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:
• Sign up 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: Fill in the required information under the "Start for free" section and create an account.
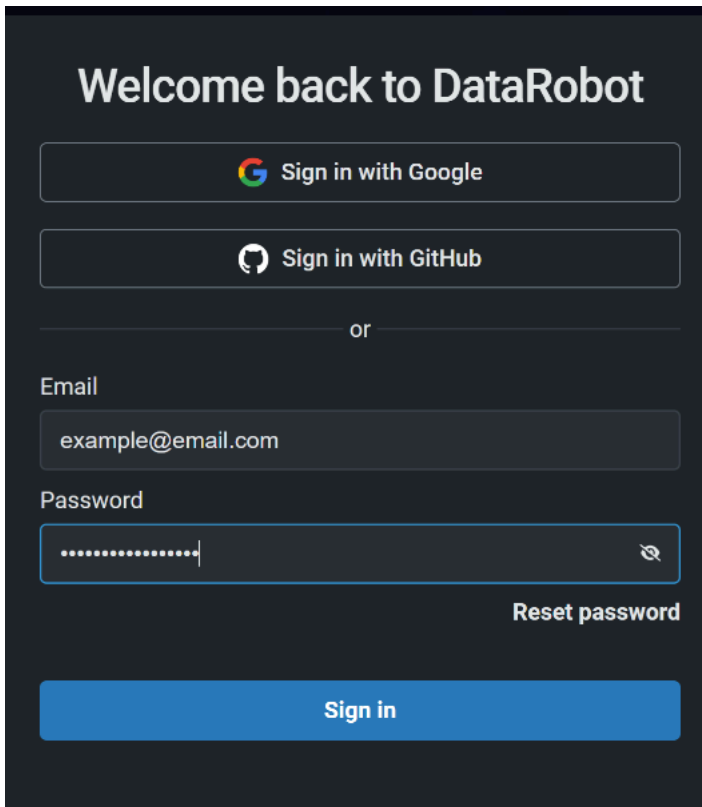

Note: To access the DataRobot platform, you must sign up using a work email address. If you do not have a relevant work email, an alternative is to create a GitHub account using your Gmail address. Once registered, you can log in to DataRobot using your GitHub credentials.

For step-by-step guidance on creating a GitHub account, please refer to the following link:

GitHub Account Setup Guide

Step 4: A new window will open; select the relevant option for signing up.
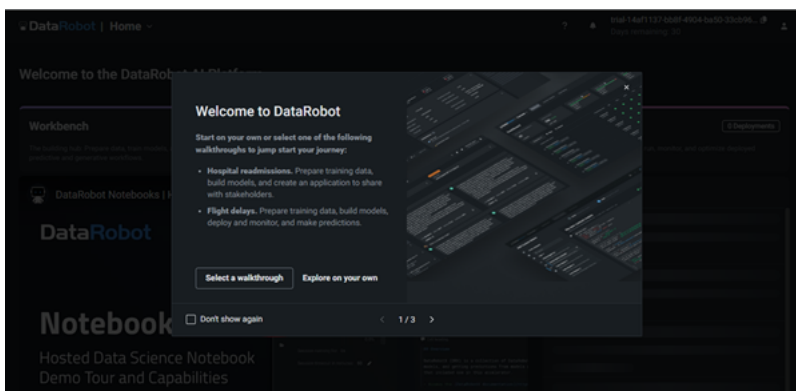
Step 5: Confirm your email by clicking **Verify Email** in your inbox.
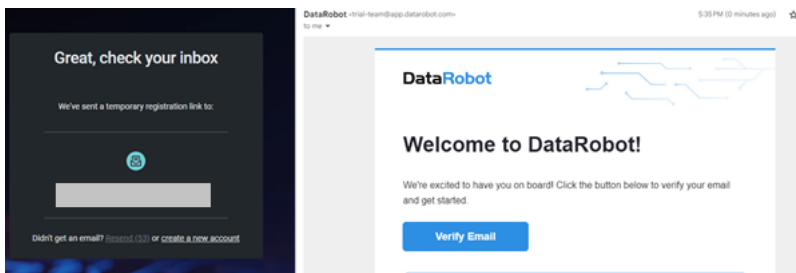
Step 6: Sign up and start your first experience of using the Generative AI tool.
The dashboard will look like the image below. You may like to familiarize yourself with the application by clicking **Select a walkthrough**.
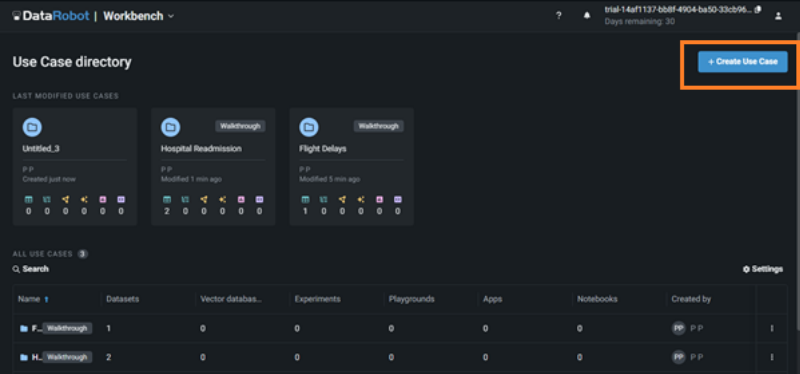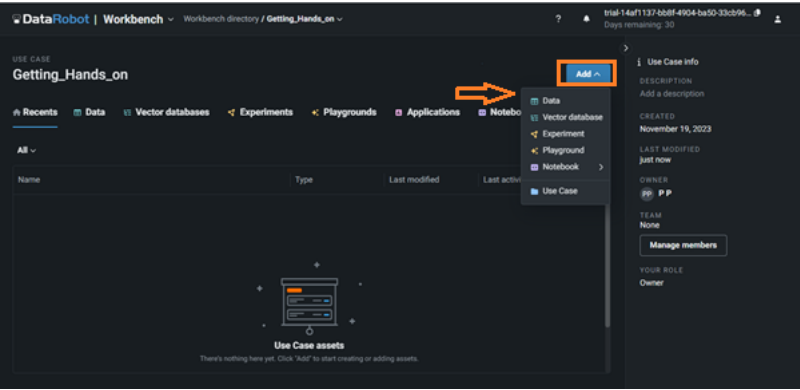

