

# 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

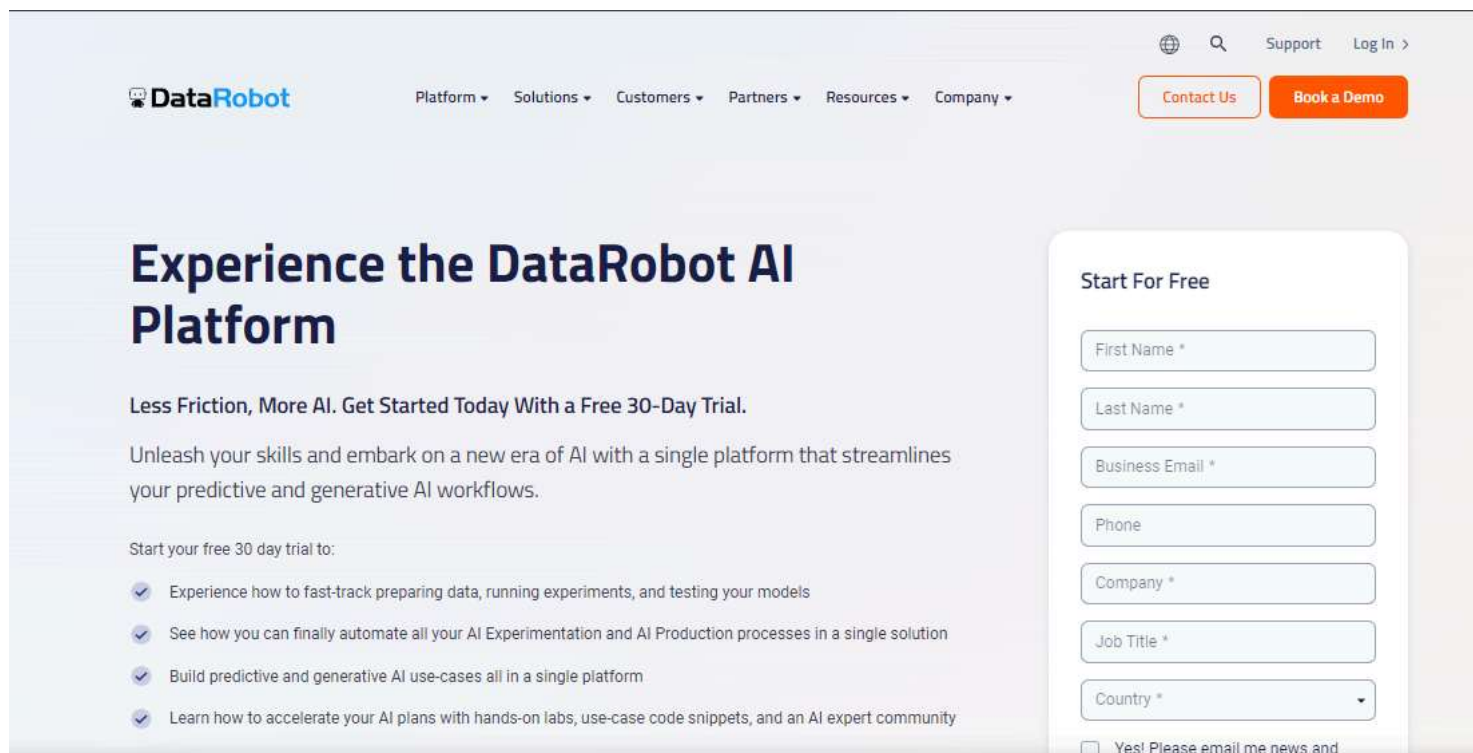
After completing this lab, you will be able to:

- 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](https://www.datarobot.com)

Step 2: Fill in the required information under the "Start for free" section and create an account.



