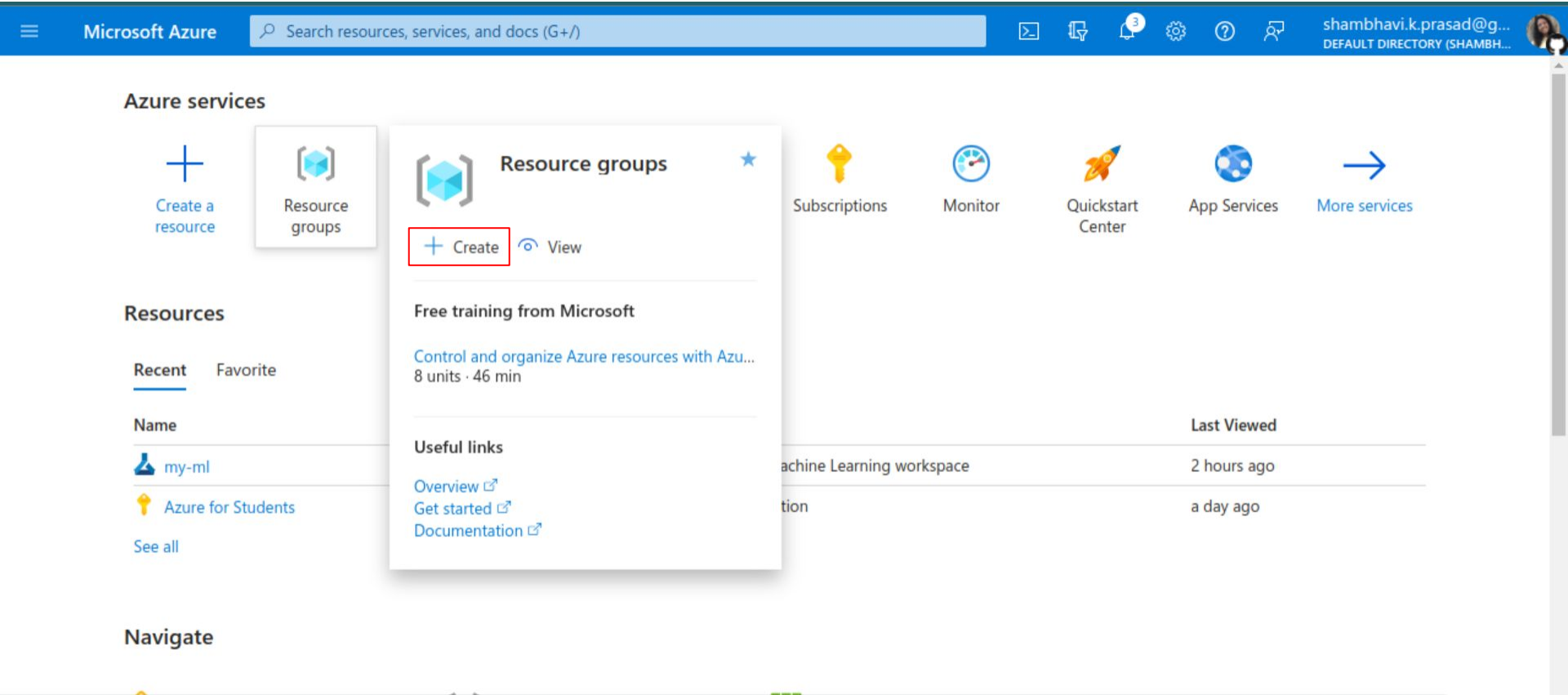
Week 1: Model Development using Azure ML Studio Designer

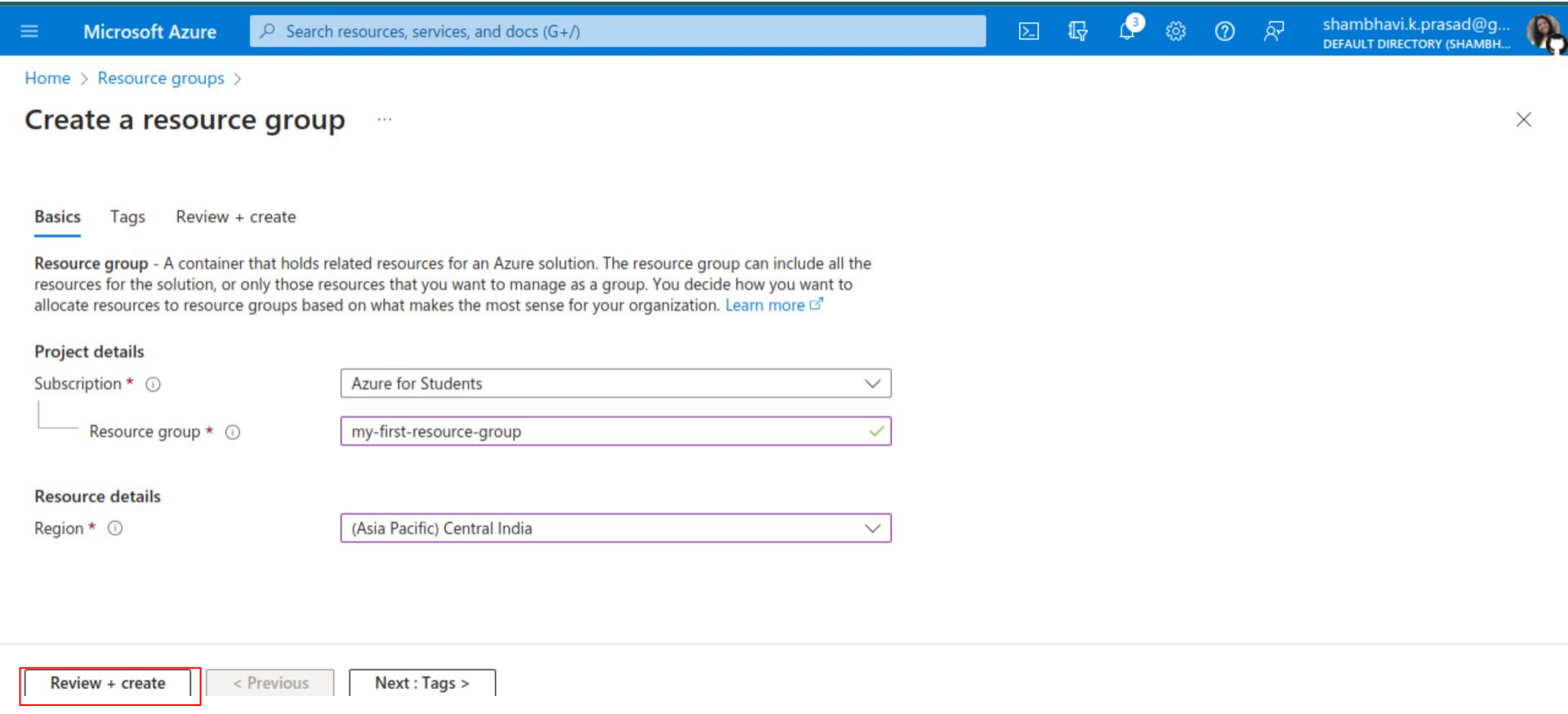
Agenda:

1. Setting up an ML pipeline using Designer (Azure ML)
2. Executing the pipeline
3. Scoring and Evaluating the model.

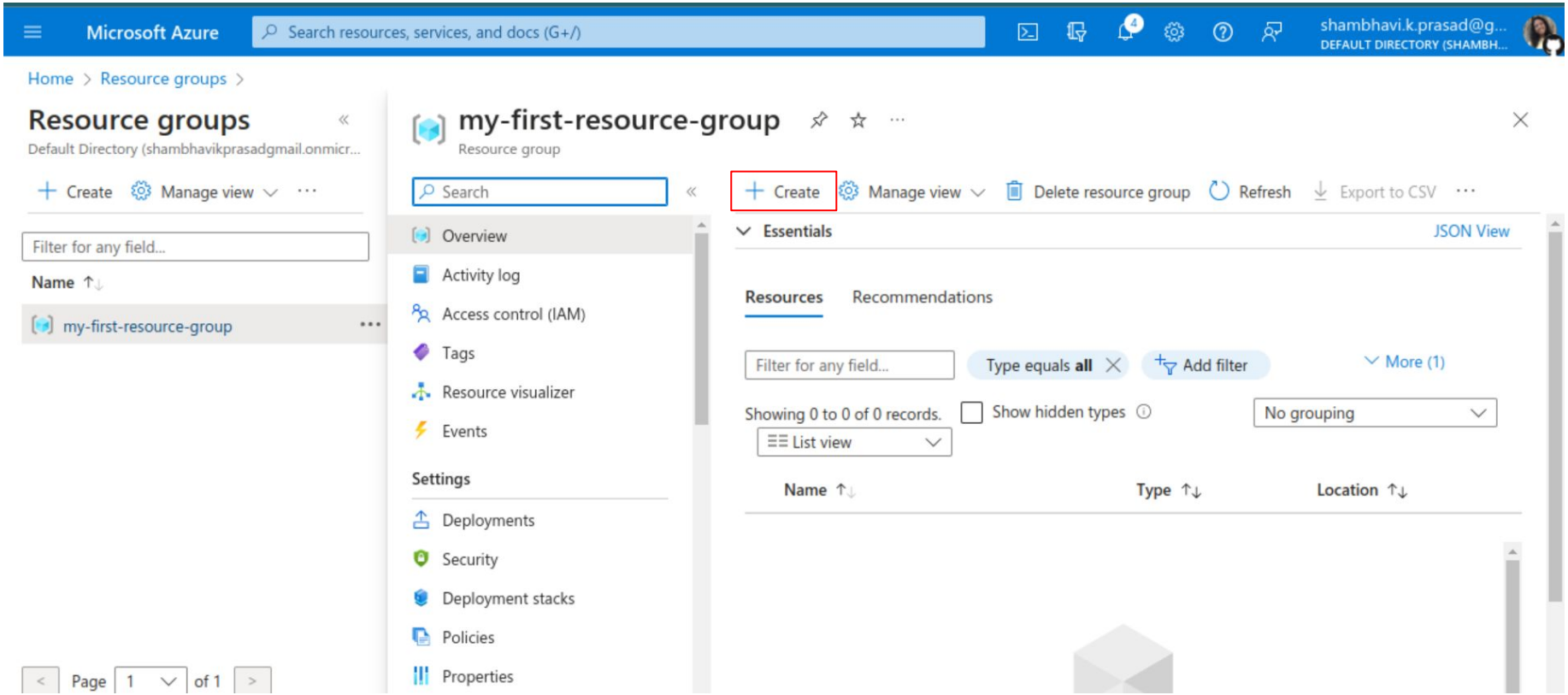
How to create a new resource group?



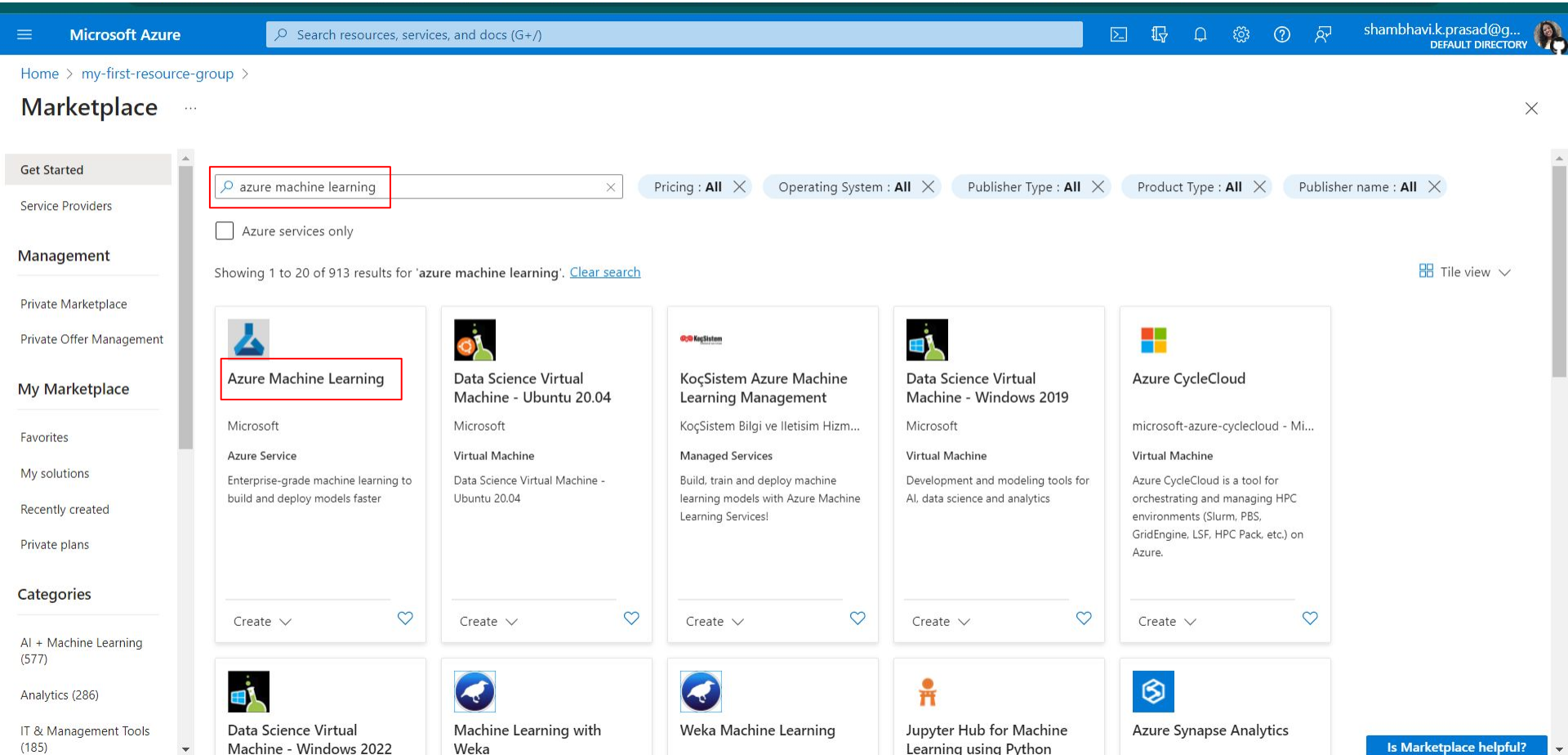
How to create a new resource group?



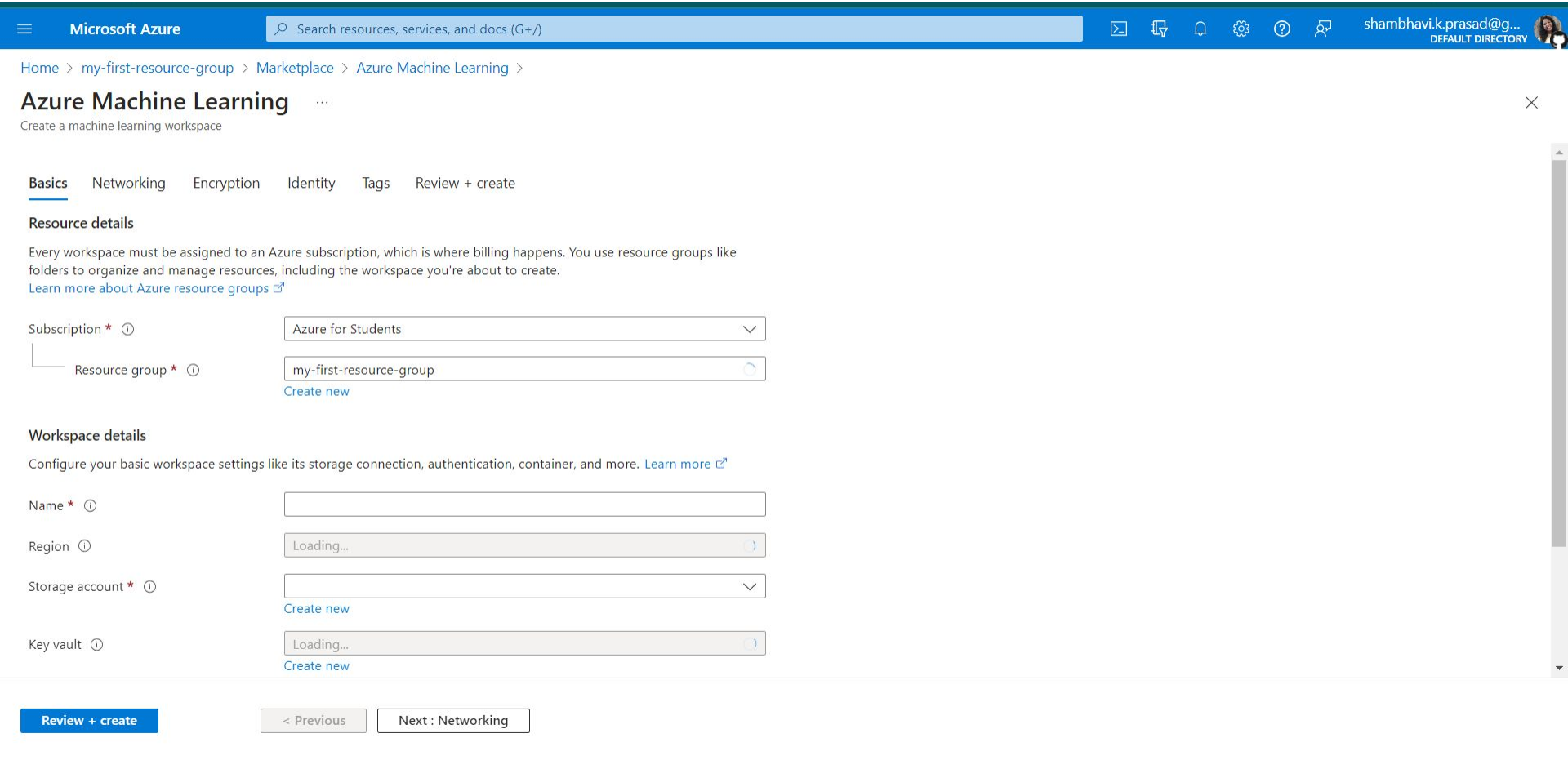
How to add a new resource to the resource group?



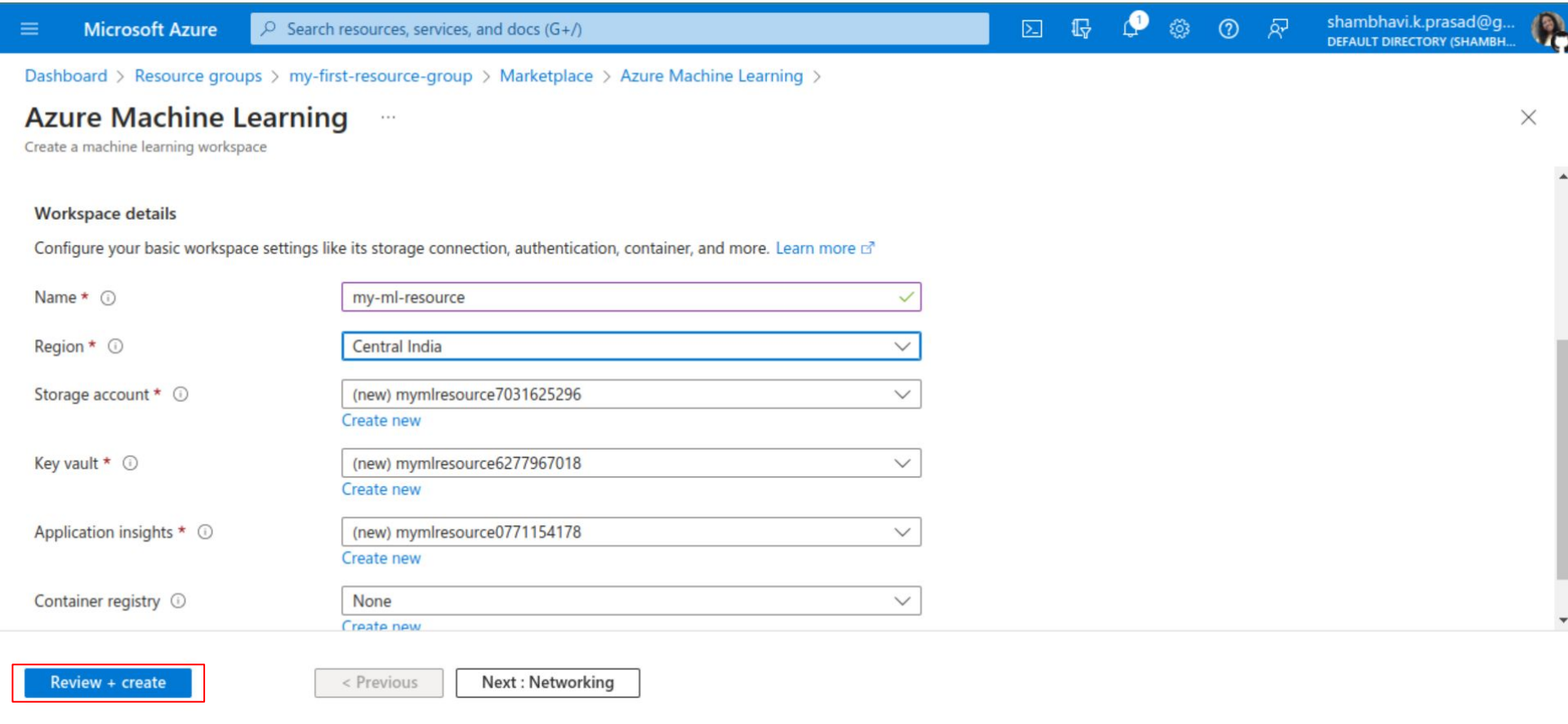
How to create a new ML resource?