Task 2: Add a data set

Step 7: The dashboard will appear shortly, and your screen will look as shown below. Click **Workbench**.
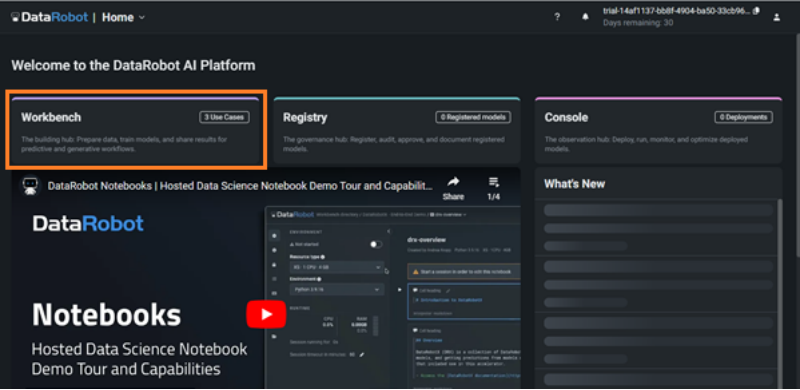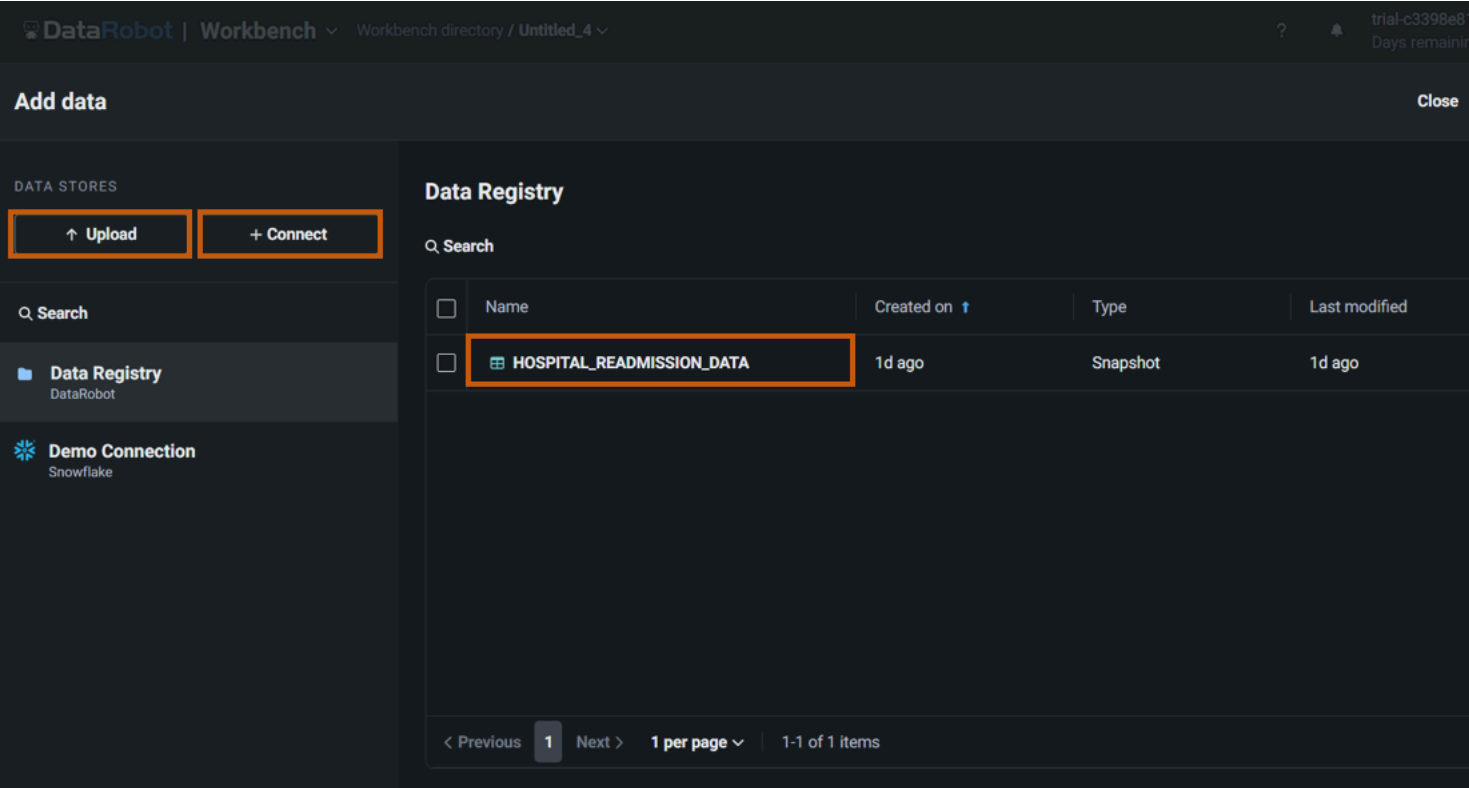
Step 8: Click **Create Use Case**.