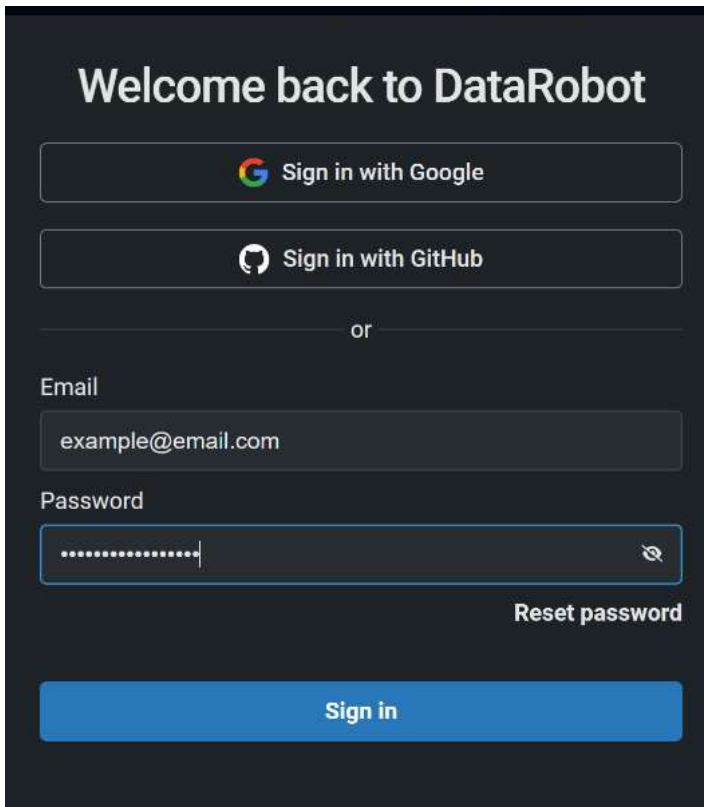
The screenshot shows the DataRobot website's homepage. The header includes the DataRobot logo, navigation links (Platform, Solutions, Customers, Partners, Resources, Company), and links for Support, Log In, Contact Us, and Book a Demo. The main content area features the heading "Experience the DataRobot AI Platform" and a subheading "Less Friction, More AI. Get Started Today With a Free 30-Day Trial." Below this, a paragraph states: "Unleash your skills and embark on a new era of AI with a single platform that streamlines your predictive and generative AI workflows." A section titled "Start your free 30 day trial to:" lists four benefits with checkmarks. On the right, a "Start For Free" form is visible, containing input fields for First Name, Last Name, Business Email, Phone, Company, Job Title, and Country, followed by a checkbox for "Yes! Please email me news and".

**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.



>Welcome back to DataRobot

Sign in with Google

Sign in with GitHub

or

Email

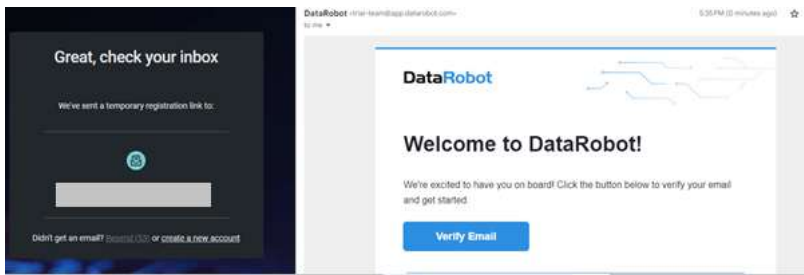
example@email.com

Password

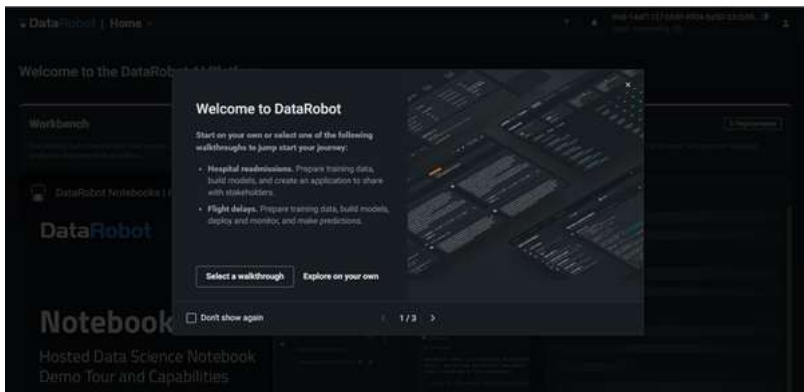
Reset password

Sign in

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

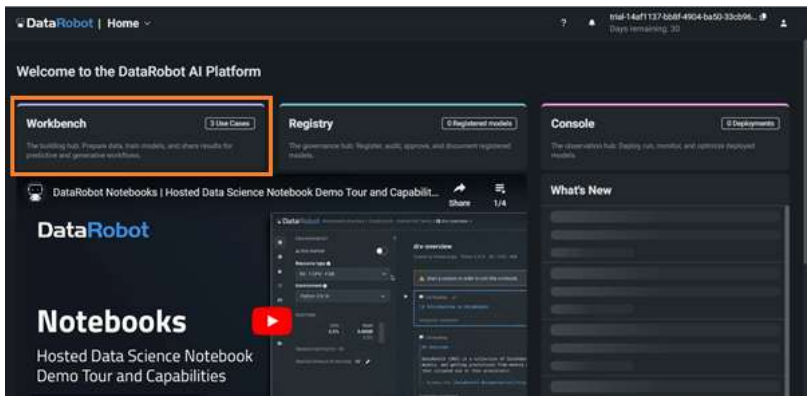


Step 6: Sign up and start your first experience of using the Generative AI tool.  
The dashboard will look like the image below. You may like to familiarize yourself with the application by clicking **Select a walkthrough**.