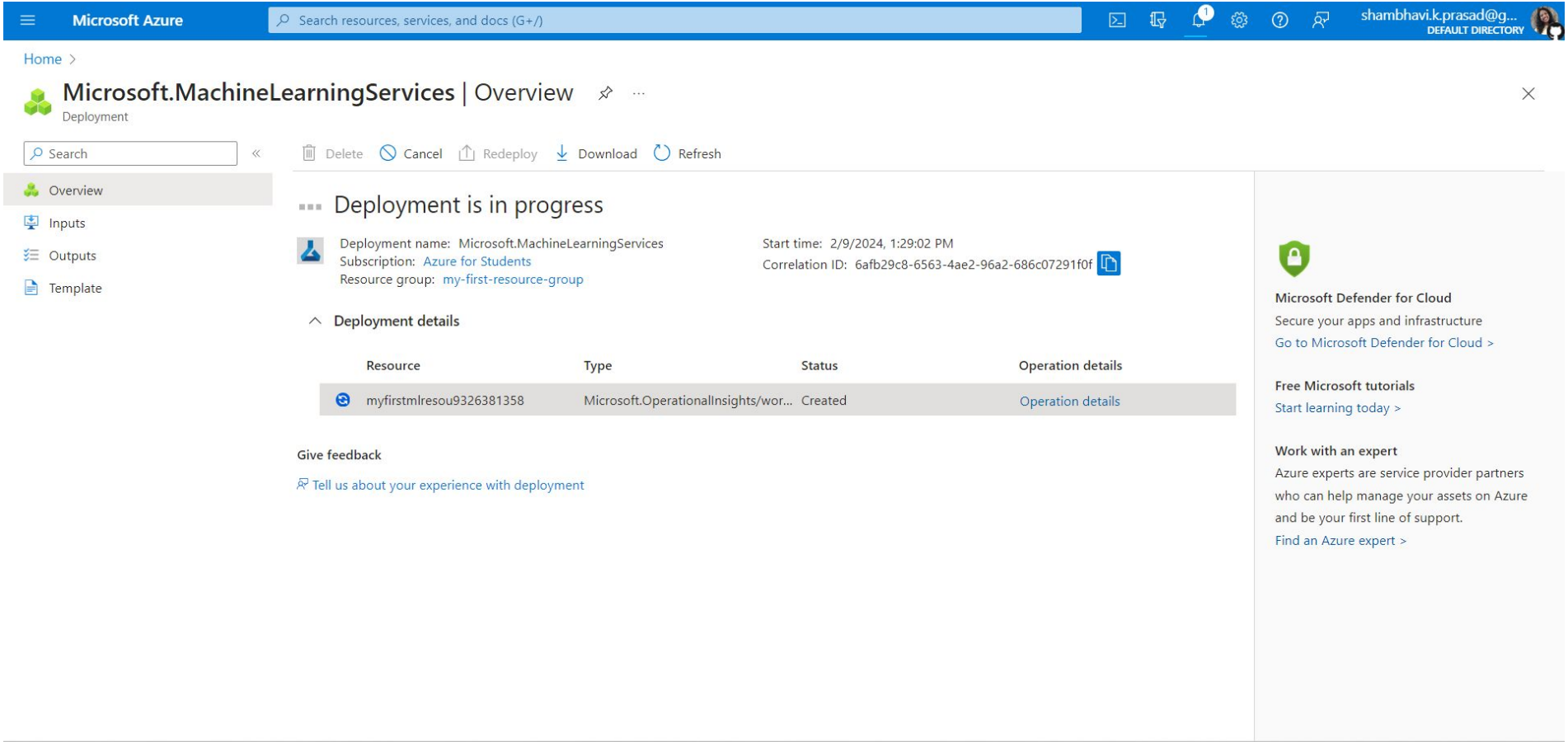
How to create a new ML resource?



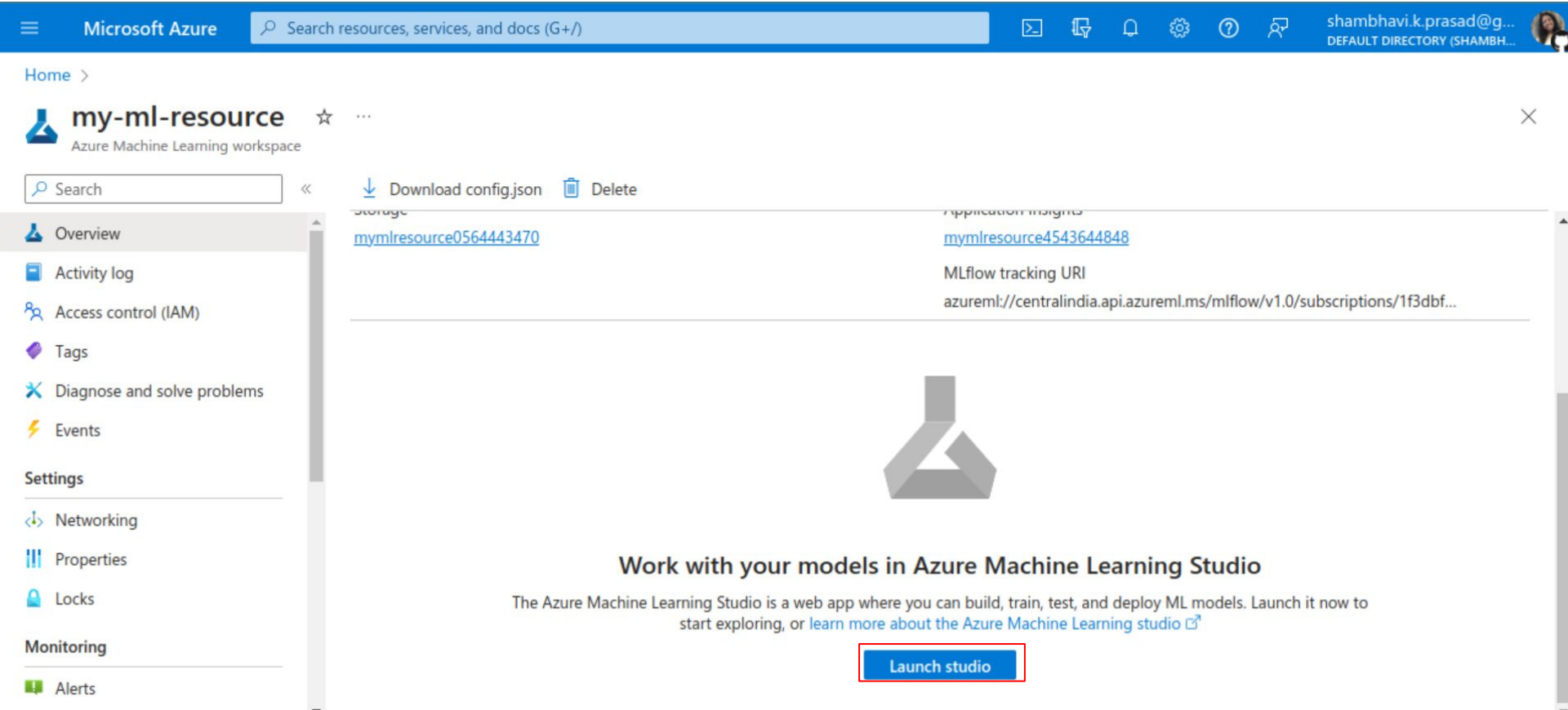
How to create a new ML resource?



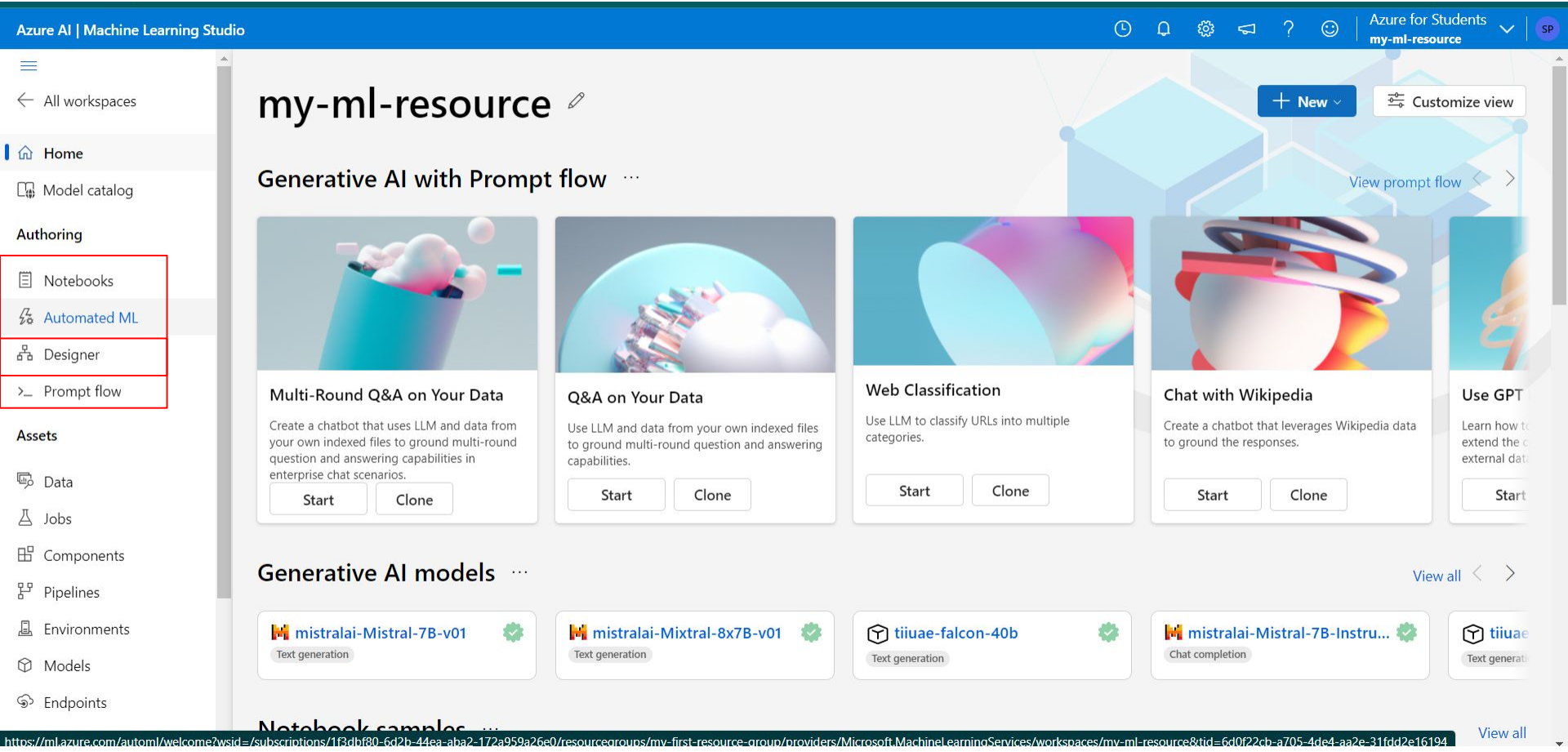
How to create a new ML resource?



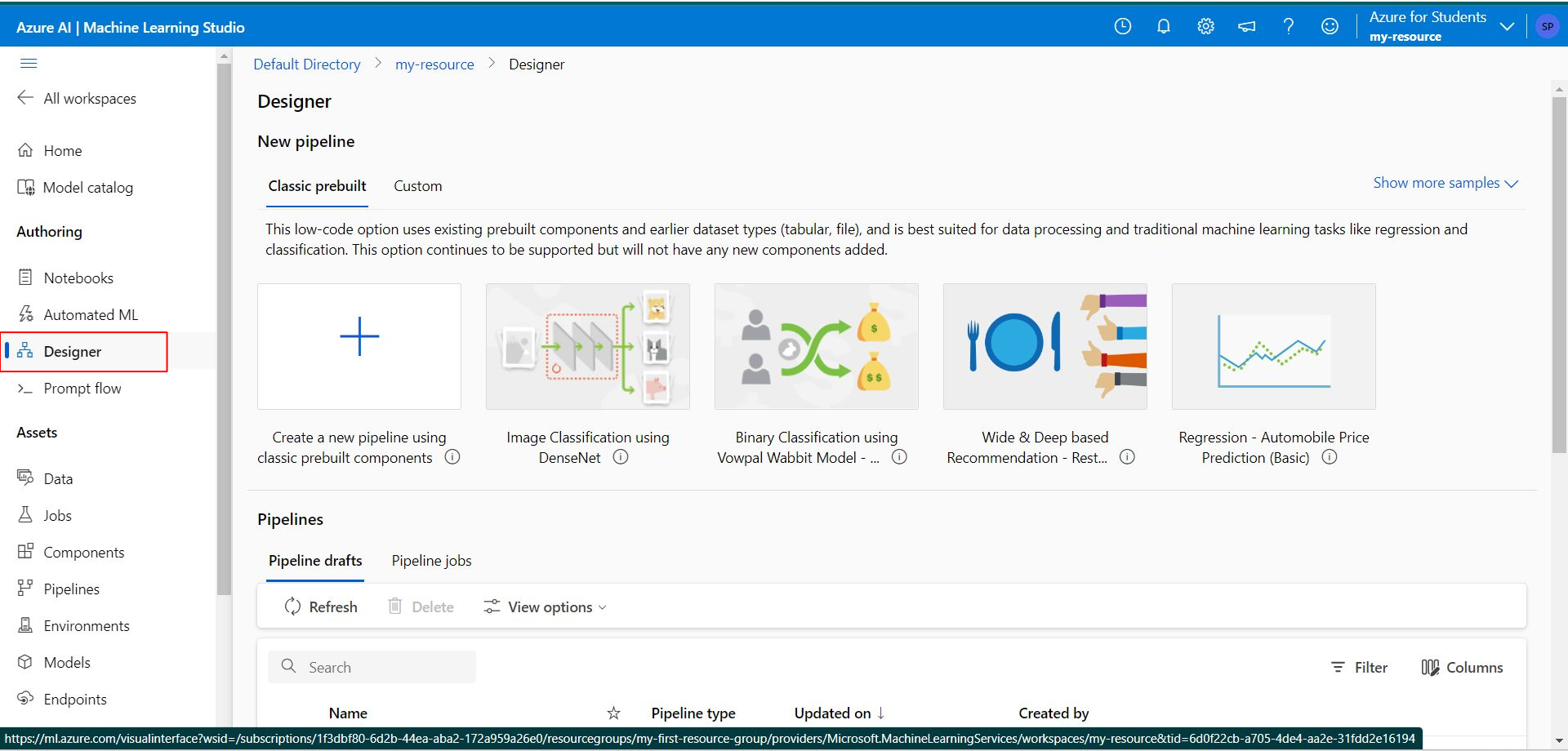
How to use the resource in a project?



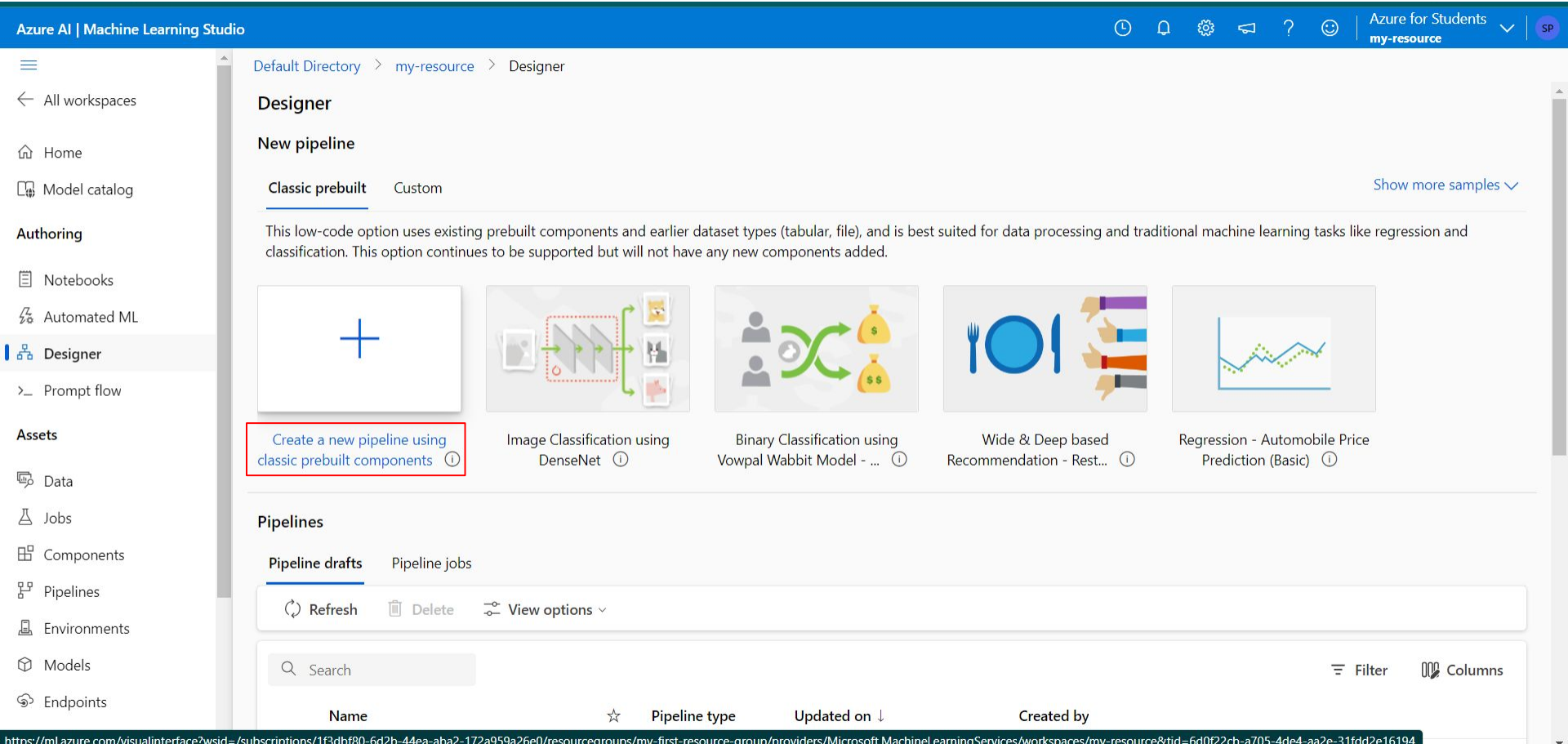
How to create a new ML job?



