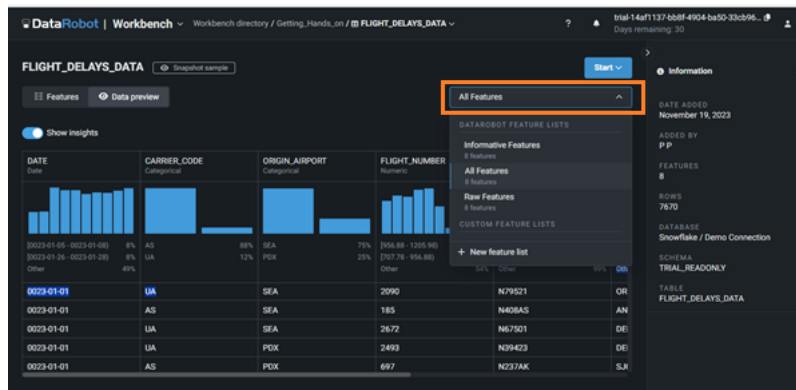
Step 11: Once you select the data set, you can see a preview of it. You can also view the data set's features, as shown below. Click **Add to Use Case**.



Step 12: After you add the data set to the use case, the workbench will appear as shown below. You can click the data set to see the feature insights.

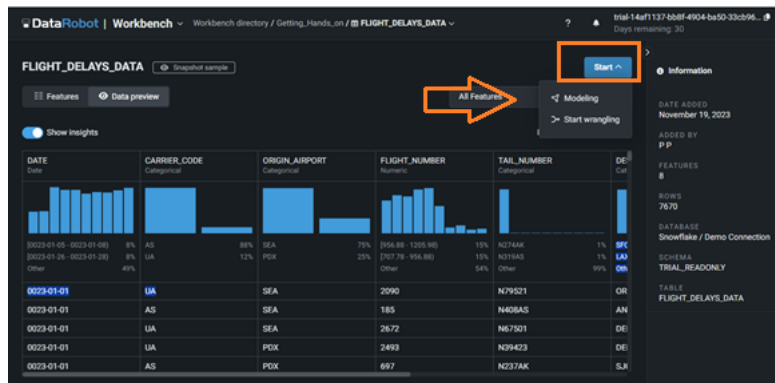


Step 13: Explore the **All Features** menu to display specific features.



Task 3: Work on Data Modeling

Step 14: Click **Start**. You will have options **Modelling** and **Start wrangling**.



Step 15: You can try data wrangling if you want to. For this lab, you will work on model building. Click **Start** and select **Modelling**. It will take a while to prepare a data set for modelling. Once done, you need to select the **Target feature**.

Set up new experiment

Dataset Target Additional settings

Preparing dataset for modeling
Once complete, select a target feature.

Experiment summary
FLIGHT_DELAYS_DATA - 2023-11-19
17:46:21

Dataset
Name: FLIGHT_DELAYS_DATA
Rows: 7670
Features: 8

Target
No target selected

Target feature
Select the feature to make predictions on.

Start typing or select a feature from the table below

Search

Feature name	Index	Var Type	Unique	Missing	Mean	Std Dev
TAKE_OFF_DELAY_...	7	Numeric	204	0	25.17	29.77
TAIL_NUMBER	4	Categorical	616	5	-	-
SCHEDULED_DEPA...	9	Categorical	20	0	-	-
SCHEDULED_DEPA...	6	Time	246	0	13:26:29	0.23 days
ORIGIN_AIRPORT	2	Categorical	2	0	-	-
FLIGHT_NUMBER	3	Numeric	558	0	854.27	579.37

Experiment summary
FLIGHT_DELAYS_DATA - 2023-11-19
17:46:21

Dataset
Name: FLIGHT_DELAYS_DATA
Rows: 7670
Features: 8

Target
No target selected

Step 16: Select **TAKE_OFF_DELAY_MINUTES** as your target feature.

Set up new experiment

Dataset Target

Target feature
Select the feature to make predictions on.

Start typing or select a feature from the table below

TAKE_OFF_DELAY_MINUTES

DATE (Day of Week)

SCHEDULED_DEPARTURE_TIME (Hour of Day)

TAIL_NUMBER

SCHEDULED_DEPA...

Unique

Missing

204

0

616

5

20

0

Experiment summary
FLIGHT_DELAYS_DATA - 2023-11-19
17:46:21

Dataset
Name: FLIGHT_DELAYS_DATA
Rows: 7670
Features: 8

Target
No target selected

Step 17: The workbench screen will be displayed as shown below. Click **Next**.

Set up new experiment

Target feature: TAKE_OFF_DELAY_MINUTES

Target type: Regression

Modeling mode: Quick Autopilot

Optimization metric: RMSE (Accuracy)

Training feature list: Informative Features

Experiment summary:

- Dataset: FLIGHT_DELAYS_DATA
- Rows: 7670
- Features: 8
- Target: TAKE_OFF_DELAY_MINUTES
- Target type: Regression
- Modeling mode: Quick Autopilot
- Optimization metric: RMSE
- Training feature list: Informative Features
- Partitioning: Random sampling
- Validation type: Cross-validation

Next >

Step 18: You can modify the model setting in **Additional Settings**; once done, click **Next** and then click **Start modelling**.