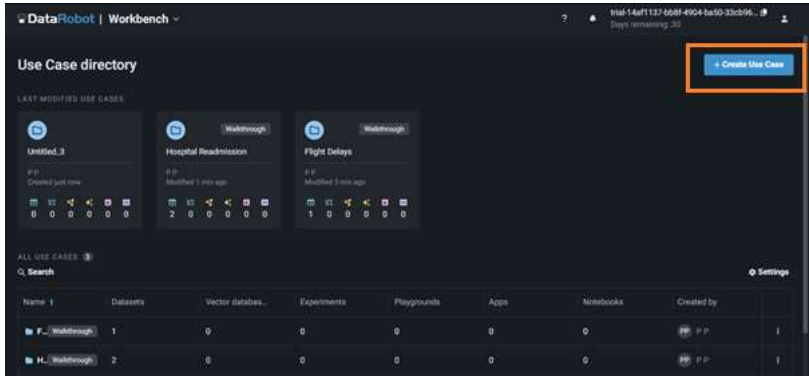


## Task 2: Add a data set

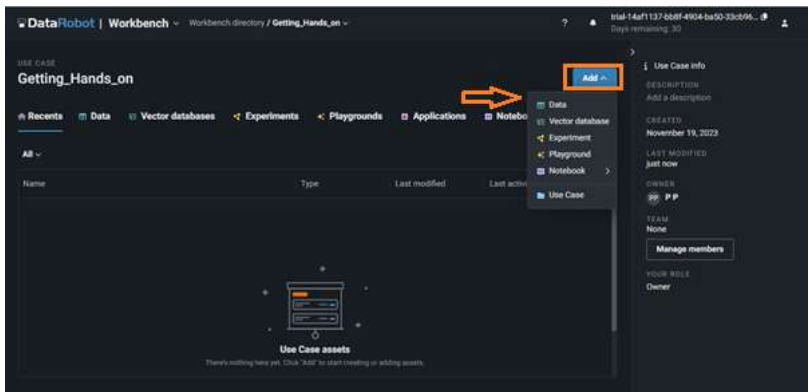
Step 7: The dashboard will appear shortly, and your screen will look as shown below. Click **Workbench**.



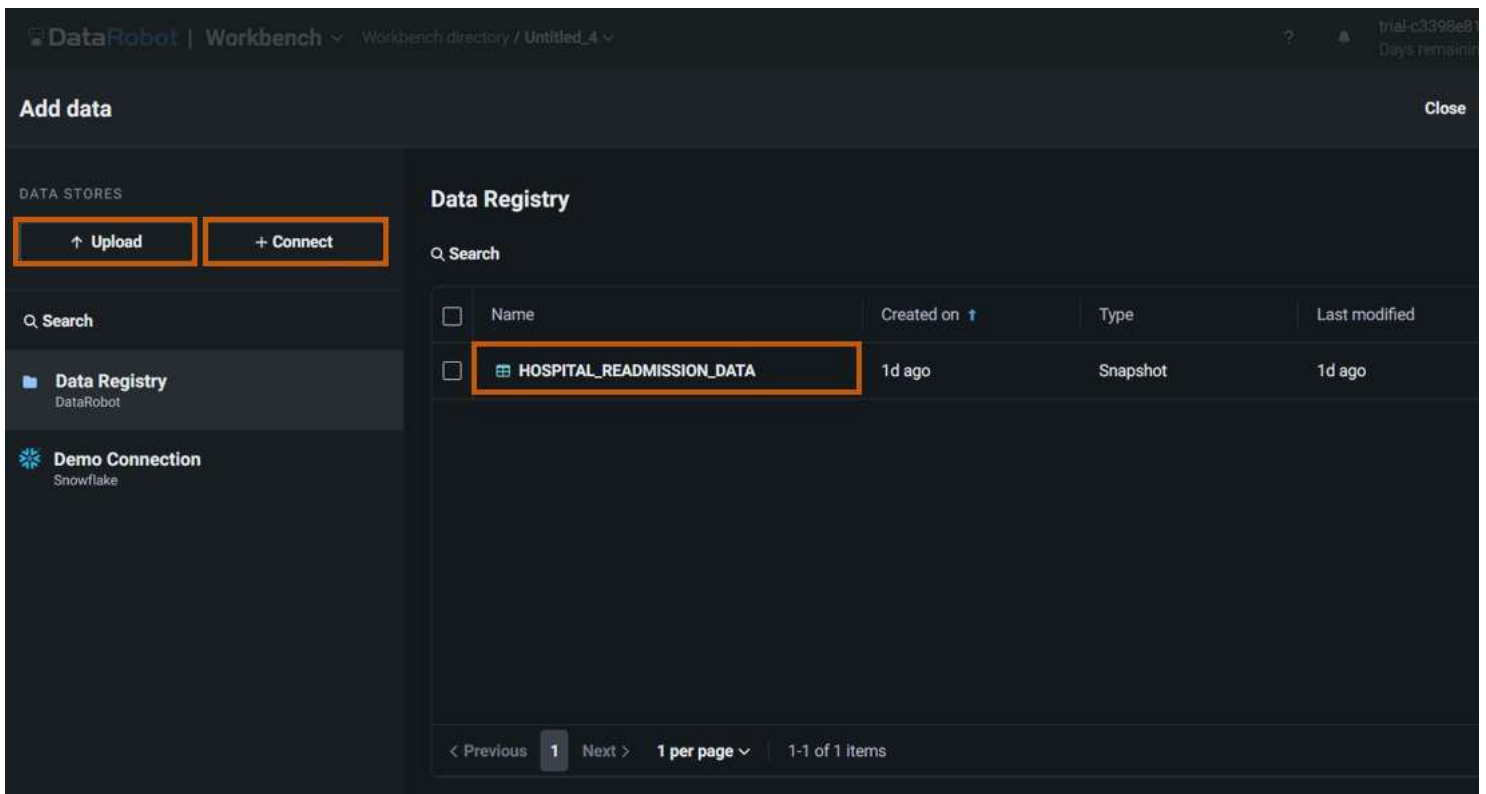
Step 8: Click **Create Use Case**.