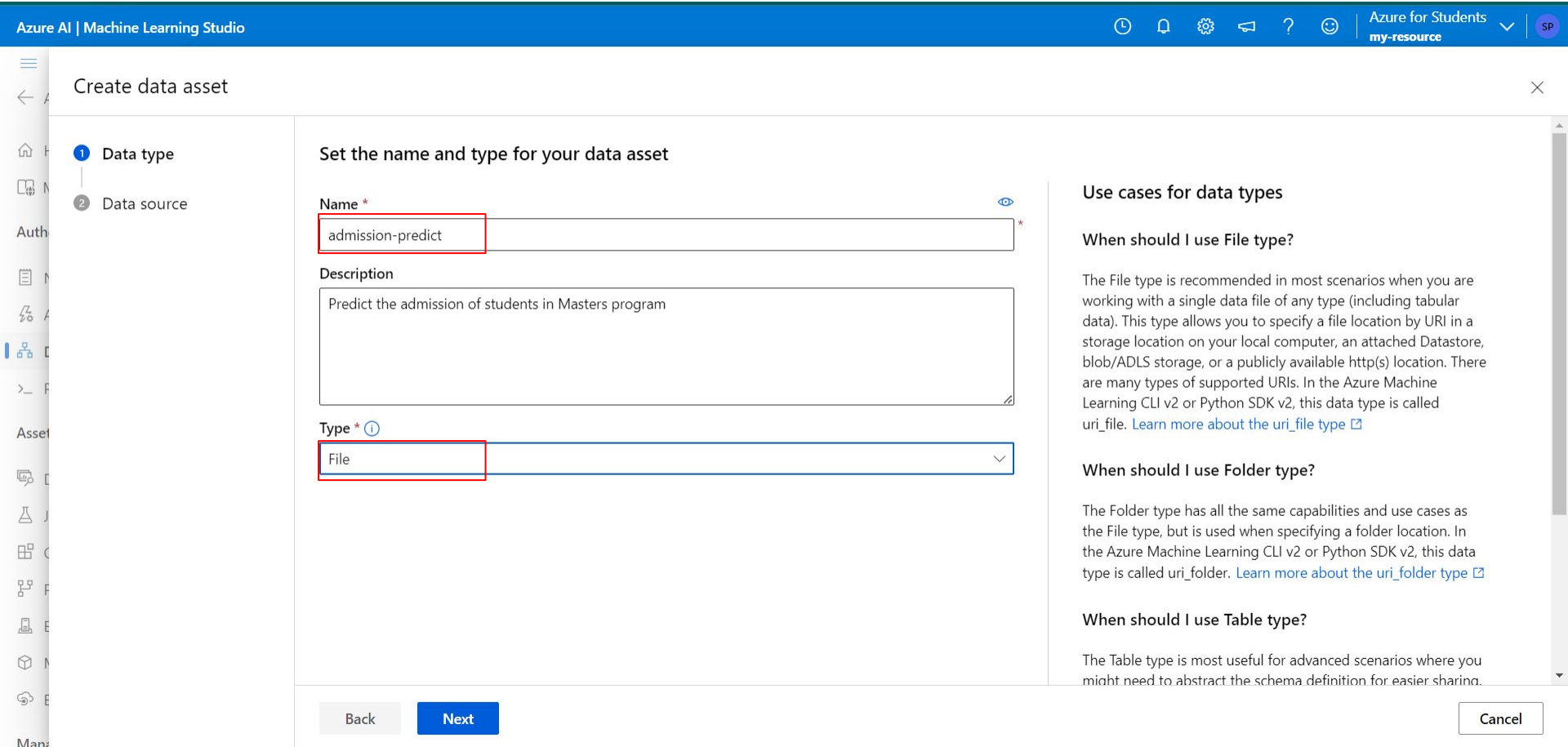
How to create a new ML job in Designer?



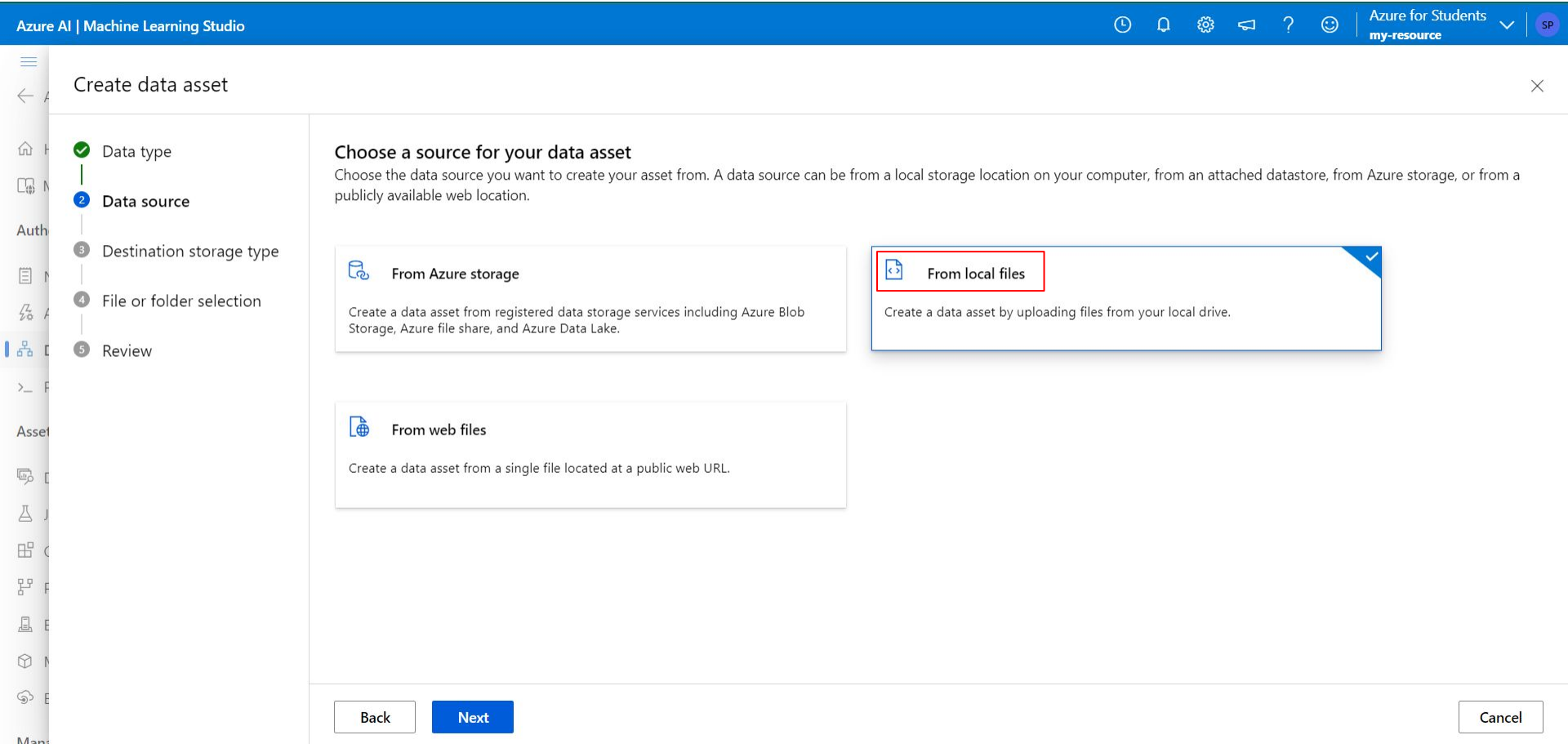
How to create a new ML job in Designer?



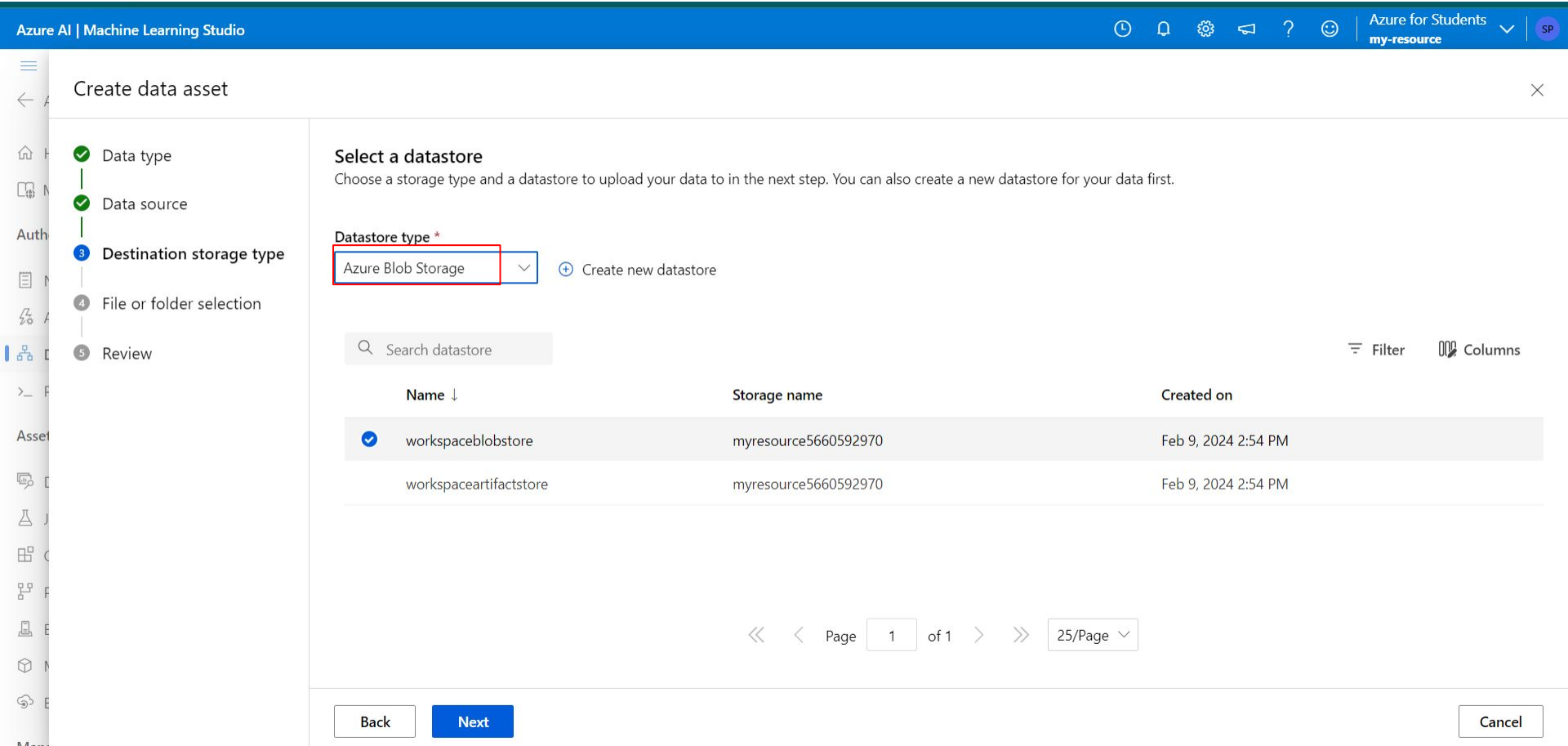
How to add a new dataset?



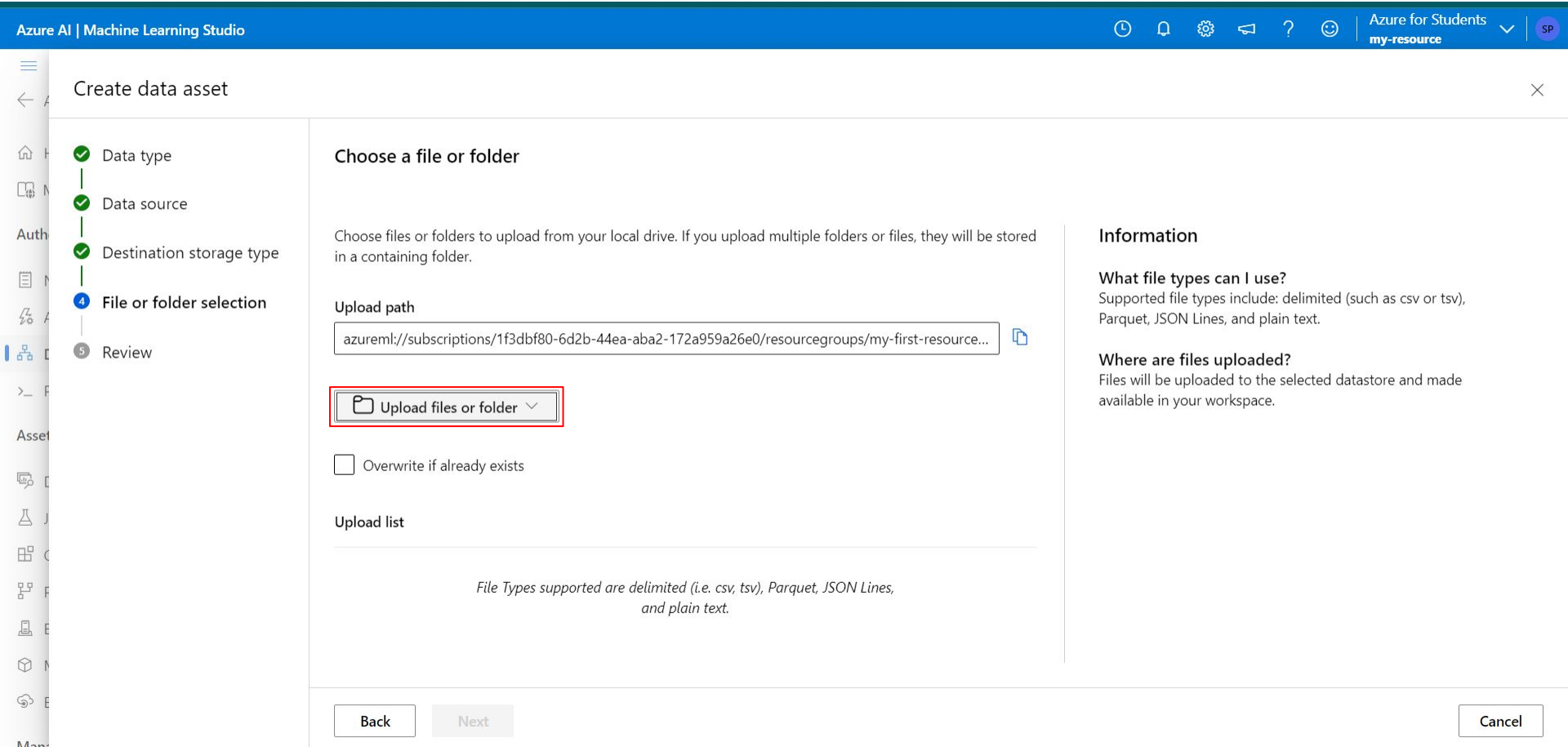
How to add a new dataset?



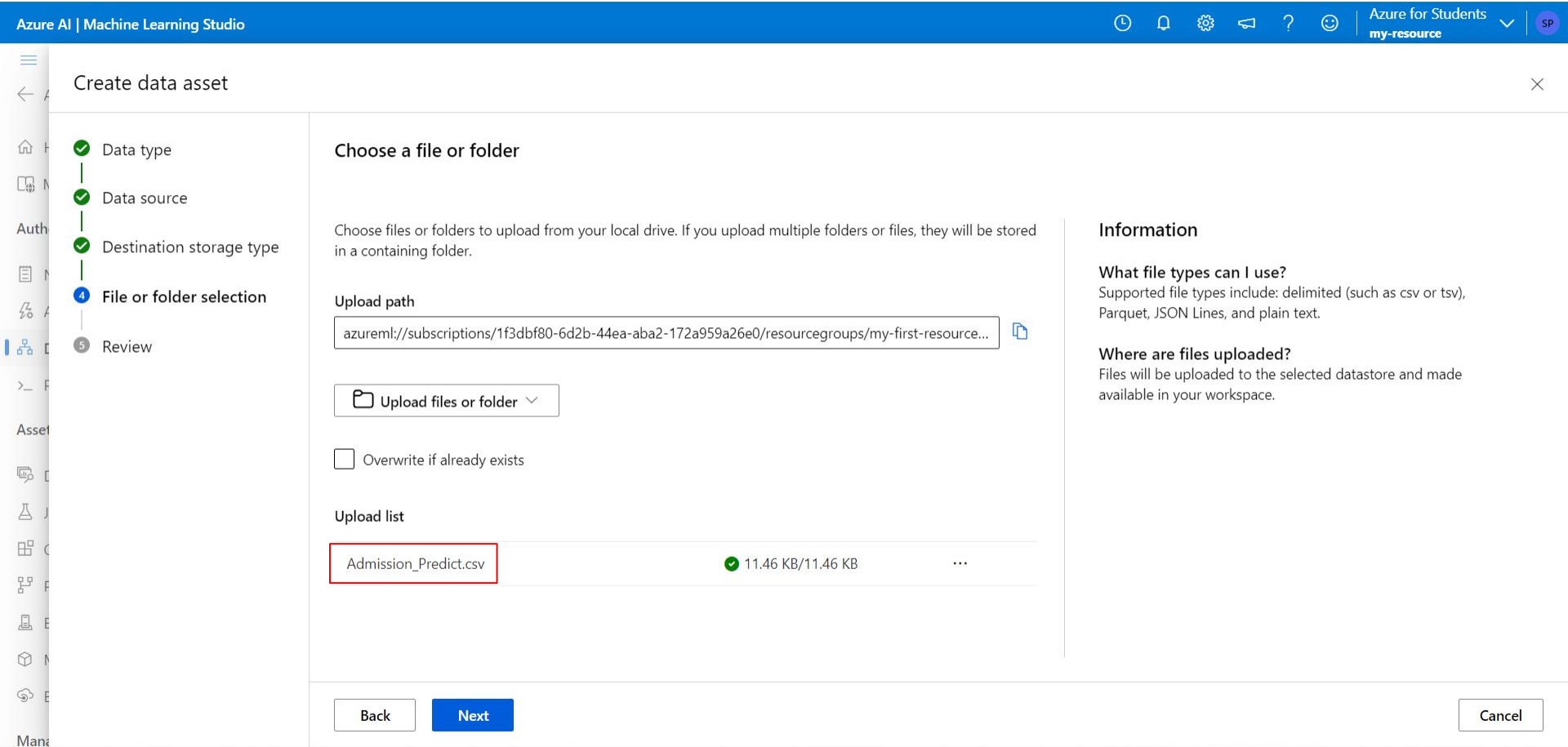
How to add a new dataset?