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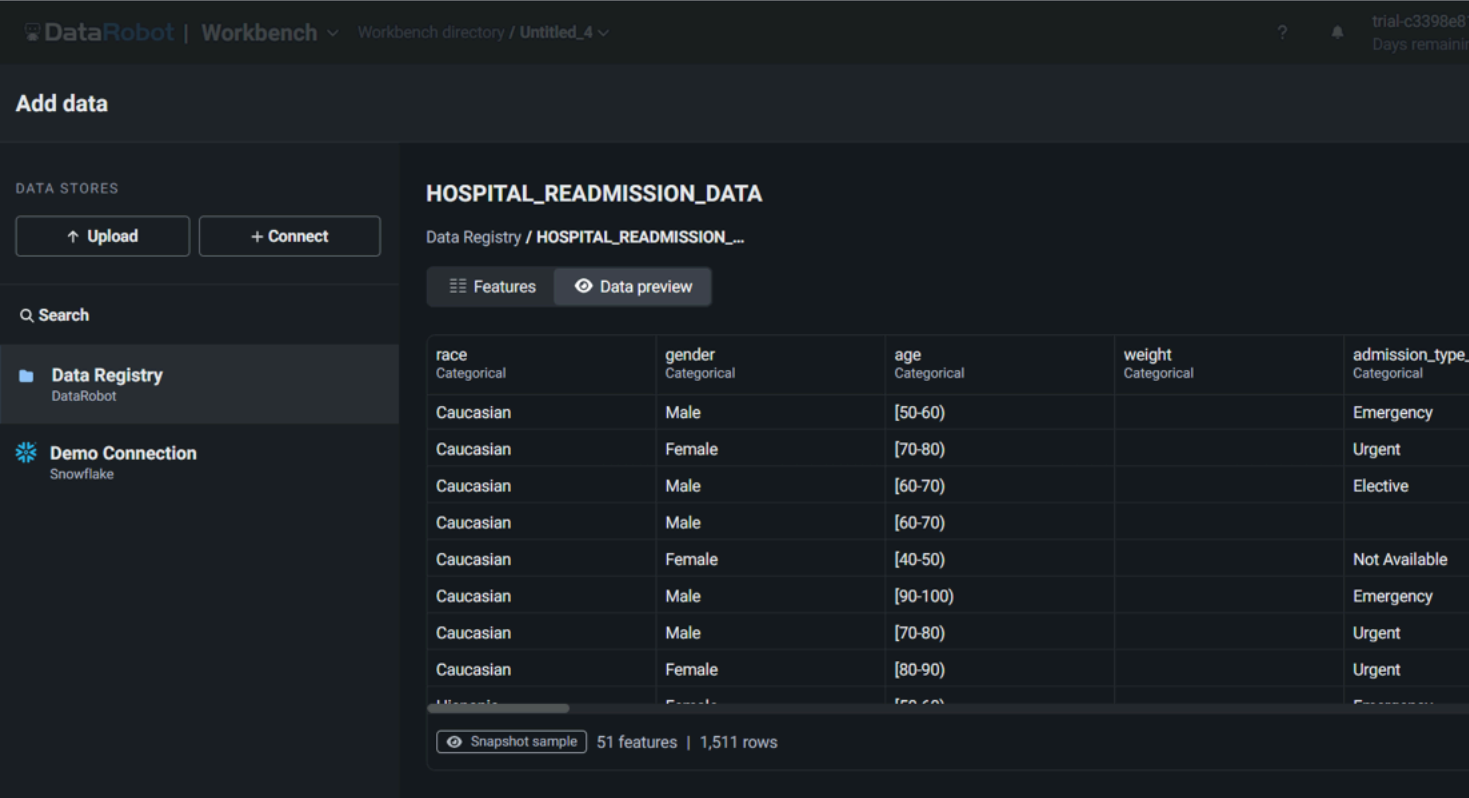
Step 12: After you add the data set to the use case, the workbench will appear as shown below. You can click the data set to see the feature insights.