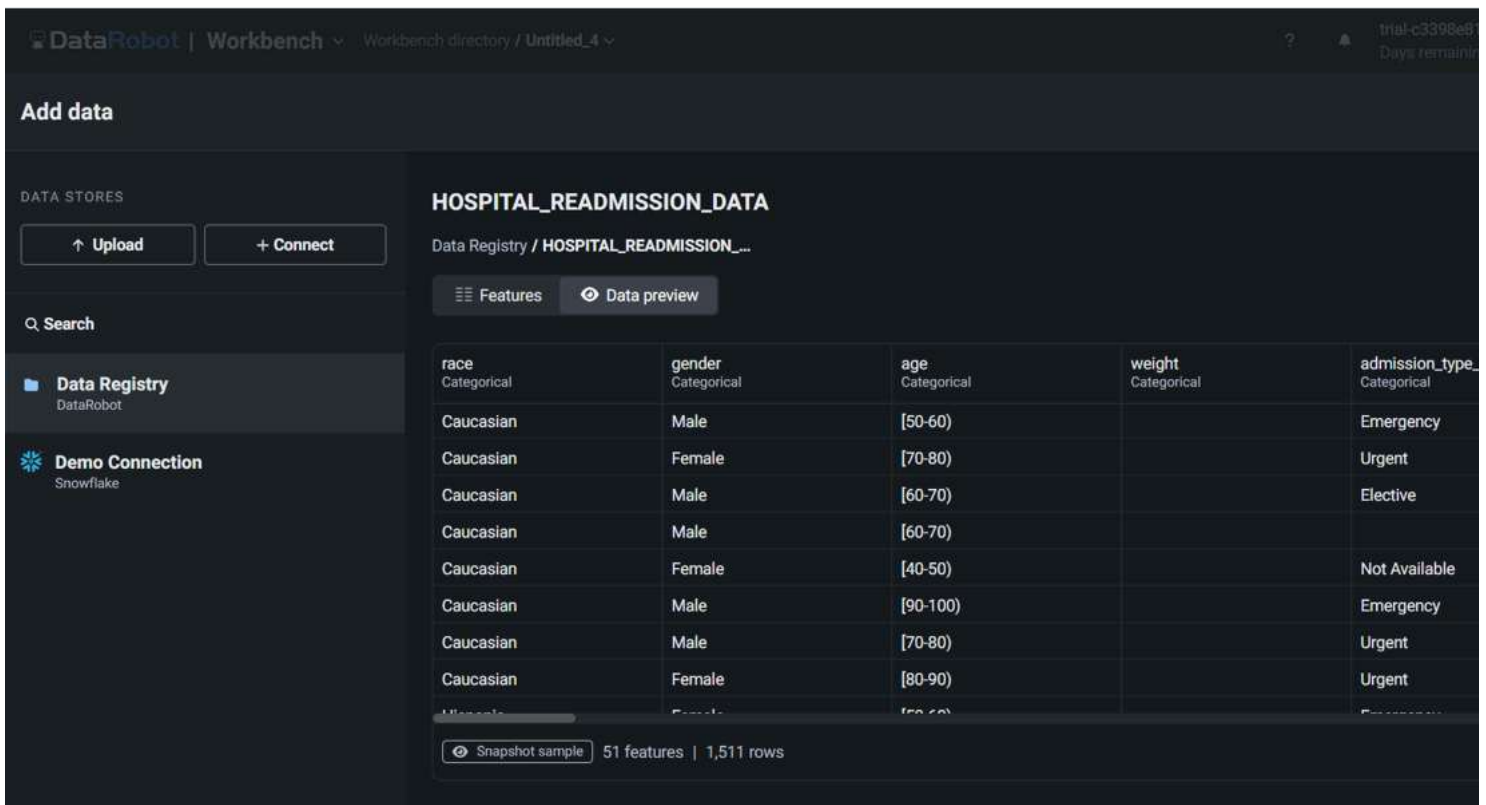
Step 9: Click **Add** and **Data** to include the data set in your use case.