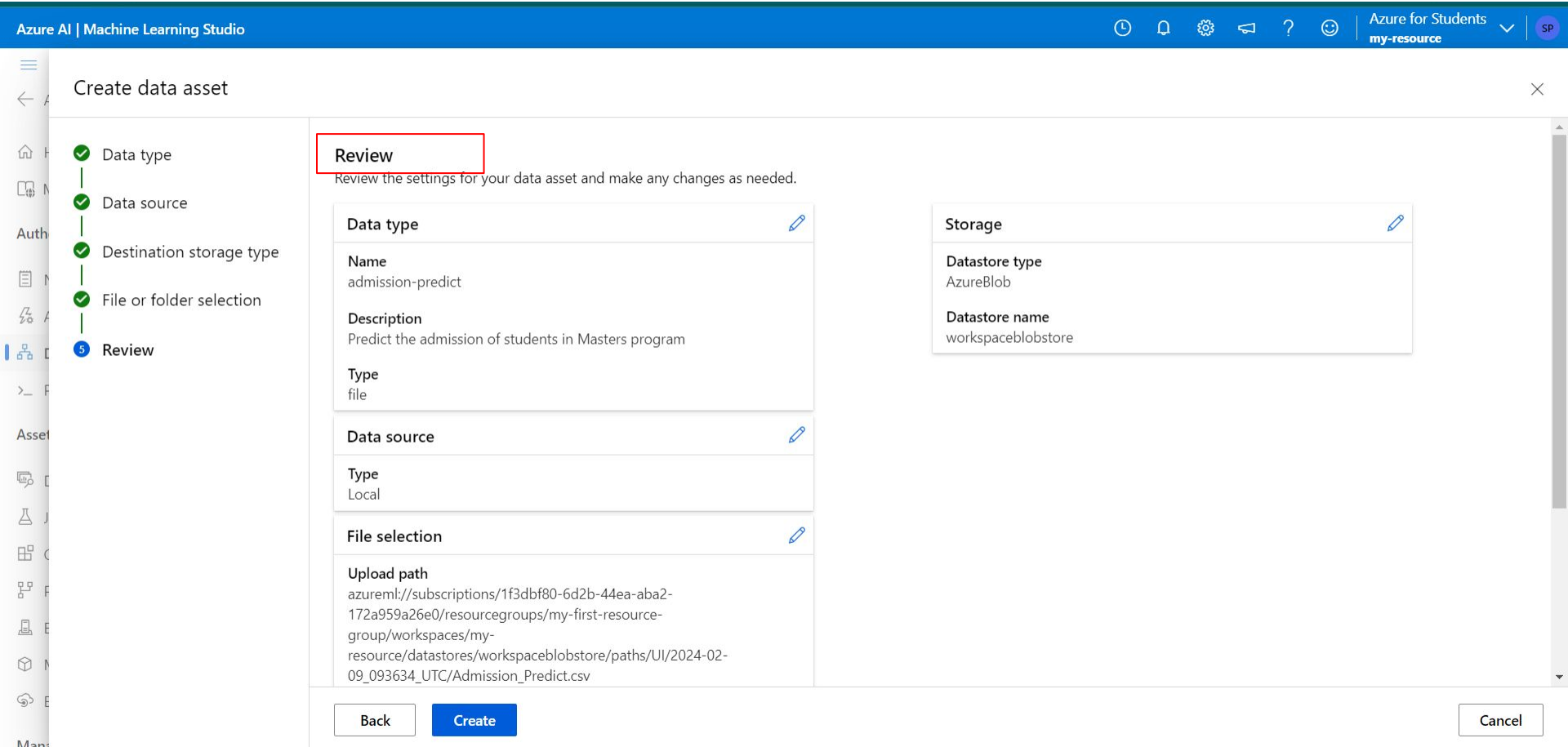
How to add a new dataset?



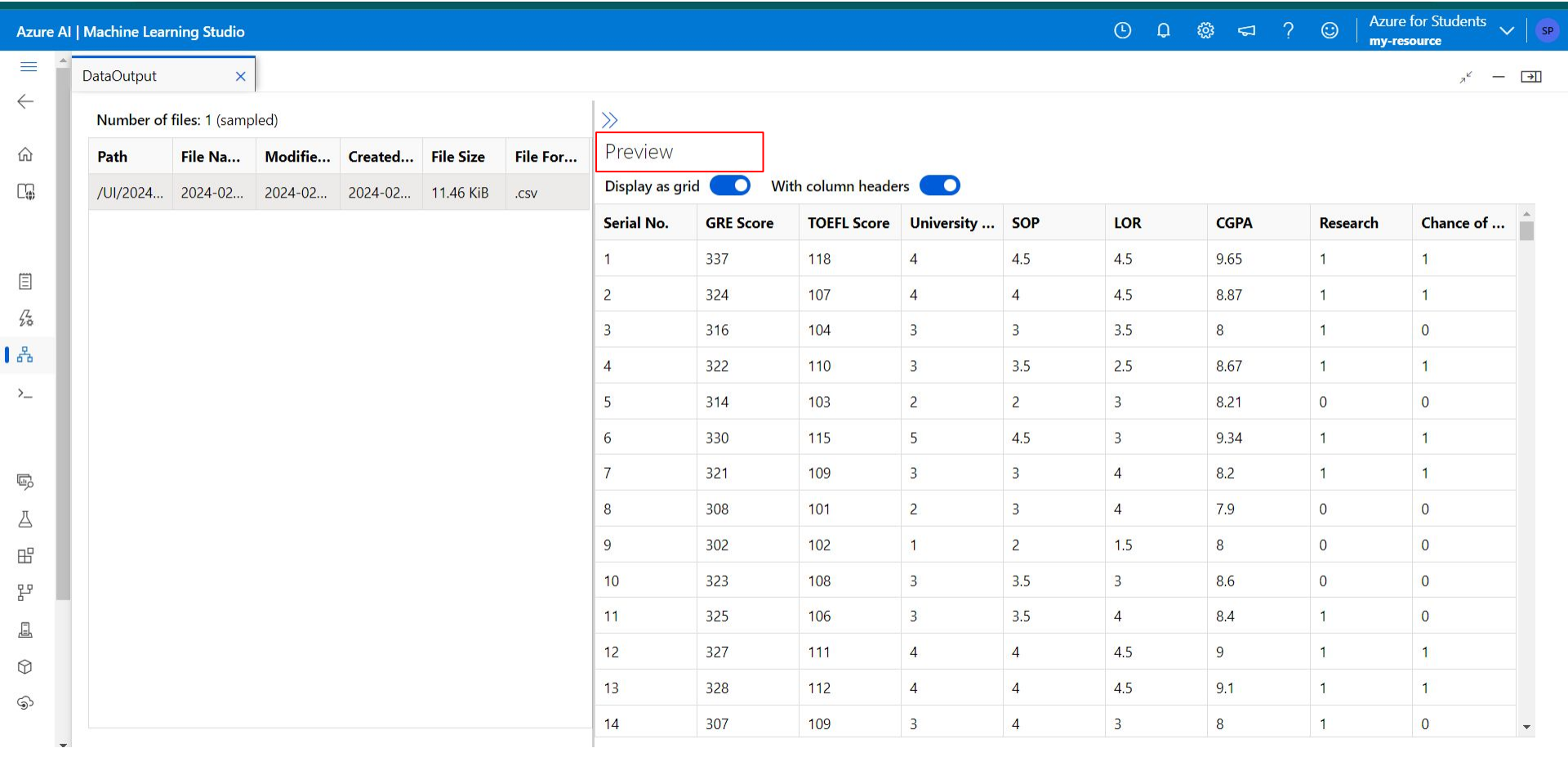
How to add a new dataset?



How to add a new dataset?



How to add a new dataset?



How to build a model pipeline?

1. A simple pipeline could be as follows:
    a. Data ingestion.
    b. Selecting the columns to be used for model-building by removing redundant columns, like Serial Number.
    c. Train-test split.
    d. Data pre-processing (say, cleaning missing data.)
    e. Training the algorithm on train dataset (say, decision tree classifier.)
    f. Scoring the model by feeding the test dataset.
    g. Evaluating the model.

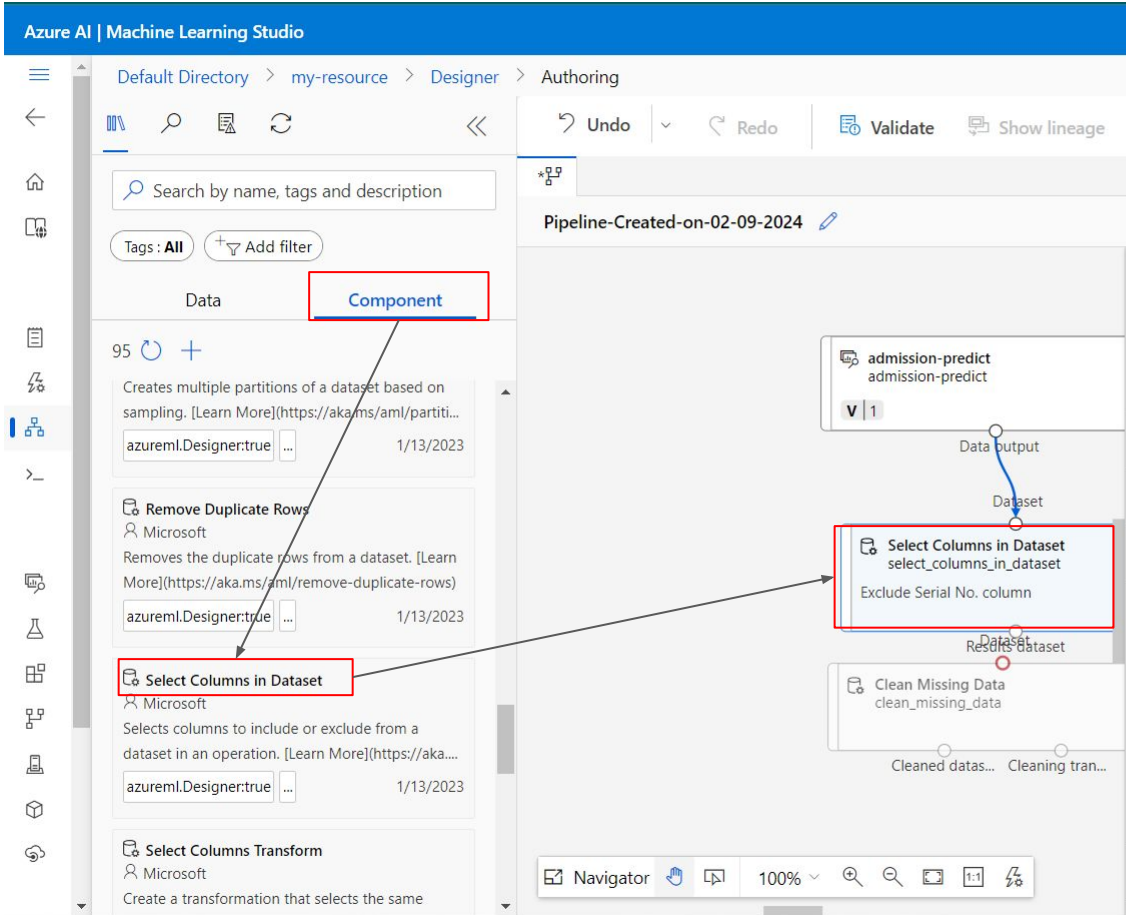
2. Proceed to build the pipeline as shown. Populate different stages of the pipeline by selecting the different components from the left panel, labelled ‘Component’.

How to build a model pipeline?

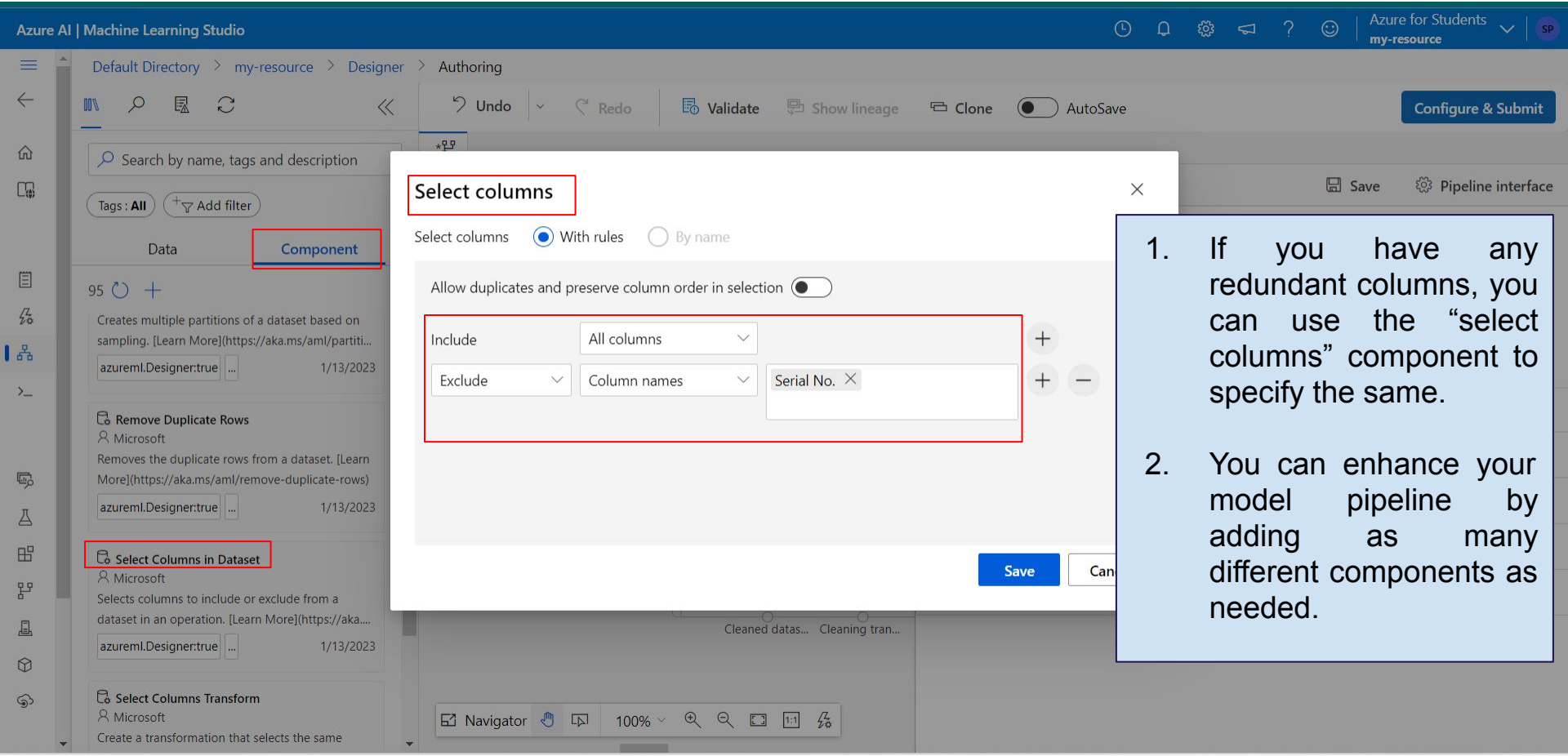
1. After the dataset has been uploaded, create the pipeline using the Designer tab under Authoring section.

2. Add different components like dataset, select columns and clean missing data, by selecting the respective components from the ‘Component’ tab on the left of the Panel.

3. Proceed to build the pipeline as shown.

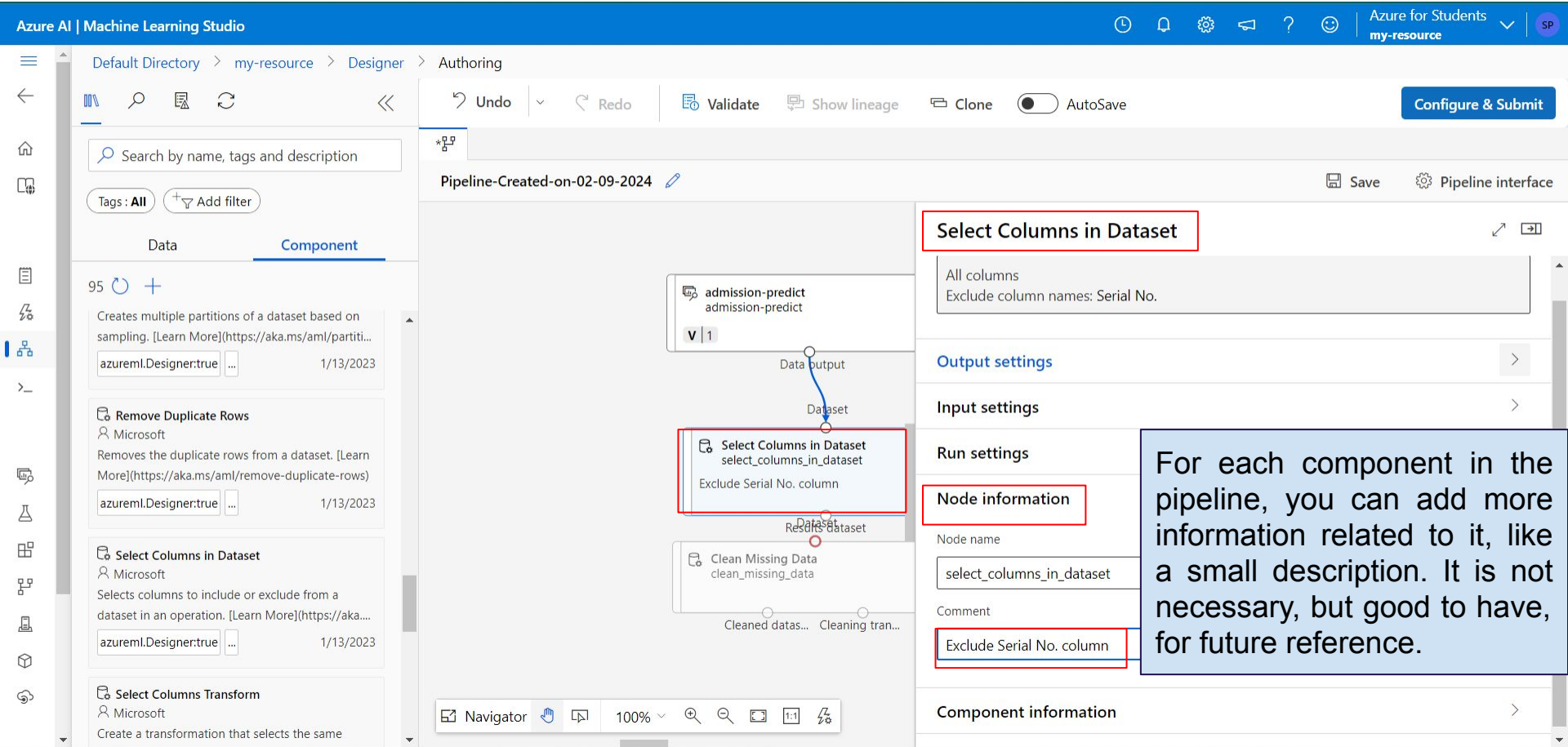


How to build a model pipeline?

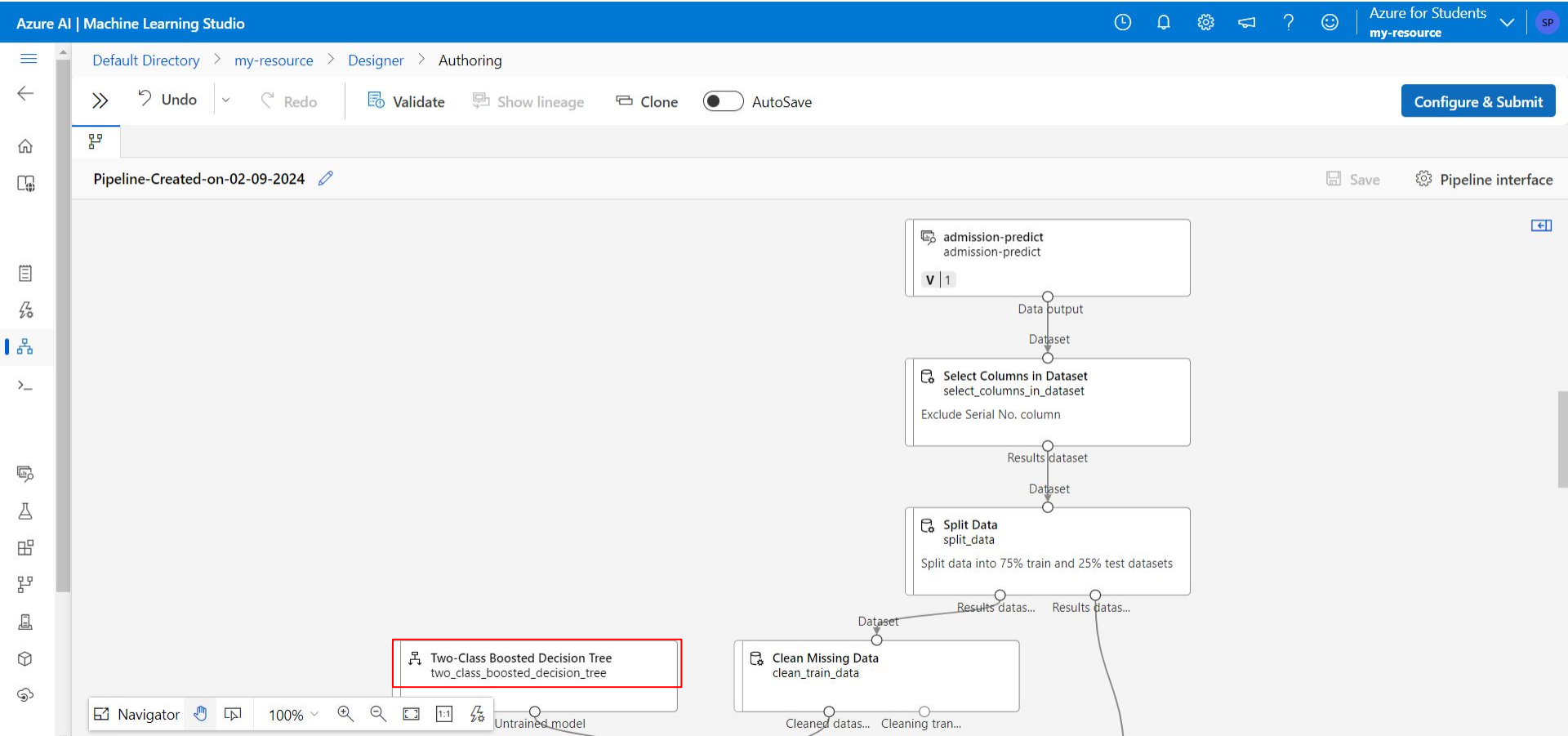


1. If you have any redundant columns, you can use the “select columns” component to specify the same.
2. You can enhance your model pipeline by adding as many different components as needed.