Set up new experiment

Validation type: Cross-validation

Cross-validation folds: 5

Holdout percentage: 20%

Start modelling

Step 19: Building models will take a while; once the modelling is complete, you can pick a model of your choice, and the DataRobot will show the **Model Overview**.

DataRobot | Workbench

Workbench directory / Getting_Hands_on / FLIGHT_DELAYS_DATA - 2023-11-19 17:46:21

Experiment Comparison

View experiment info

Filter models

Validation • RMSE

Models

Models are building...

Building...

Sage Regressor

Informative Features

64%

Building...

Elastic-Net Regressor (mixing alpha=0.5 / Least...

Informative Features

64%

Model Overview

Light Gradient Boosting on ElasticNet Predictions

Training scores: RMSE

Validation 26.8995

Cross-validation 29.3099

Holdout 27.4150

Training settings

Training feature list

Training sample size 64%

Informative Features

Blueprint

Feature Impact

SHAP Prediction Explanations

Download Prediction Explanations

Prediction Explanations preview for validation data

Light Gradient Boosting on ElasticNet Predictions

Informative Features

100%

26.8907

Light Gradient Boosting on ElasticNet Predictions

Informative Features

64%

26.8995

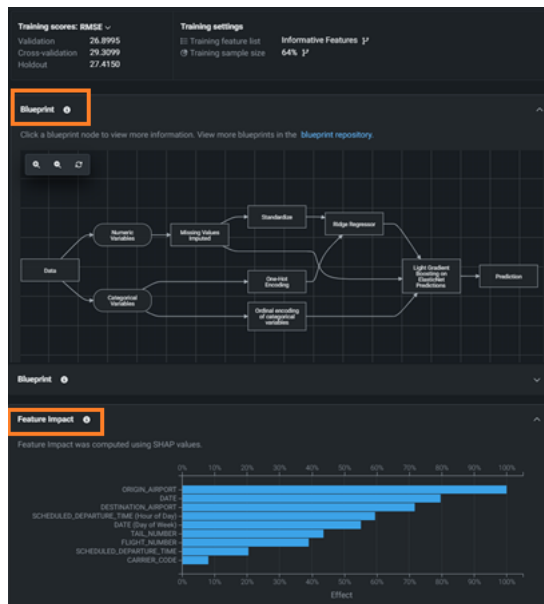
Light Gradient Boosting on ElasticNet Predictions

Informative Features

64%

26.9699

Step 20: You can explore various model overview components like **Blueprint**, **Feature Impact**, and so on.



Step 21: If you have test or unseen data, you can also make predictions by clicking **Make Predictions** under **Model actions**.

Step 22: You can also click **Generate compliance report**.

The screenshot shows the AWS SageMaker console interface. At the top, the breadcrumb is 'Getting_Hands_on / < FLIGHT_DELAYS_DATA - 2023-11-19 17:46:21'. The main heading is 'Overview'. Below it, the model name 'Boosting on ElasticNet Predictions' is shown with a star icon. On the right, a blue button labeled 'Model actions ^' is highlighted with an orange box. A dropdown menu is open, showing five options: 'Register model' (trophy icon), 'Make predictions' (chart icon, highlighted with an orange box), 'Create app' (server icon), 'Generate compliance report' (document icon), and 'Delete model' (trash icon). On the left, under 'Metrics: RMSE', three values are listed: 26.8995, 29.3099, and 27.4150. Under 'Training settings', 'Training feature list' and 'Training sample size' are shown. 'Informative Features' is listed as 64%.

The screenshot shows the 'Make new predictions' dialog. The 'Prediction dataset' section has a note 'Prediction datasets cannot exceed 5.00 GB.' and a dashed box for file upload with the text 'Drop file(s) here to upload or Choose file ^'. The 'Prediction options' section has a toggle for 'Include additional feature values' which is turned on. Below it, a note says 'When enabled, results include specified feature values taken from the prediction dataset.' There are two radio buttons: 'Add all features' (selected) and 'Add specified features'. Below the radio buttons, a text box says 'All features selected'. A dropdown menu is open over the 'Choose file ^' link, showing three options: 'Upload local file' (upload icon), 'Use model training data' (table icon), and 'Select from data registry' (folder icon).

Step 23: Then you can click **Download compliance report** for your use case.

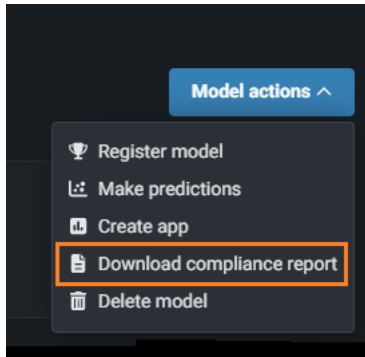


Table of Contents

- 1 How To Use This Document
- 2 DataRobot Model Development Documentation
- 3 Executive Summary and Model Overview
 - 3.1 Model Stakeholders
 - 3.2 Model Development Purpose and Intended Use
 - 3.3 Model Description and Overview
 - 3.4 Overview of Model Results

Conclusion

In this lab, you have signed up in DataRobot, added a data set in a use case, and worked on data modelling.

Author(s)

[Dr. Pooja](#)

© IBM Corporation. All rights reserved.



Skills Network