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 *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.

**Workbench**

Workbench directory / **Untitled\_4**

?
🔔
trial-c3398e81-5e8 Days remaining: 2

< Use Case directory

USE CASE

Untitled\_4

⋮
Add Data
▼

⋮
All
Data 1
Vector databases
Experiments
Playgrounds
Applications

Type ▼
🔍 Search
⚙️ Settings

Name	Created By	Last Modified ↑	Type	Source	Rows	
HOSPITAL_READMISSION_DATA	Pratiksha V.	now	Snapshot	Snowflake	10000	⋮

< Previous
1
Next >
1 per page ▼
1-1 of 1 items

Use Case

DESCRIPTION

Add a description

CREATE DATE

July 29, 2024

LAST MODIFIED

just now

OWNER

Pratiksha V.

TEAM

None

MANAGER

Manage

YOUR ROLE

Owner

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

**Workbench**

Workbench directory / **Untitled\_4** / **HOSPITAL\_READMISSION\_DATA**

?
🔔
trial-c3398e81-5e8 Days remaining: 2

HOSPITAL\_READMISSION\_DATA

Jul 29th, 2024 10:39 AM
📷 Snapshot ▼
Data actions ▼

👁 Data preview
**📋 Features**
📋 Feature lists

🔍 Show insights
Show features from: All Features ^
+ Create feature list

🔍 Search

DATAROBOT FEATURE LISTS

All Features

51

Informative Features

40

Raw Features

51

race

Categorical

Caucasian

21%

Male

46%

AfricanAmerican

6%

Other

age

Categorical

[70-80]

26%

[60-70]

22%

Other

52%

weight

Categorical

==Missing==

96%

[75-100]

2%

Other

2%

admission\_type\_id

Categorical

Emergency

49%

Urgent

19%

Other

32%

discharge\_id

Categorical

Discharged

49%

Discharged

19%

Other

32%

Snapshot sample

51 features | 1,511 rows

### 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.





DataRobot | Workbench Workbench directory / Untitled\_4

trial-c3398e81-5e Days remaining: 2

## Set up new experiment

Dataset Target Additional settings

**Target feature**  
Select the feature to make predictions on.

readmitted

Target type: Binary classification

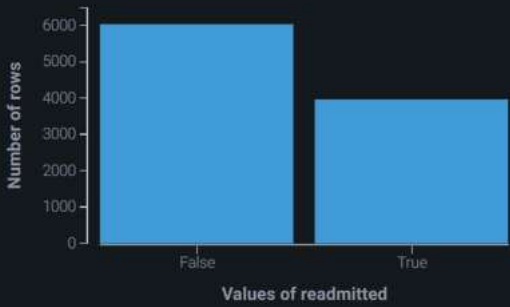
Positive class: ☐ 0 ☒ 1

**Modeling mode**  
Set the mode used for selecting which blueprints to build when training models.

Quick Autopilot

**Optimization metric**  
Set the metric used when training models to evaluate and optimize accuracy.

LogLoss (Accuracy) Recommended



**Experiment summary**  
HOSPITAL\_READM  
07-30 11:23:57

**Dataset**  
Name

**Rows**  
Features

**Target**  
Feature  
Target type  
Positive class  
Modeling mode  
Optimization metric  
Training feature list

**Partitioning**

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

DataRobot | Workbench Workbench directory / Untitled\_4

trial-c3398e81-5e Days remaining: 2

## Set up new experiment

Dataset Target Additional settings Exit

Data partitioning Time series modeling Preview Additional settings

**Partitioning method**  
Select the method for assigning rows to partitions when training models.

Stratified sampling  
Rows are assigned to ensure similar target distribution across each partition.

**Validation type**

☒ Cross-validation  
Trains models on a specified number of folds, maximizing data use but also increasing run time.

☐ Training-validation-holdout  
Splits data into three partitions: trains models on the training set, assess performance on the validation set, and evaluates the model on unseen data in the holdout set.