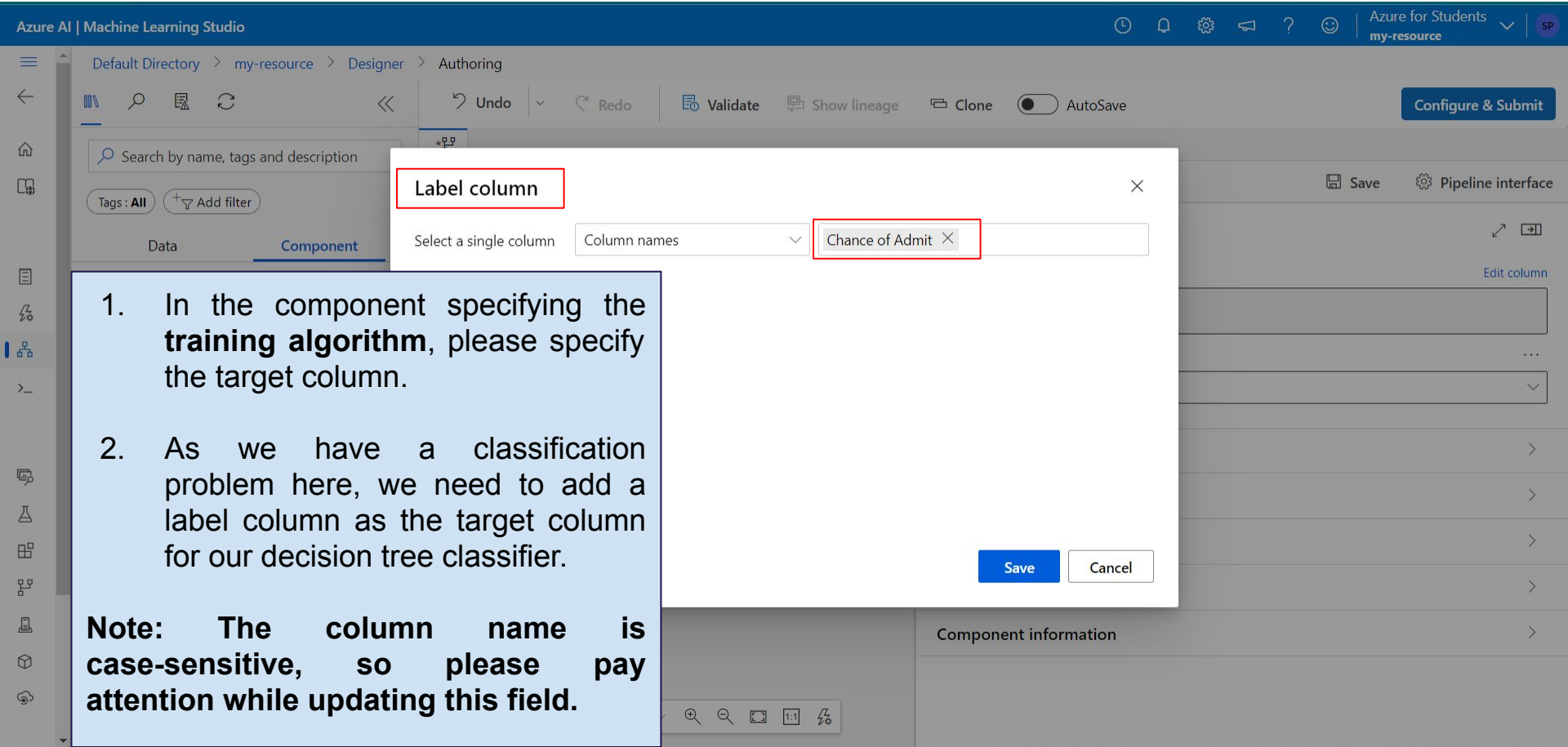
How to build a model pipeline?



How to build a model pipeline?



How to build a model pipeline?

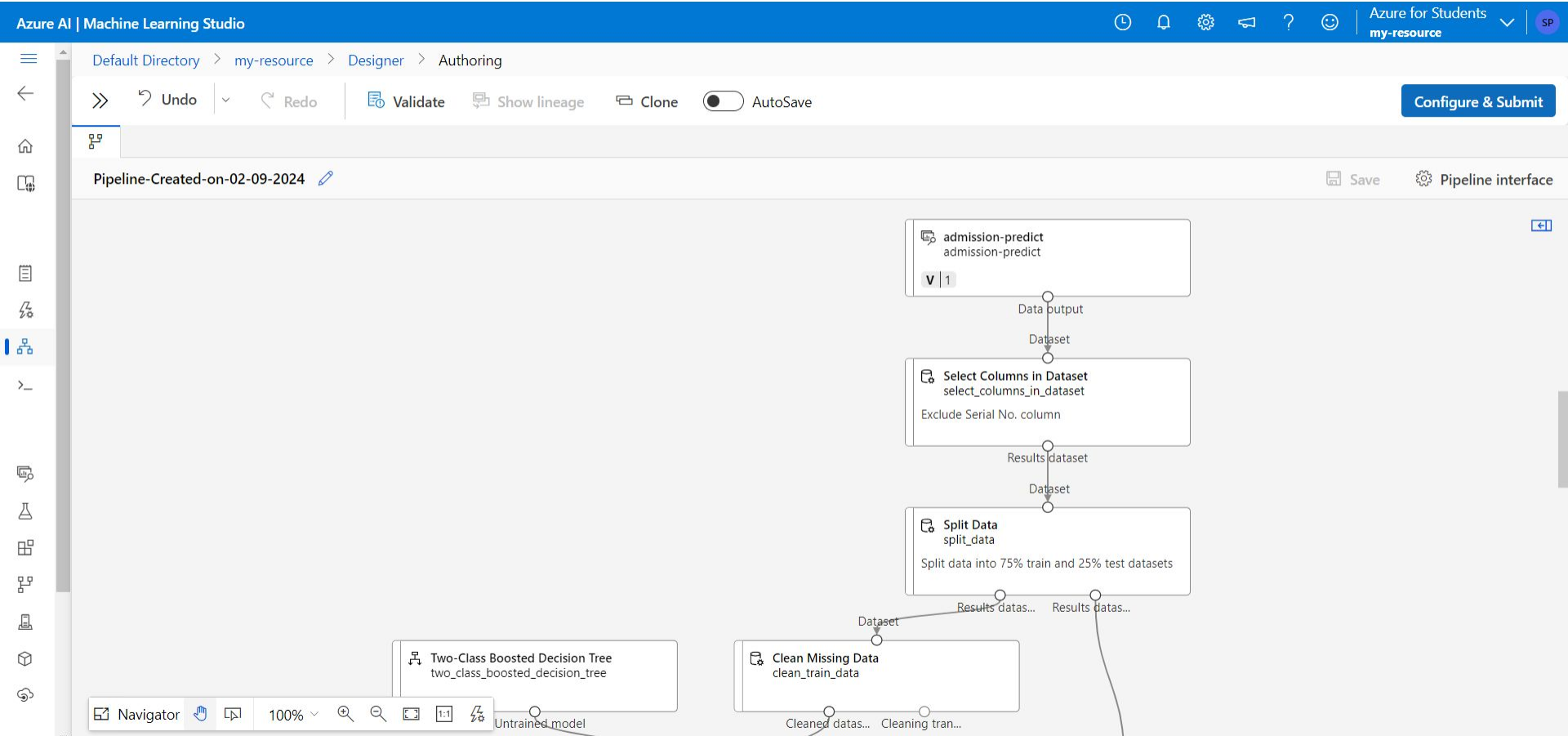


1. In the component specifying the **training algorithm**, please specify the target column.

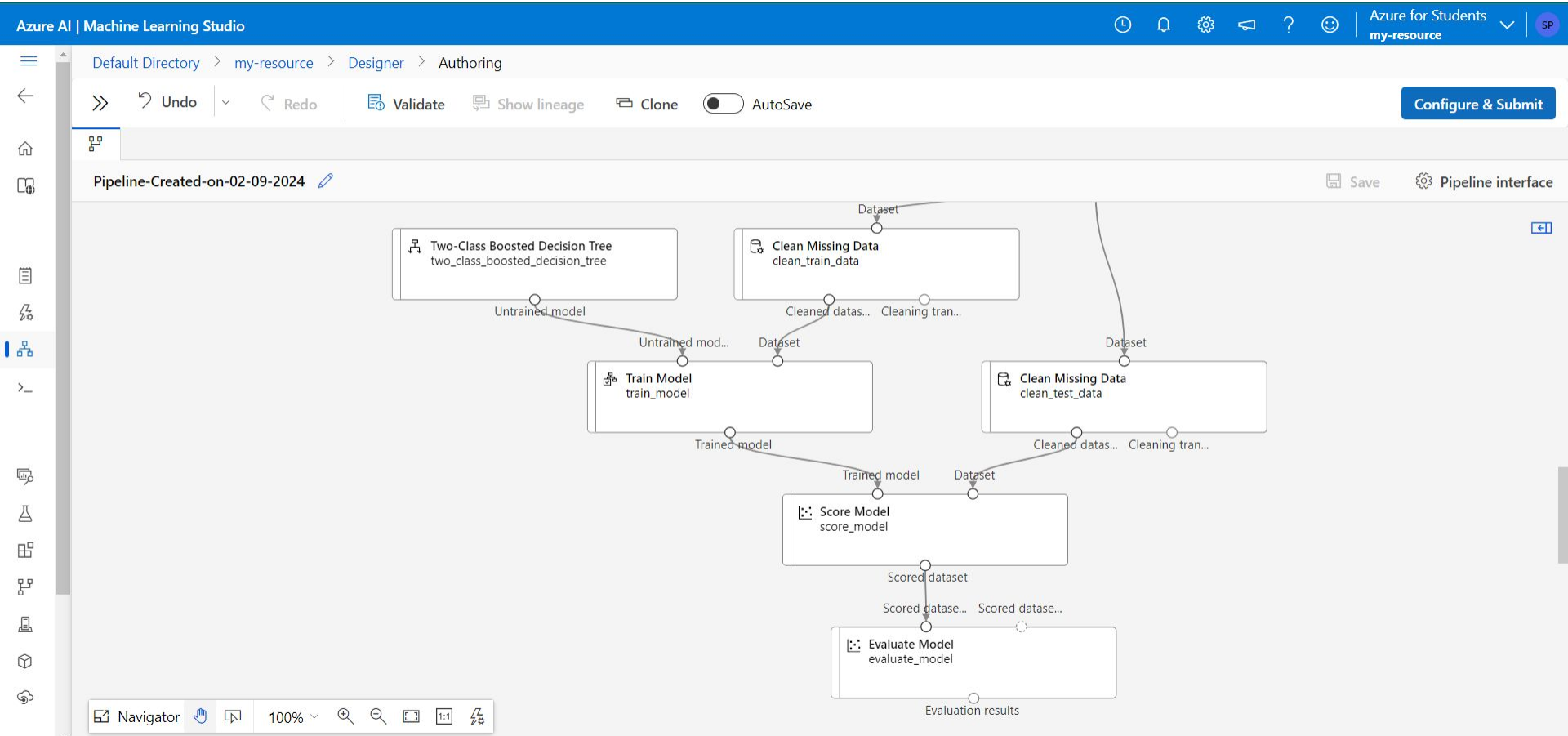
2. As we have a classification problem here, we need to add a label column as the target column for our decision tree classifier.

Note: The column name is case-sensitive, so please pay attention while updating this field.

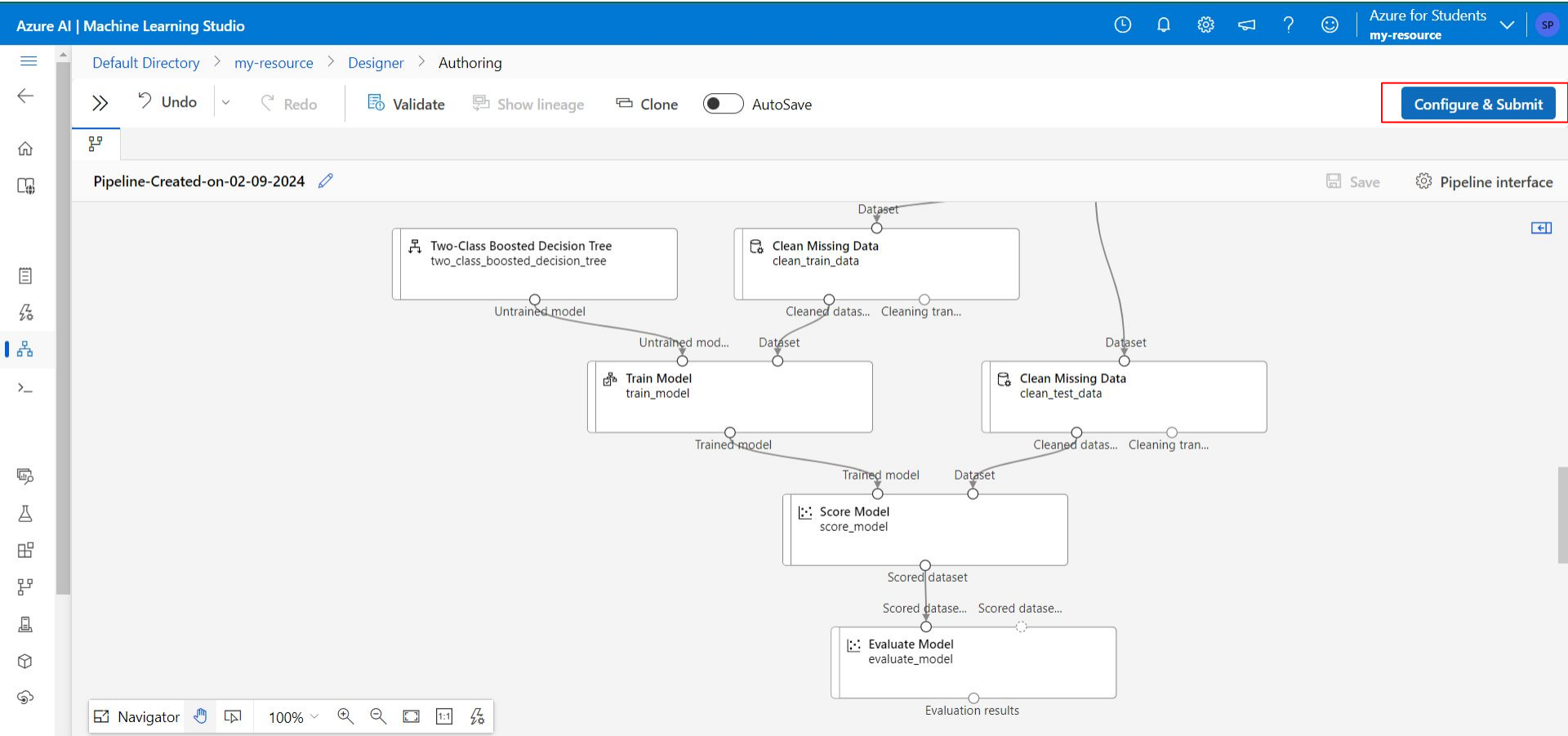
Model Pipeline



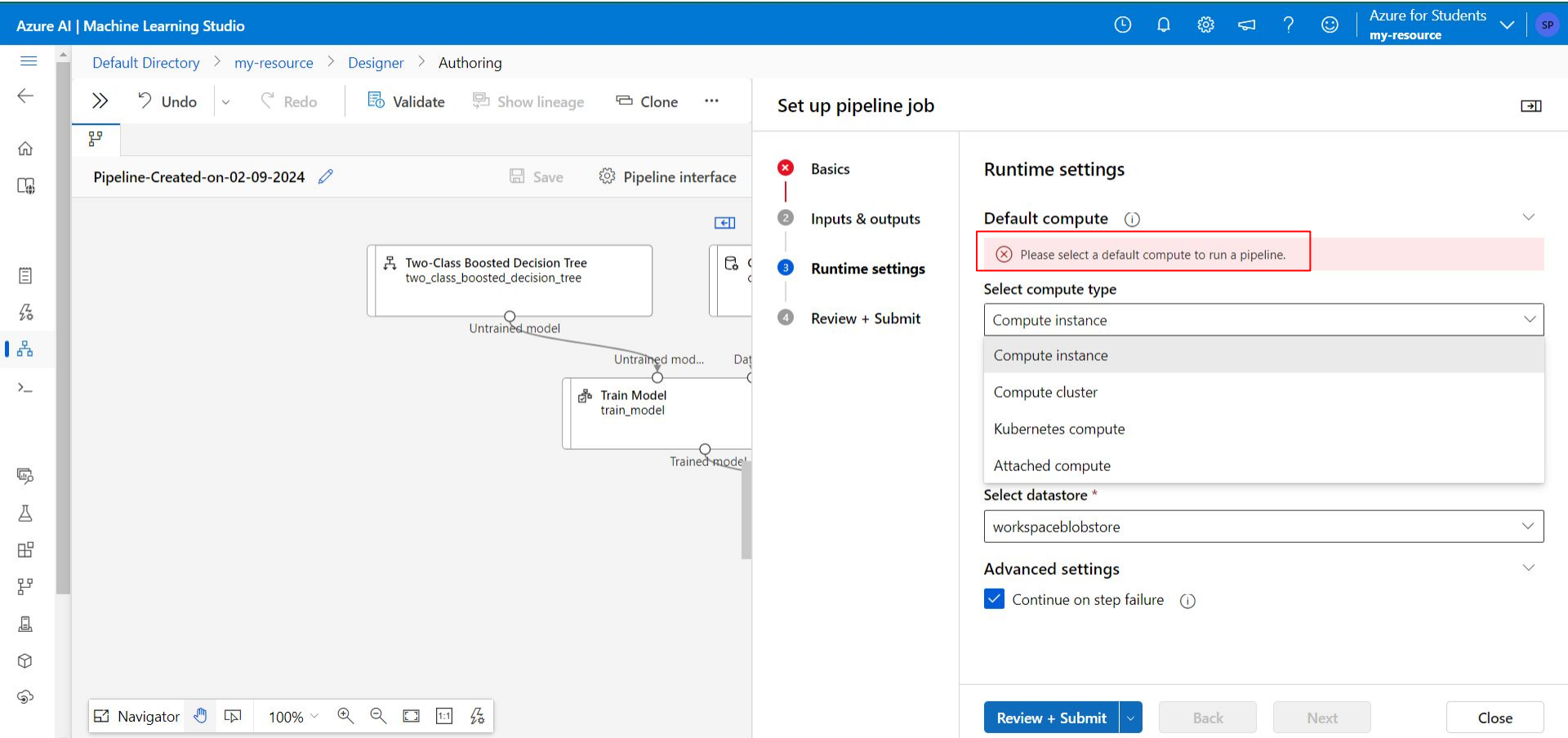
Model Pipeline



How to run the model pipeline?