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**. 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.

Step 15: Once done, you need to select the **Target feature**. Select **readmitted** as your target feature.

Step 16: The workbench screen will be displayed as shown below. Click **Next**.

Step 17: You can modify the model setting in **Additional Settings**; once done, click **Next** and then click **Start modelling**.

Step 18: Building models will take a while.

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

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

Blueprint ⓘ

Elastic-Net Classifier (mixing alpha=0.5 / Binomial Deviance)

One-Hot Encoding | Missing Values Imputed | Standardize | Matrix of word-grams occurrences | Elastic-Net Classifier (mixing alpha=0.5 / Binomial Deviance)

ElasticNet Generalized Linear Model

Click a blueprint node to view more information. View more blueprints in the blueprint repository.

Feature Impact ⓘ

Data slice   Compute method   Sort by
All data ⌄   SHAP ⌄   Feature impact • Decr... ⌄   ⬤ Use quick-compute ⓘ   🔍 Search   Actions ⌄

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

Model Overview

Elastic-Net Classifier (mixing alpha=0.5 / Binomial Deviance) ☆

Training scores: LogLoss ⌄

Validation	0.6089
Cross-validation	🔖 Score
Holdout	0.6111

Training settings

⊟ Training feature list   **Informative Features** ⌥
⏱ Training sample size   **64% (6,400 rows)** ⌥

🔖 Regi
📊 Mak
📊 Crea
📄 Gen
🗑 Dele

Blueprint ⓘ

‹ **Model Overview**

Make Predictions

Make new predictions

Prediction dataset
Prediction datasets cannot exceed 5.00 GB.

Drop file(s) here to upload or   📁 **Choose file** ⌃

    ⬆ Upload local file
    ⊞ Use model training data
    📁 Data registry

Prediction options

⬤ Include additional feature values in predi
When enabled, results include specified feature values taken from the prediction dataset.

◉ Add all features   ◯ Add specified features

All features selected

◯ Include Prediction Explanations ⓘ

Compute and download predictions

Download recent predictions
Predictions are stored and available to download for 48 hours.

No recent predictions

Step 22: You can also click **Generate compliance report** and **download compliance report** for your use case.

Conclusion

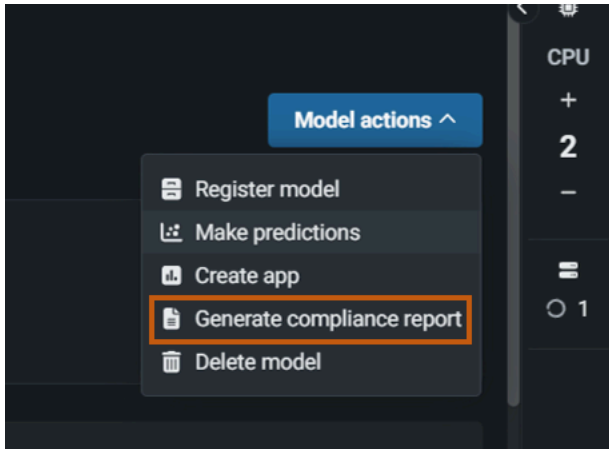In this lab, you have signed up in DataRobot, added a data set in a use case, and worked on data modelling.

Author(s)

Dr. Pooja