**Cross-validation folds**  
Enter a value from 2 - 50.

**Holdout percentage**  
Set the subset of data that is unavailable during training and validation. Enter a value

**Experiment summary**  
HOSPITAL\_READM  
07-30 11:23:57

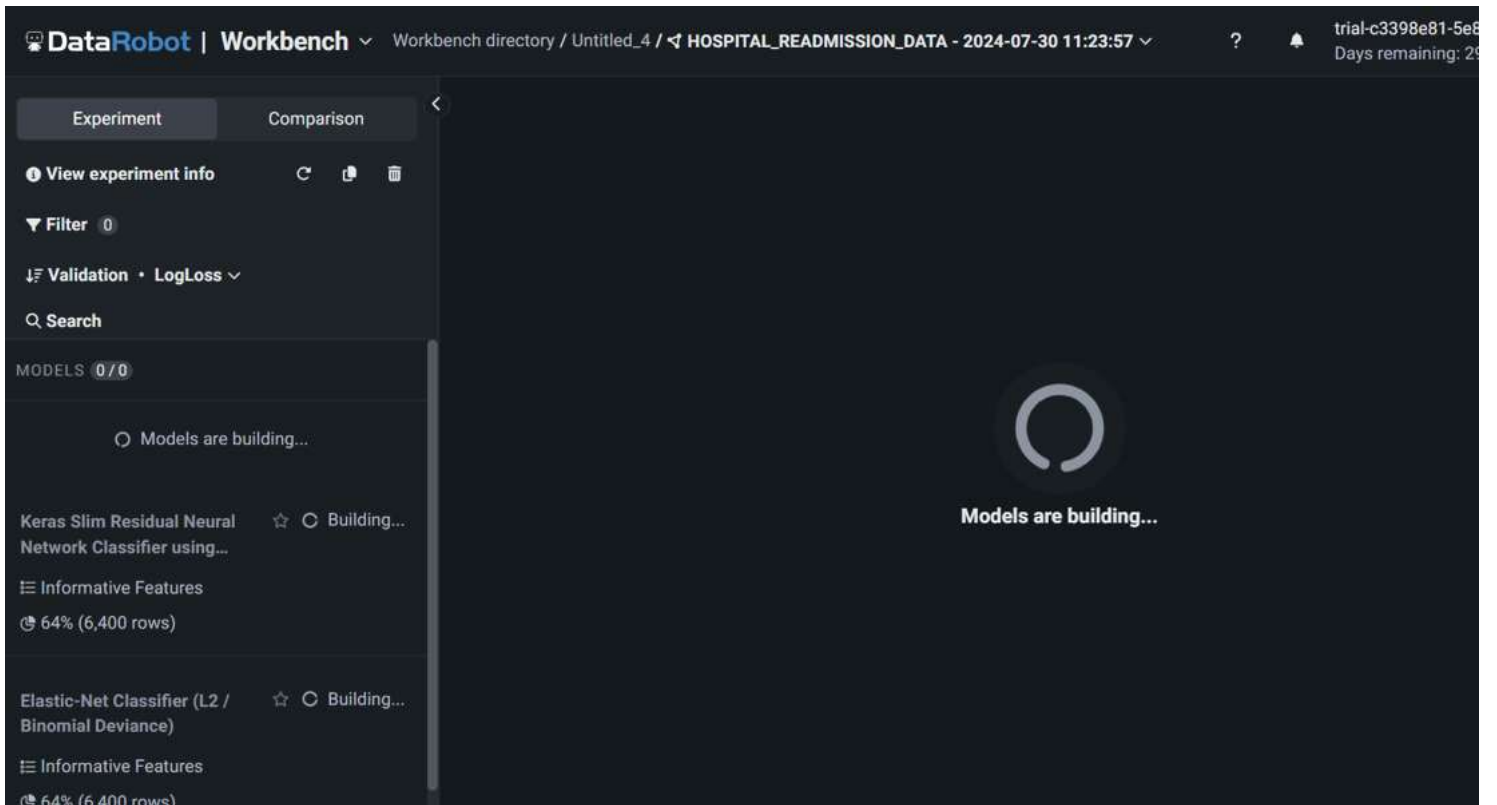