How to run a model pipeline?



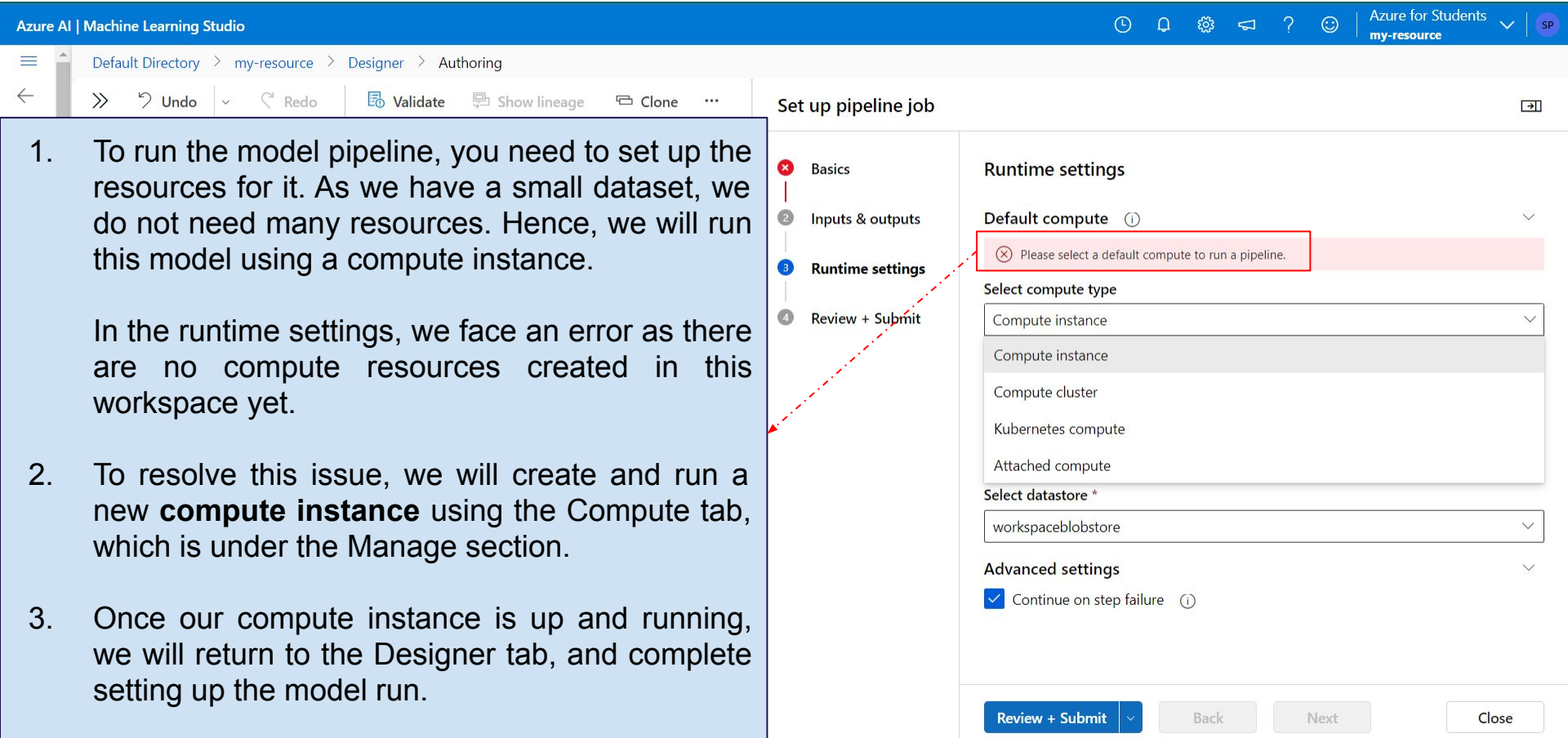
How to run a model pipeline?

1. To run the model pipeline, you need to set up the resources for it. As we have a small dataset, we do not need many resources. Hence, we will run this model using a compute instance.

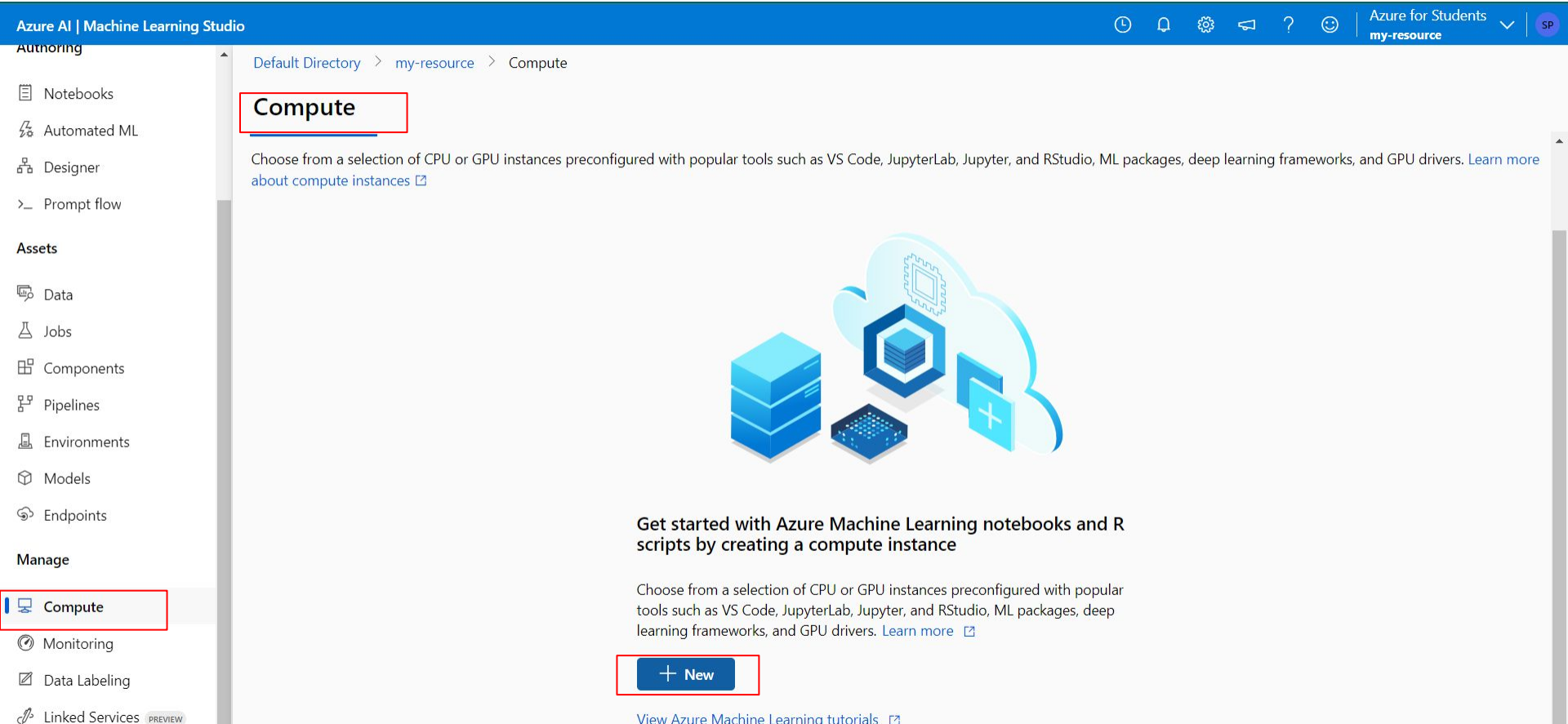
   In the runtime settings, we face an error as there are no compute resources created in this workspace yet.

2. To resolve this issue, we will create and run a new **compute instance** using the Compute tab, which is under the Manage section.

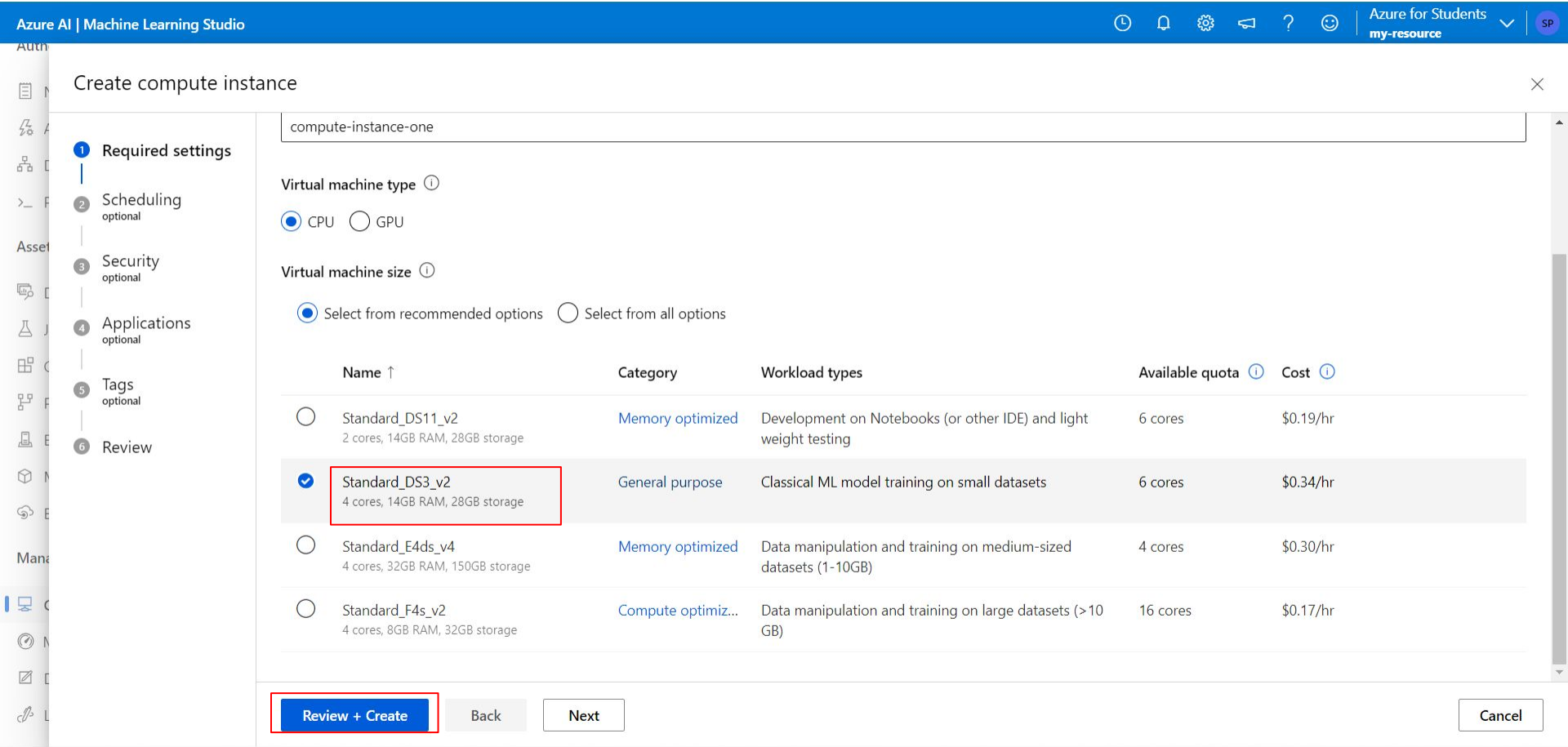
3. Once our compute instance is up and running, we will return to the Designer tab, and complete setting up the model run.



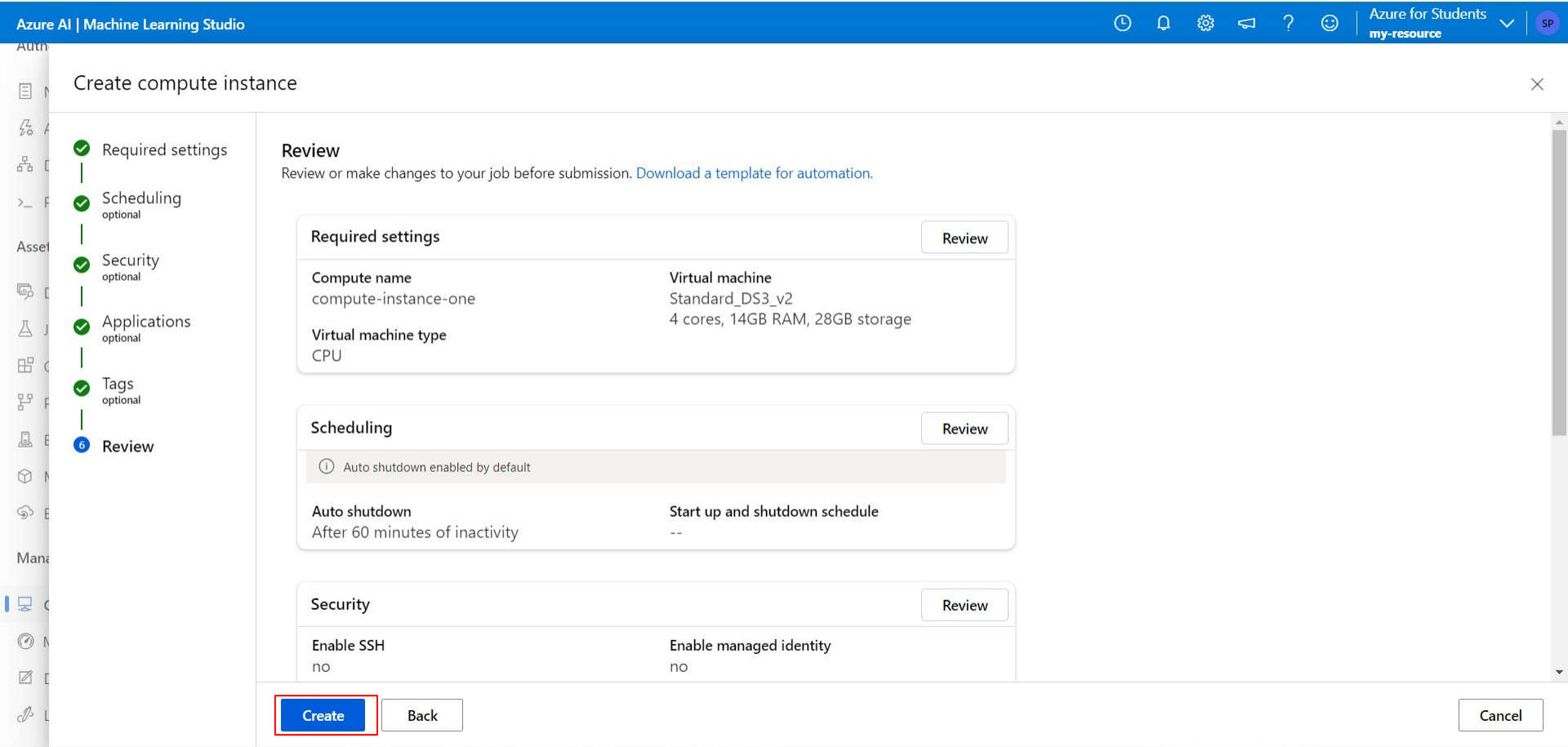
How to set up the compute resource?



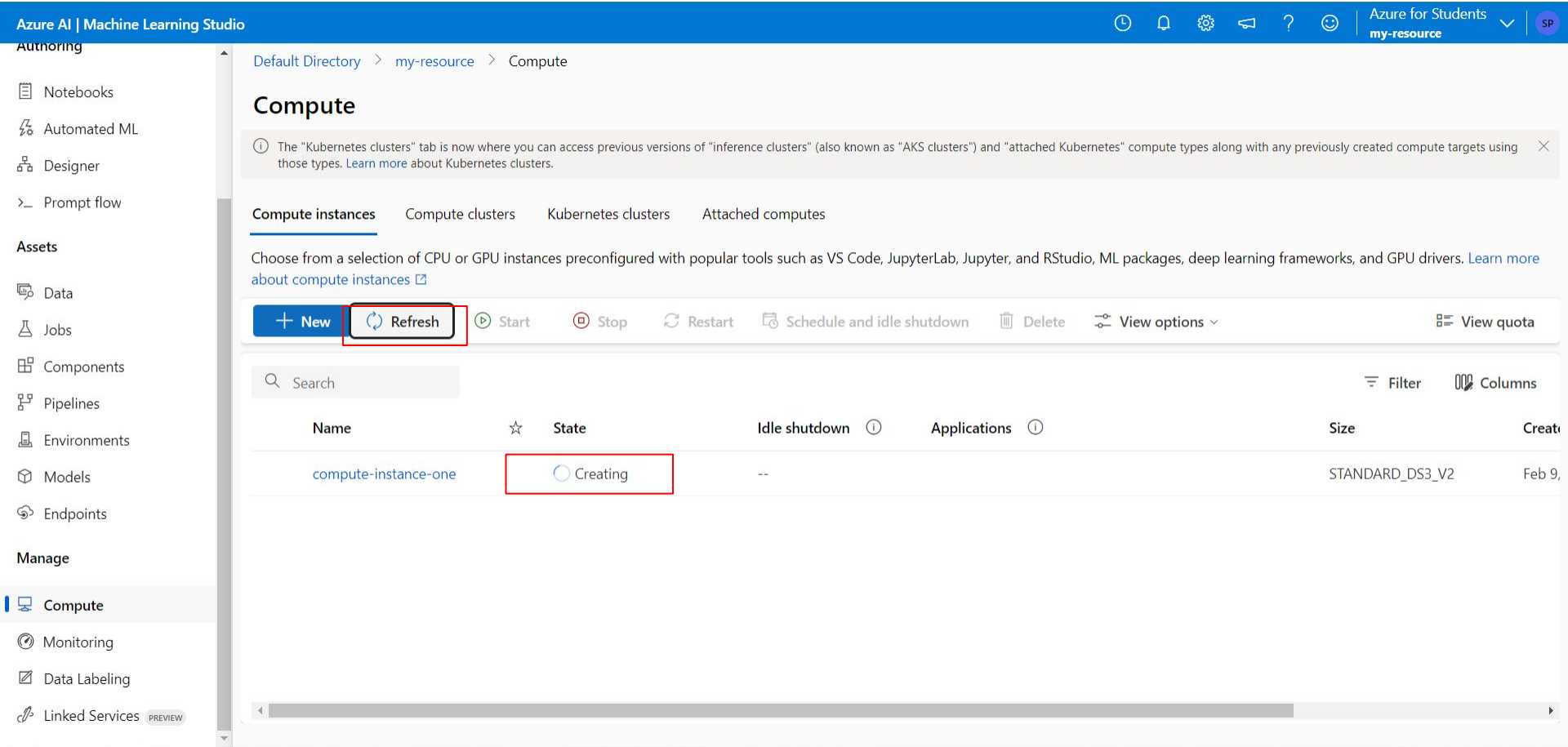
How to set up the compute resource?



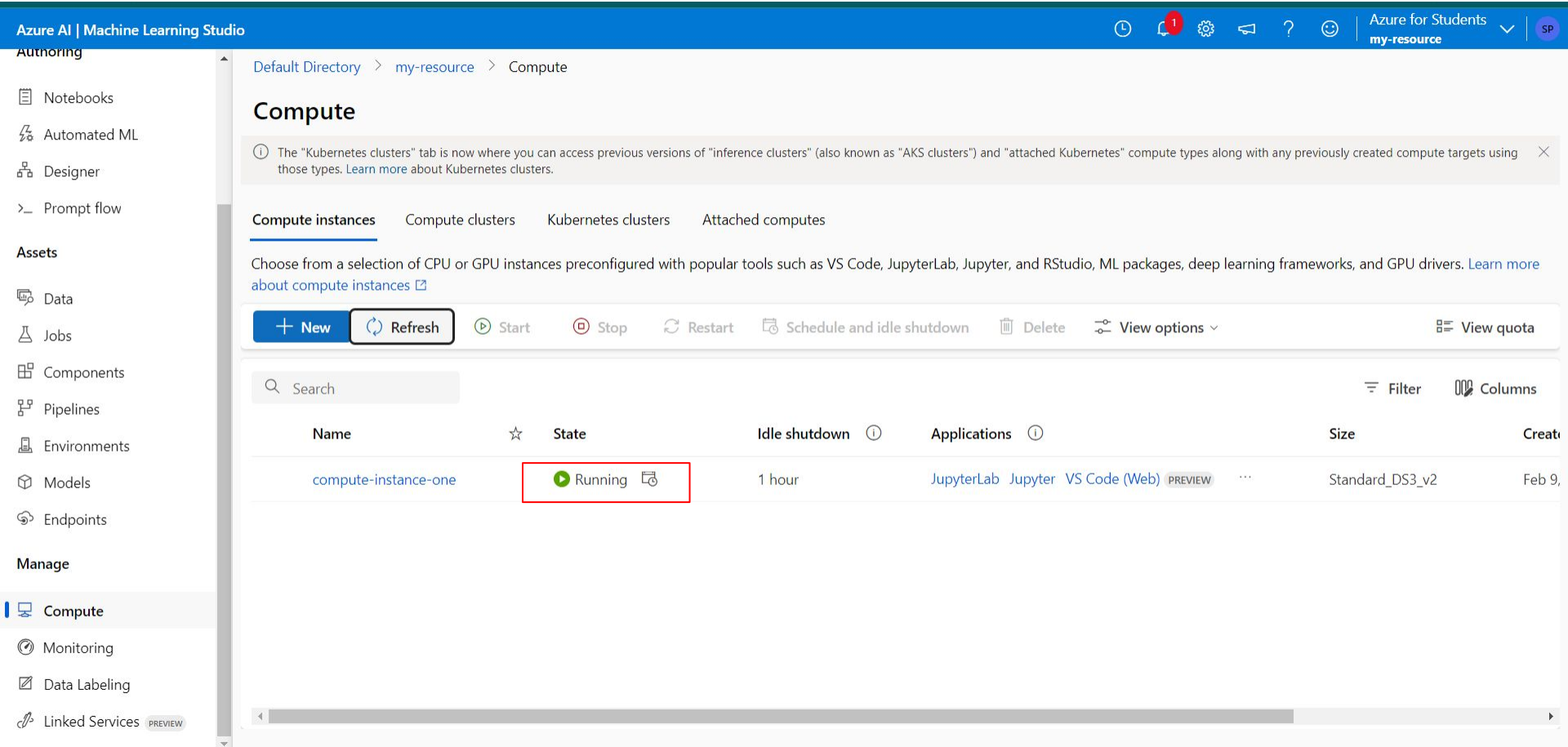
How to set up the compute resource?



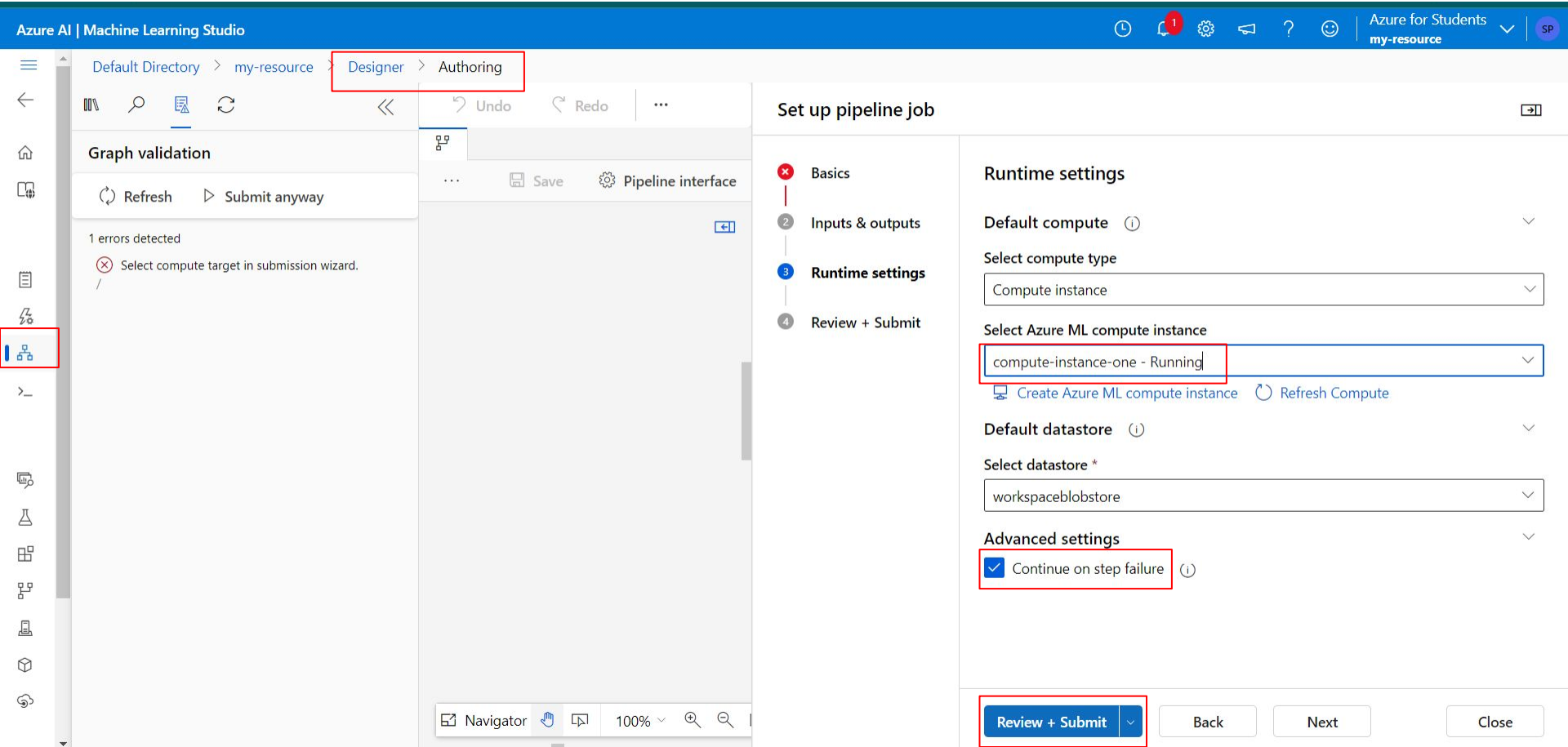
How to set up the compute resource?



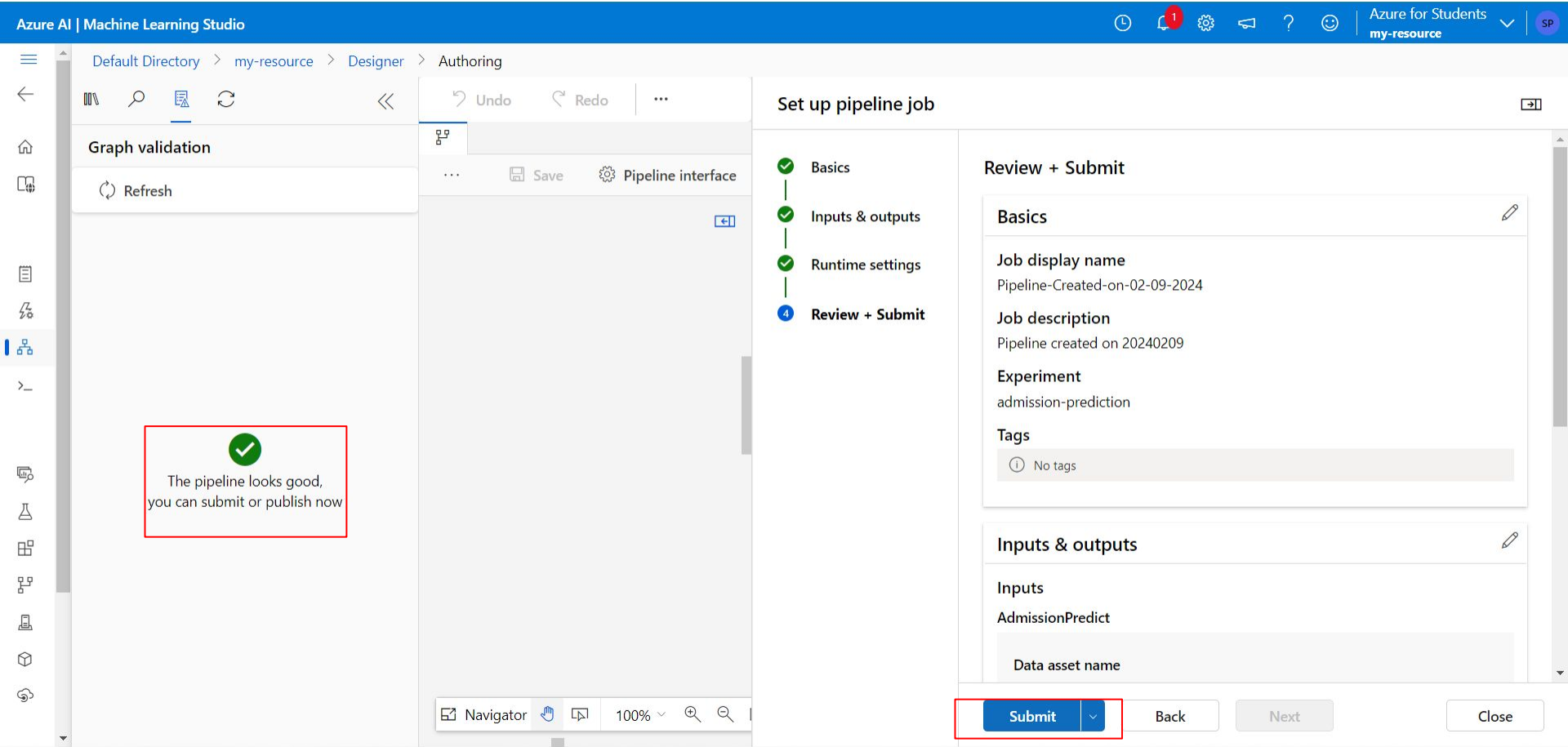
How to set up the compute resource?



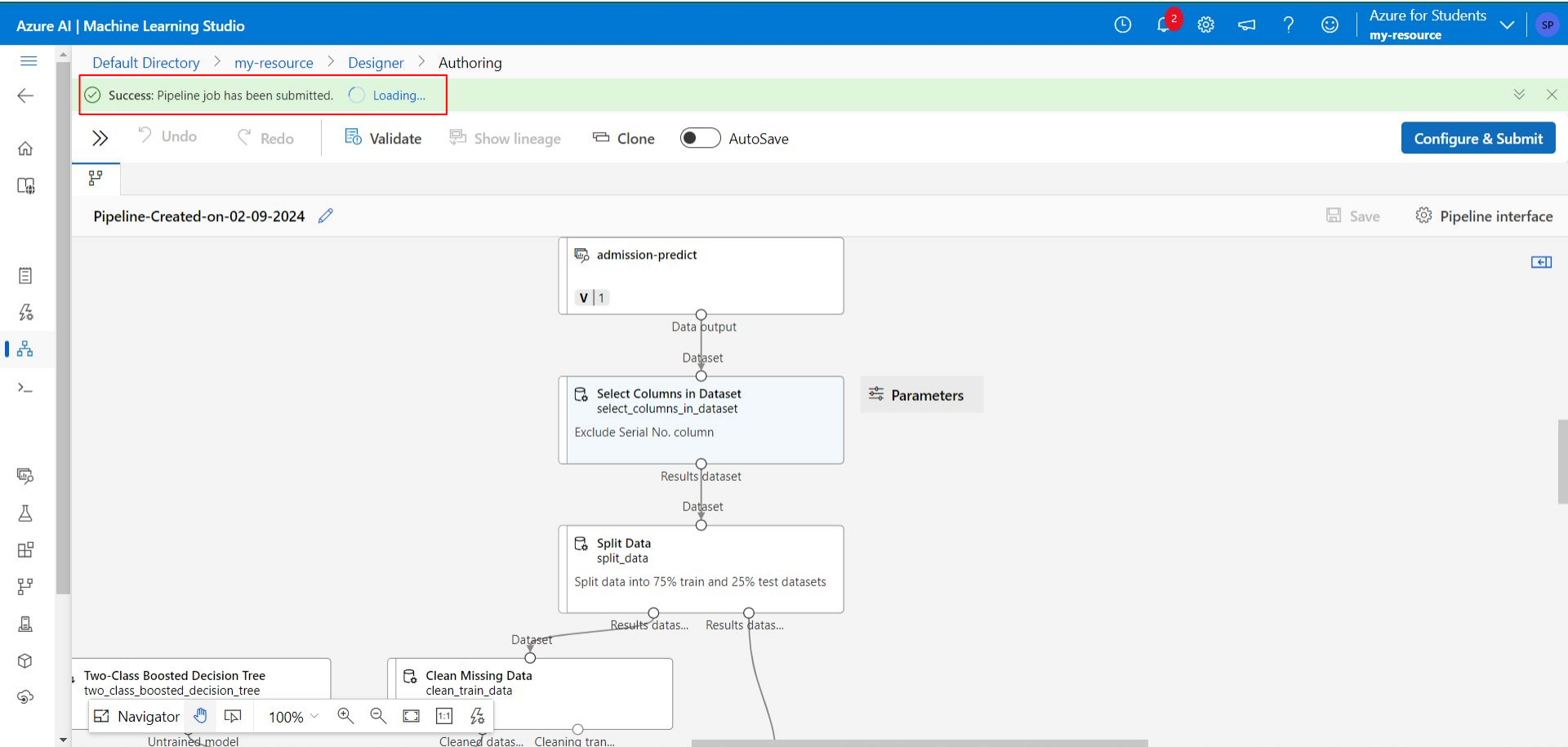
How to run a model pipeline?



How to run a model pipeline?



How to run a model pipeline?



How to check the status of model run?

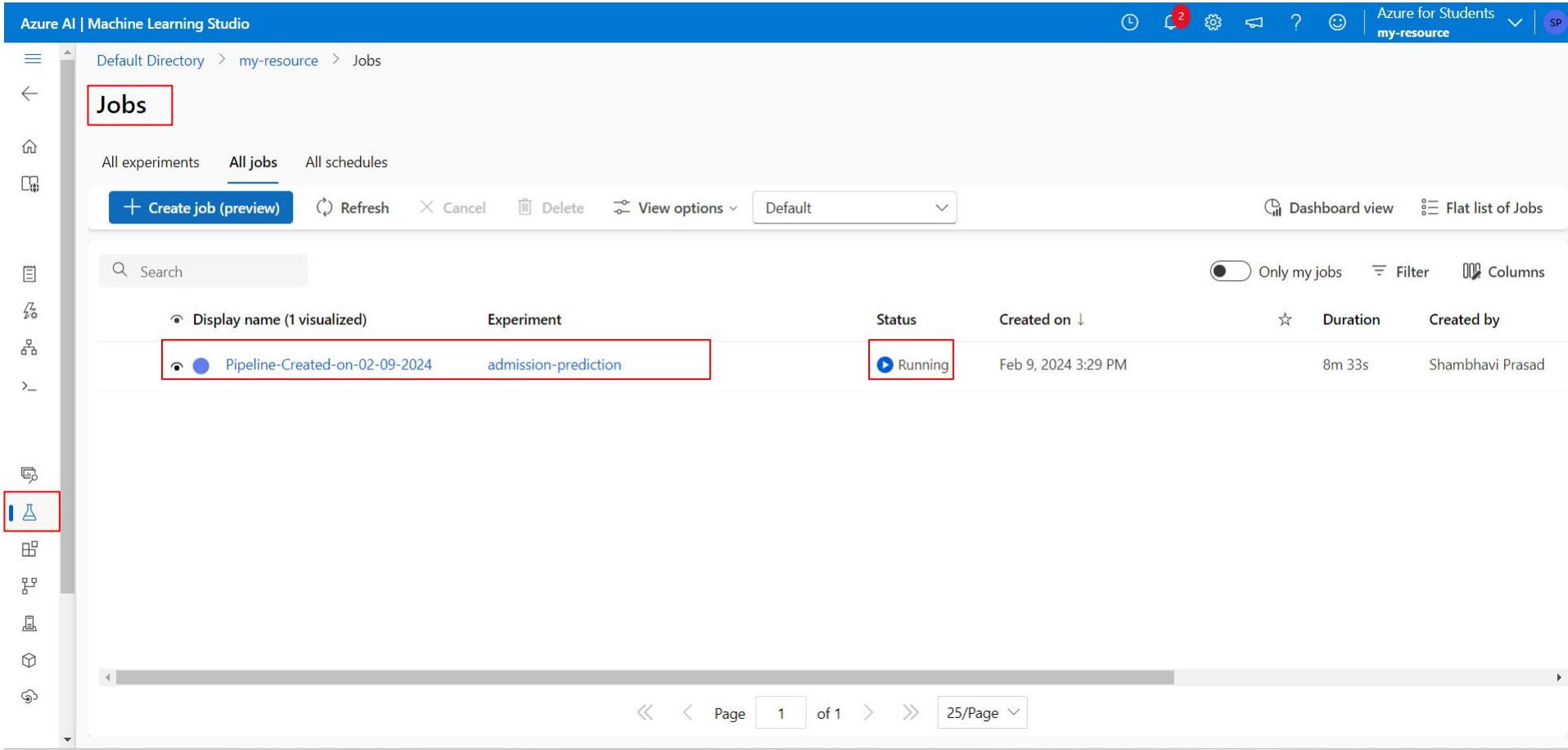
1. After you have started running the model, you can check its status by going to the Jobs tab under the Assets section.

2. Here, you will see all the jobs that have been submitted for execution, along with the respective experiment that they are a part of. In this case, the name of the experiment is ‘admission-prediction’, and the name of the job is the pipeline name that we have just created.

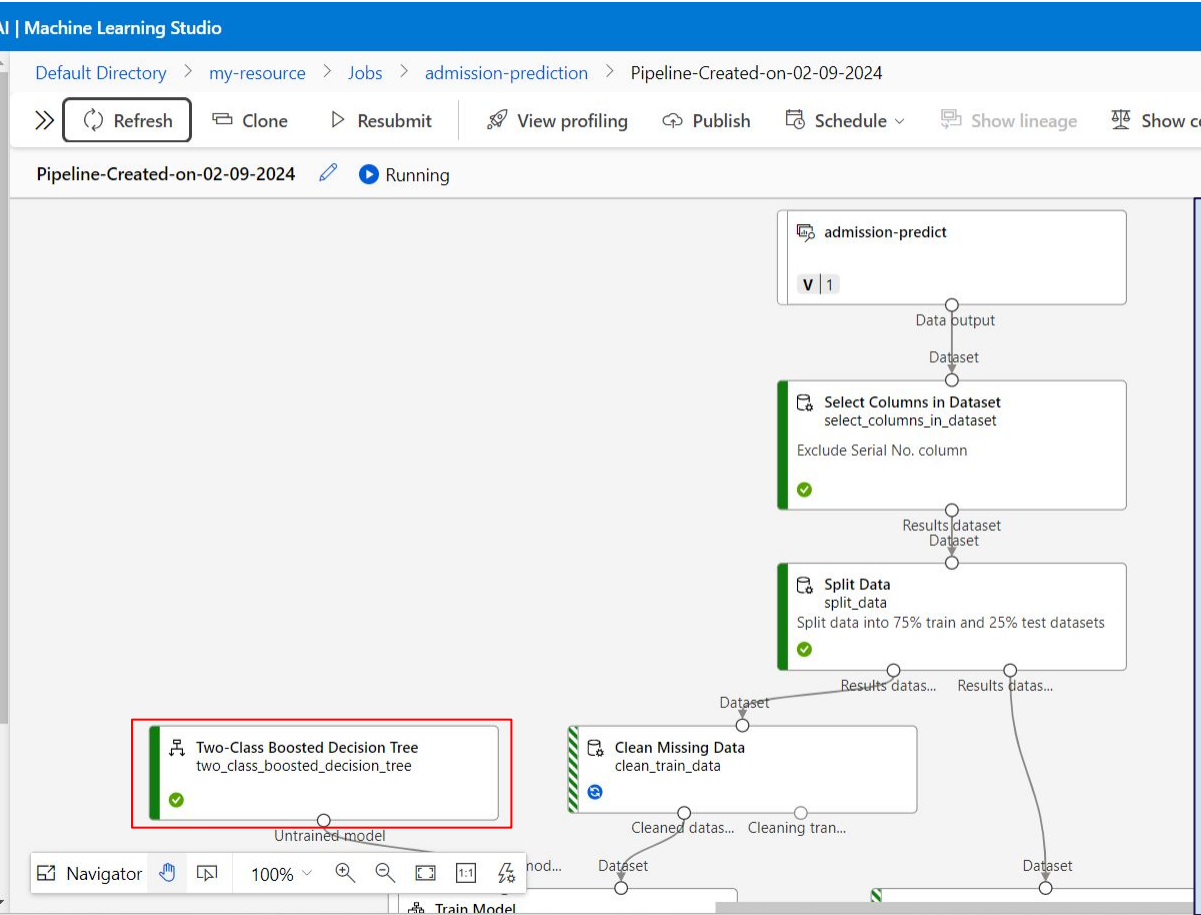
3. Under the Jobs tab, you can select your desired job, to view its progress. You can also see the related metrics of the model post its execution.