**Dataset**  
Name

**Rows**  
Features

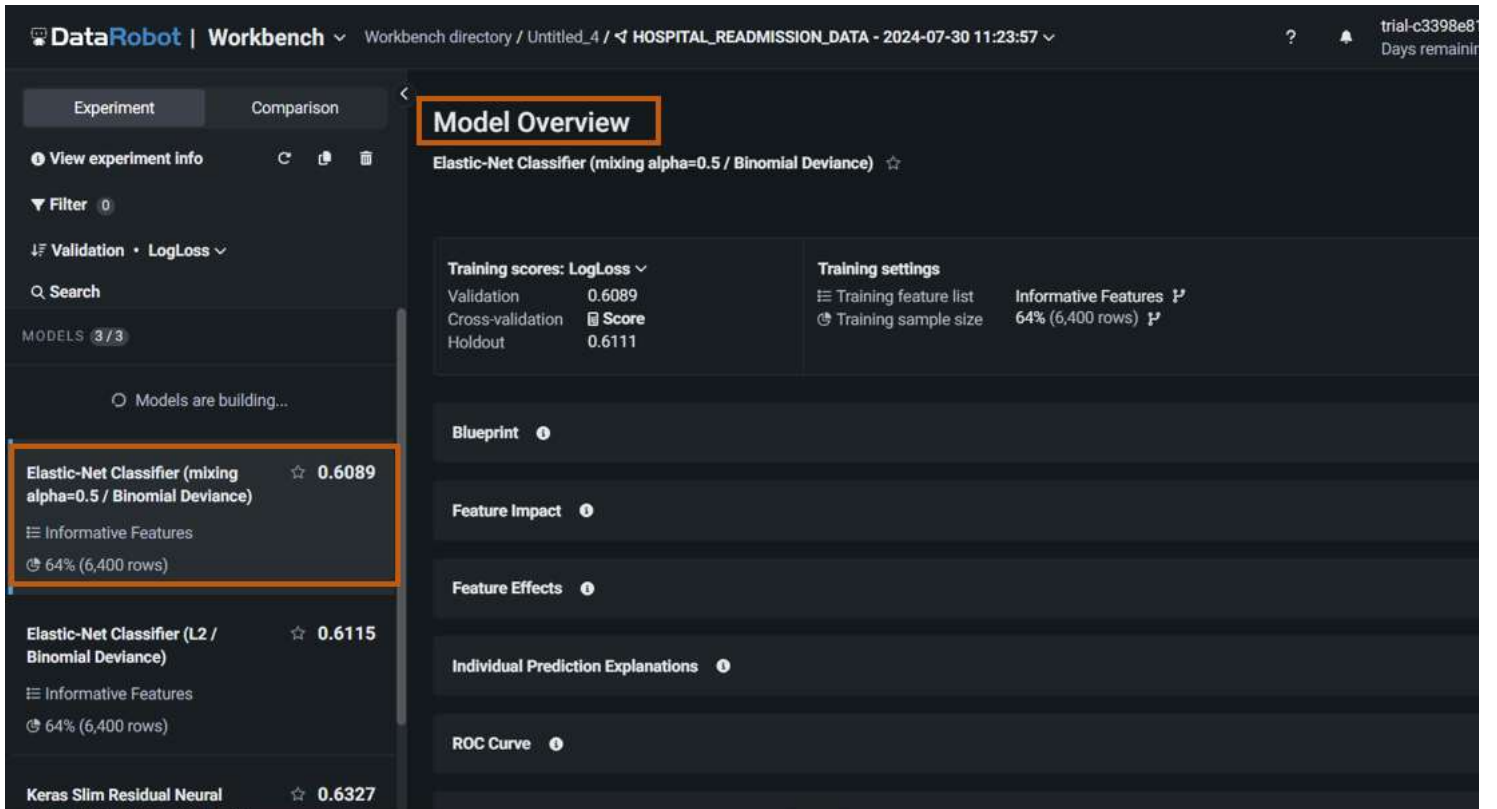
**Target**  
Feature  
Target type  
Positive class  
Modeling mode  
Optimization metric  
Training feature list

**Partitioning**

Step 18: Building models will take a while.

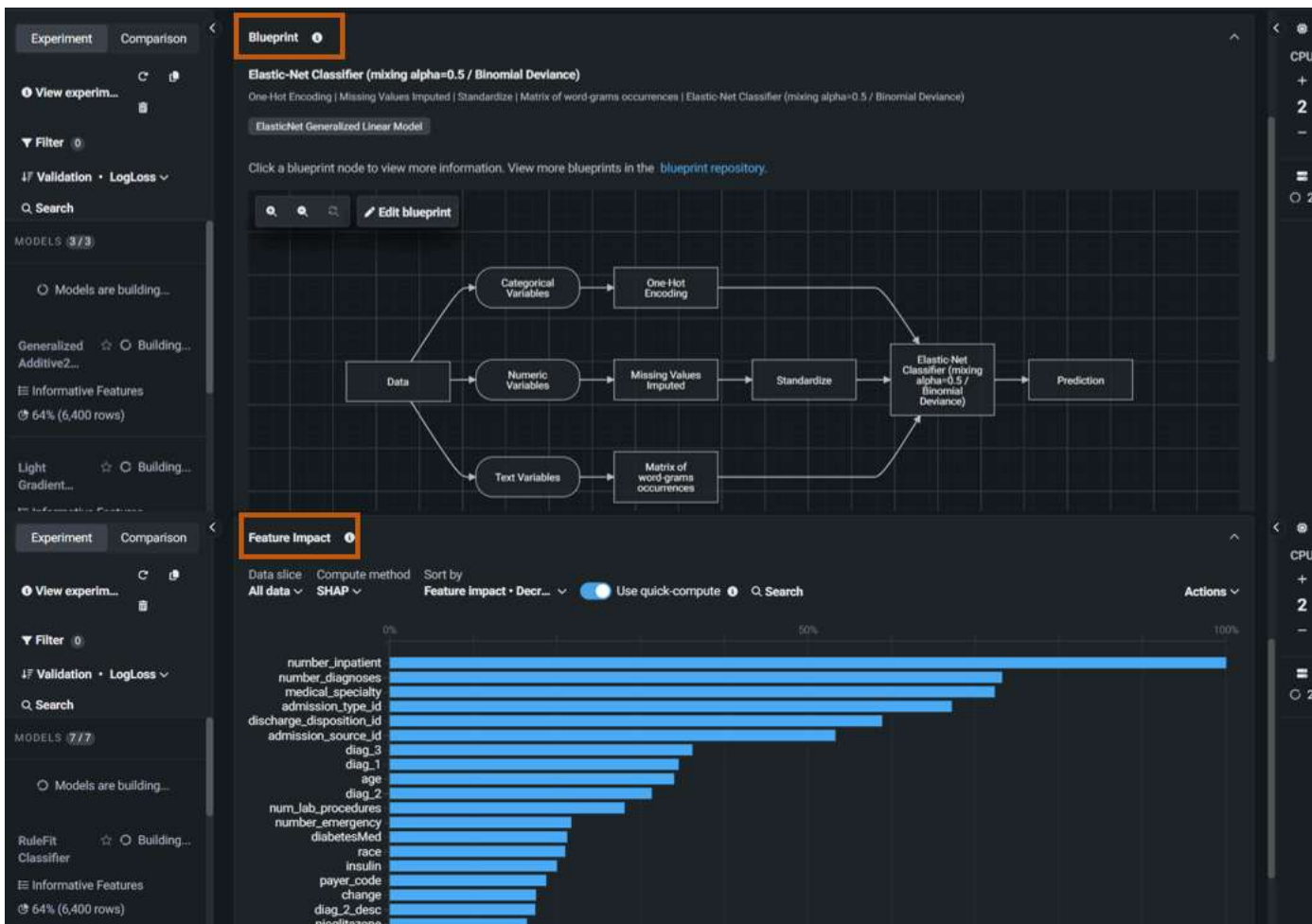


Step 19: once the modelling is complete, you can pick a model of your choice, and the DataRobot will show the **Model Overview**.



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





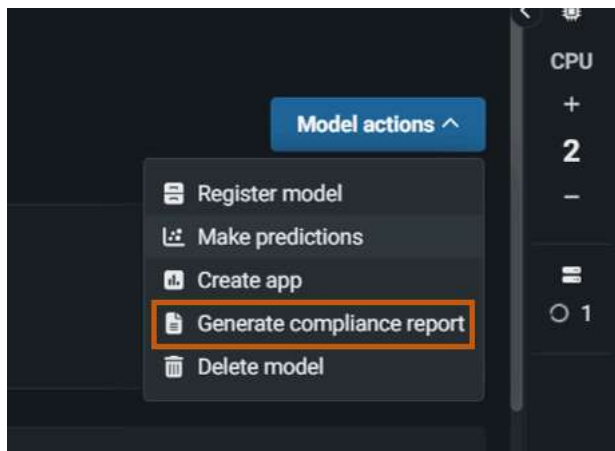
Step 21: If you have test or unseen data, you can also make predictions by clicking **Make Predictions** under **Model actions**.

The screenshot shows the 'Model Overview' page for an Elastic-Net Classifier (mixing alpha=0.5 / Binomial Deviance). The page is divided into several sections:

- Training scores:** LogLoss, Validation (0.6089), Cross-validation (Score), Holdout (0.6111).
- Training settings:** Training feature list, Informative Features (64% (6,400 rows)), Training sample size.
- Blueprint:** A link to view the model's architecture.
- Make Predictions:** A section for making new predictions, including a prediction dataset upload area and prediction options.
- Download recent predictions:** A section for downloading recent predictions, with a note that predictions are stored and available to download for 48 hours.

The 'Make Predictions' section includes a 'Prediction dataset' upload area with a 'Choose file' button. Below this, there are 'Prediction options' including a toggle for 'Include additional feature values in prediction' (checked), a radio button for 'Add all features' (selected), and a radio button for 'Add specified features'. There is also a toggle for 'Include Prediction Explanations' (unchecked).

Step 22: You can also click **Generate compliance report** and **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
  - 3.5 Model Interdependencies
- 4 Model Data Overview
  - 4.1 Feature Association
  - 4.2 Data Source Overview and Appropriateness
  - 4.3 Input Data Extraction, Preparation, and Quality & Completeness

## 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](#)



# Skills Network