4. By clicking on the job name, it will redirect you to the Models tab, and show the component being run currently.

How to view progress of model run?



How to view progress of model run?

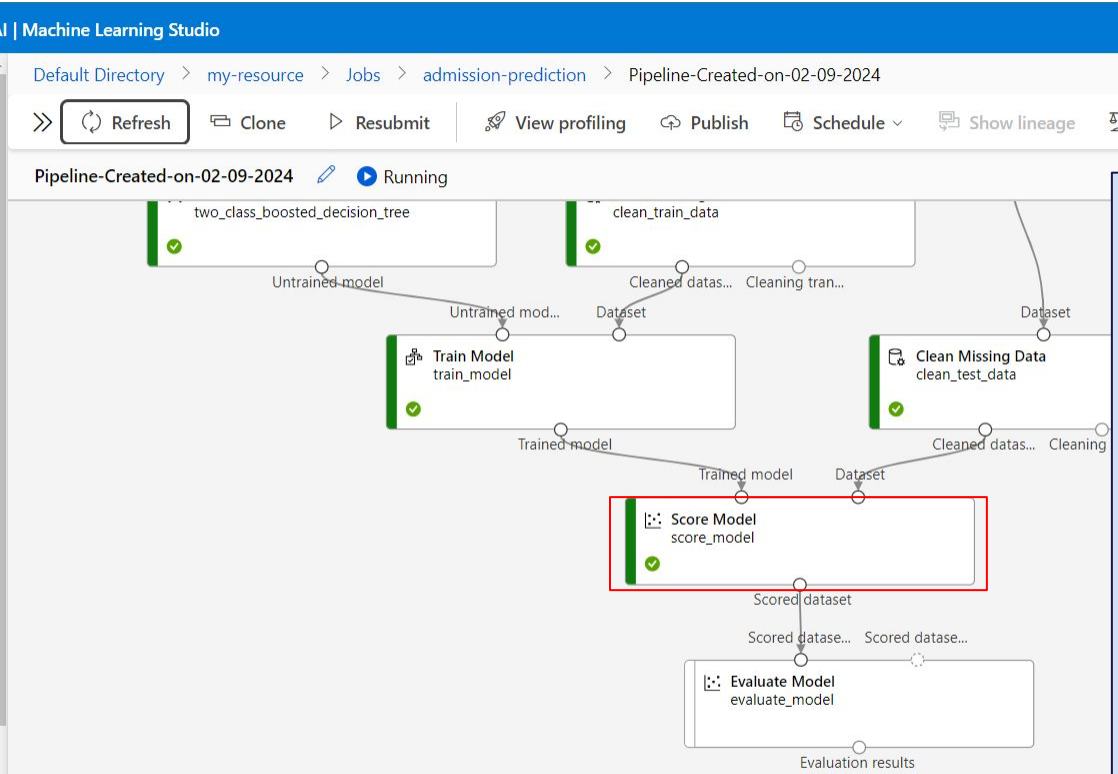


1. In the model pipeline, the solid green line indicates that the particular component has completed execution. A dashed line indicates that the process is still running.

2. At the end of a successful run, all the model components should have a solid green line.

3. You can refresh the model status at any time by clicking at the refresh button at the top-left.

How to view progress of model run?

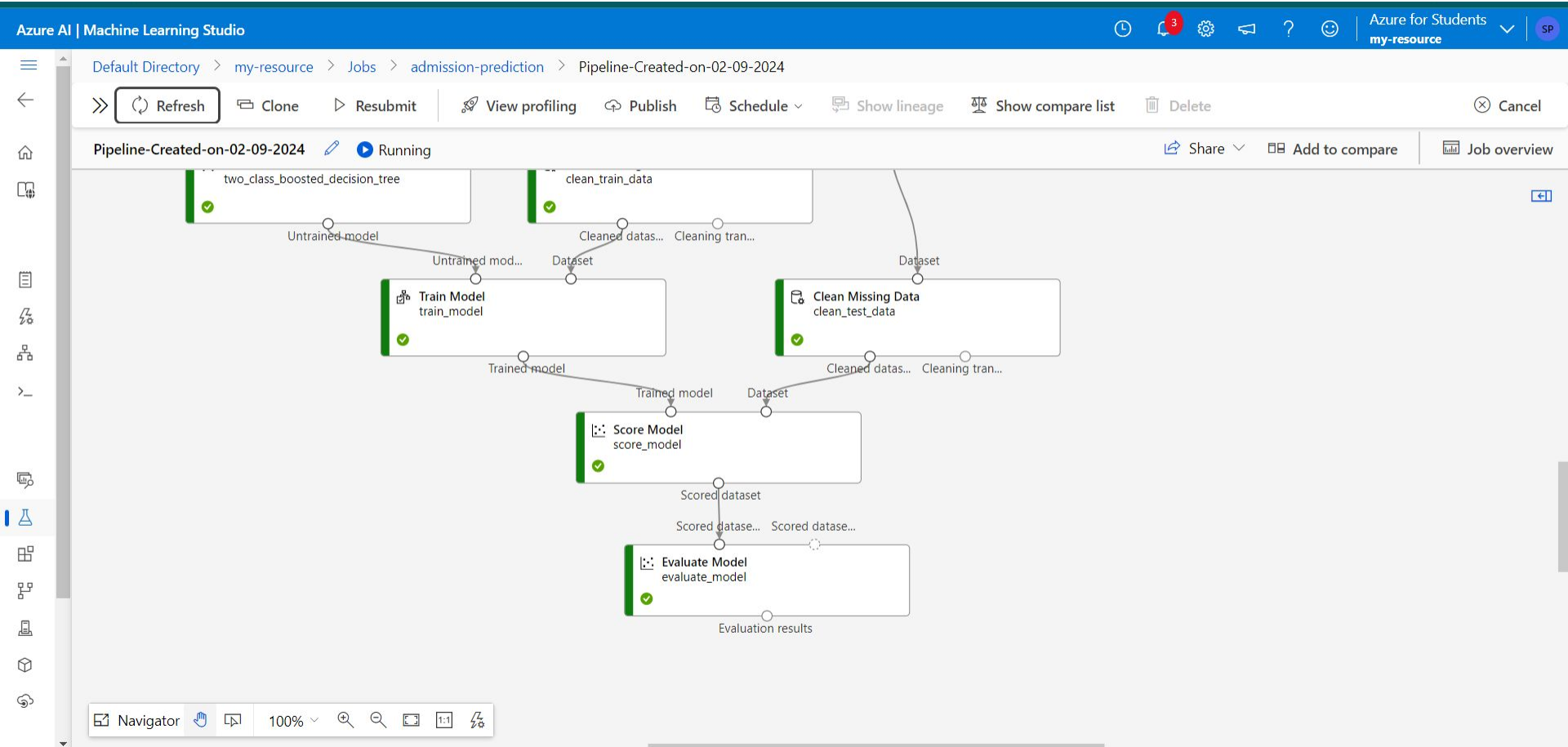


1. Please note, the score_model component in this model takes two inputs: the trained model and the test dataset. We had cleaned the test dataset separately from the train dataset.

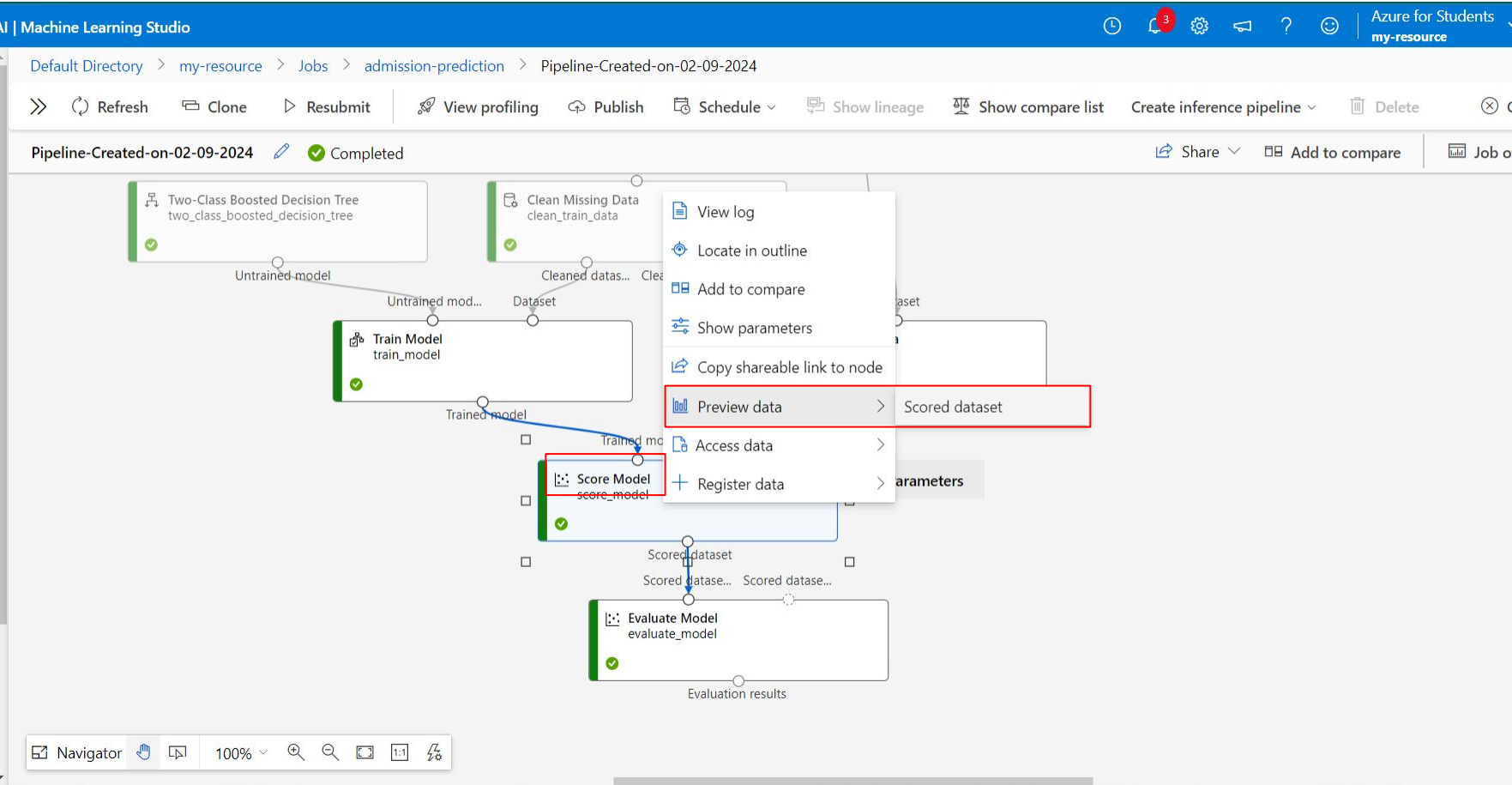
2. After scoring the model, we will evaluate it. This step completes the model development and execution process.

3. Please note, we will close this process here, as model deployment will be covered later, and is out of scope for the current case study.

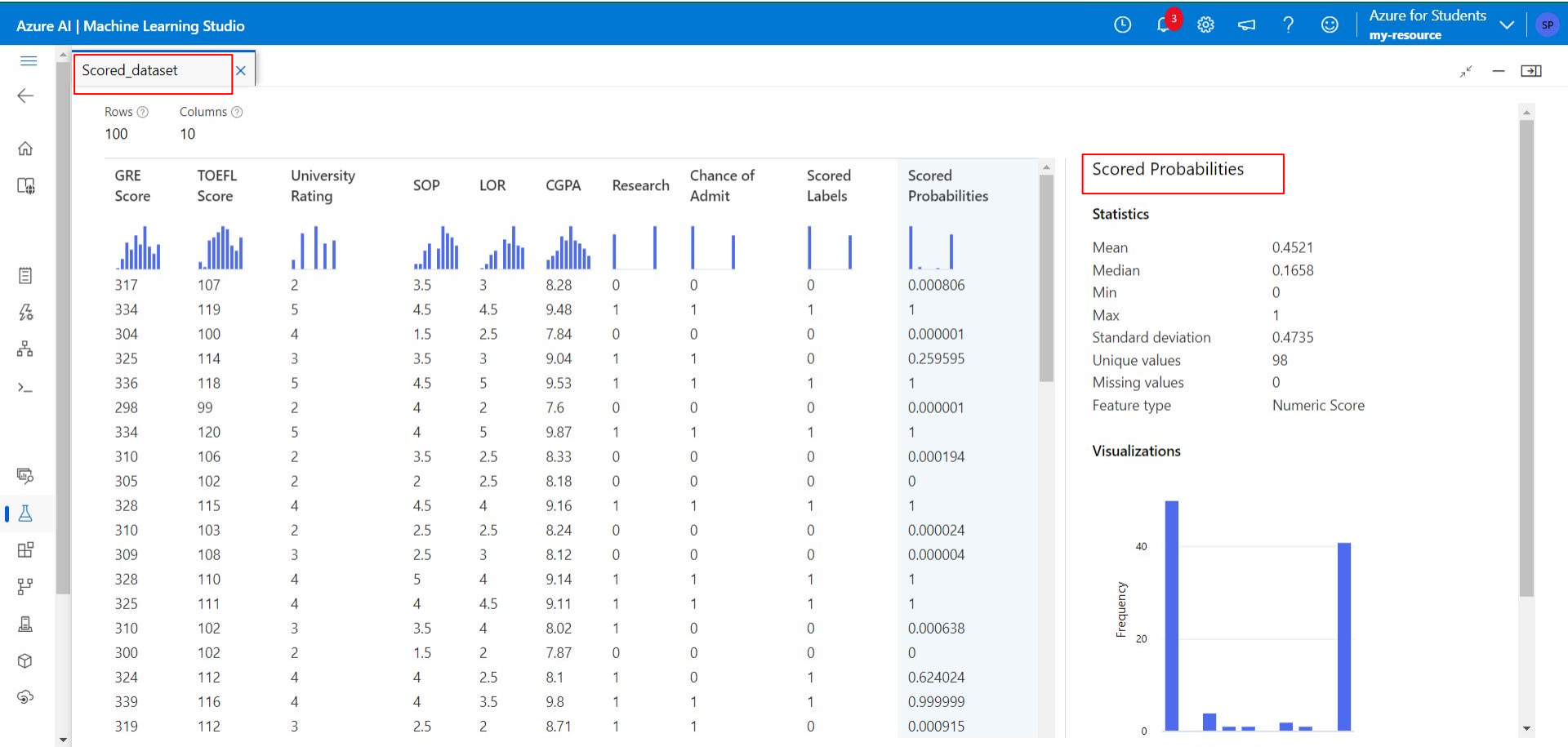
How to view progress of model run?



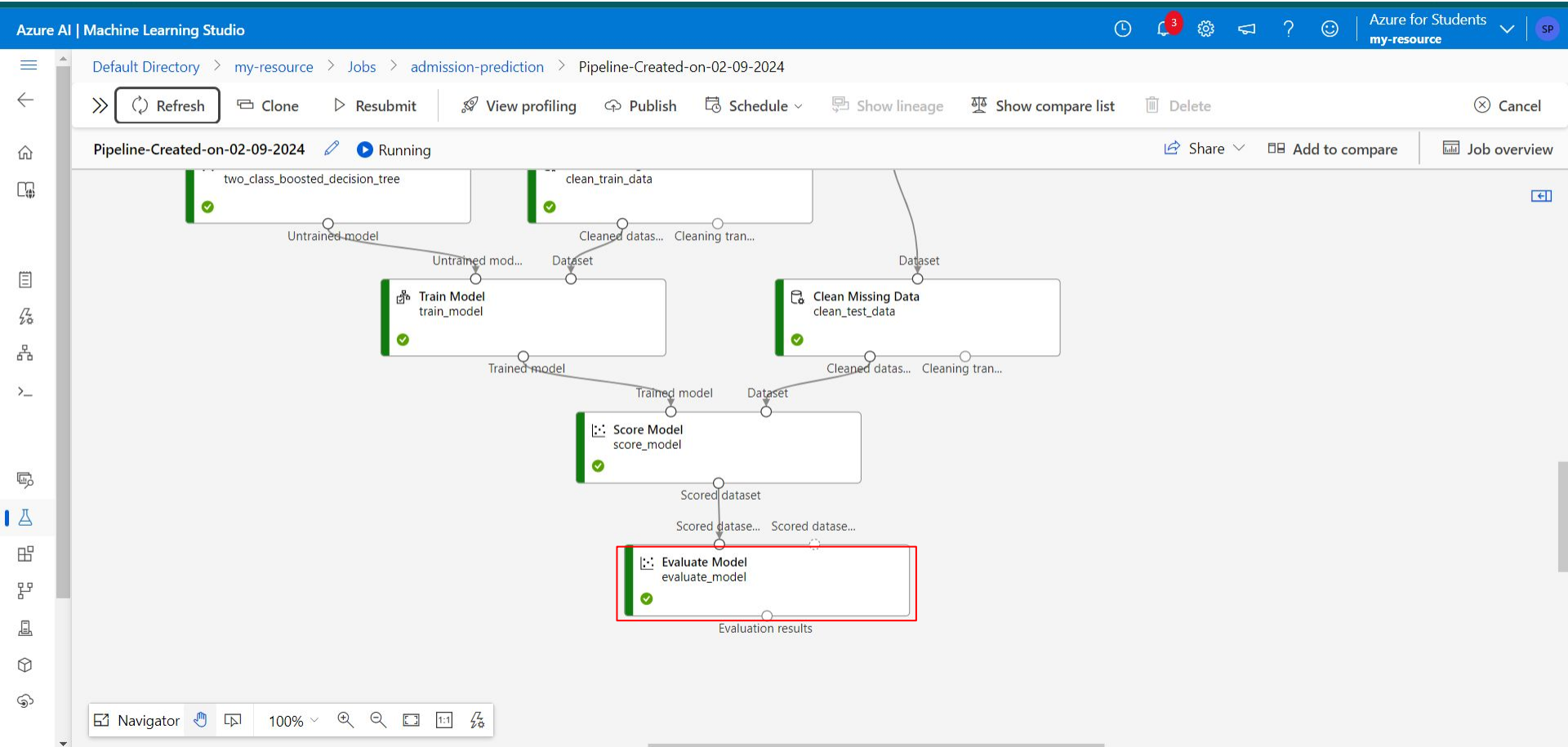
How to view the model predictions?



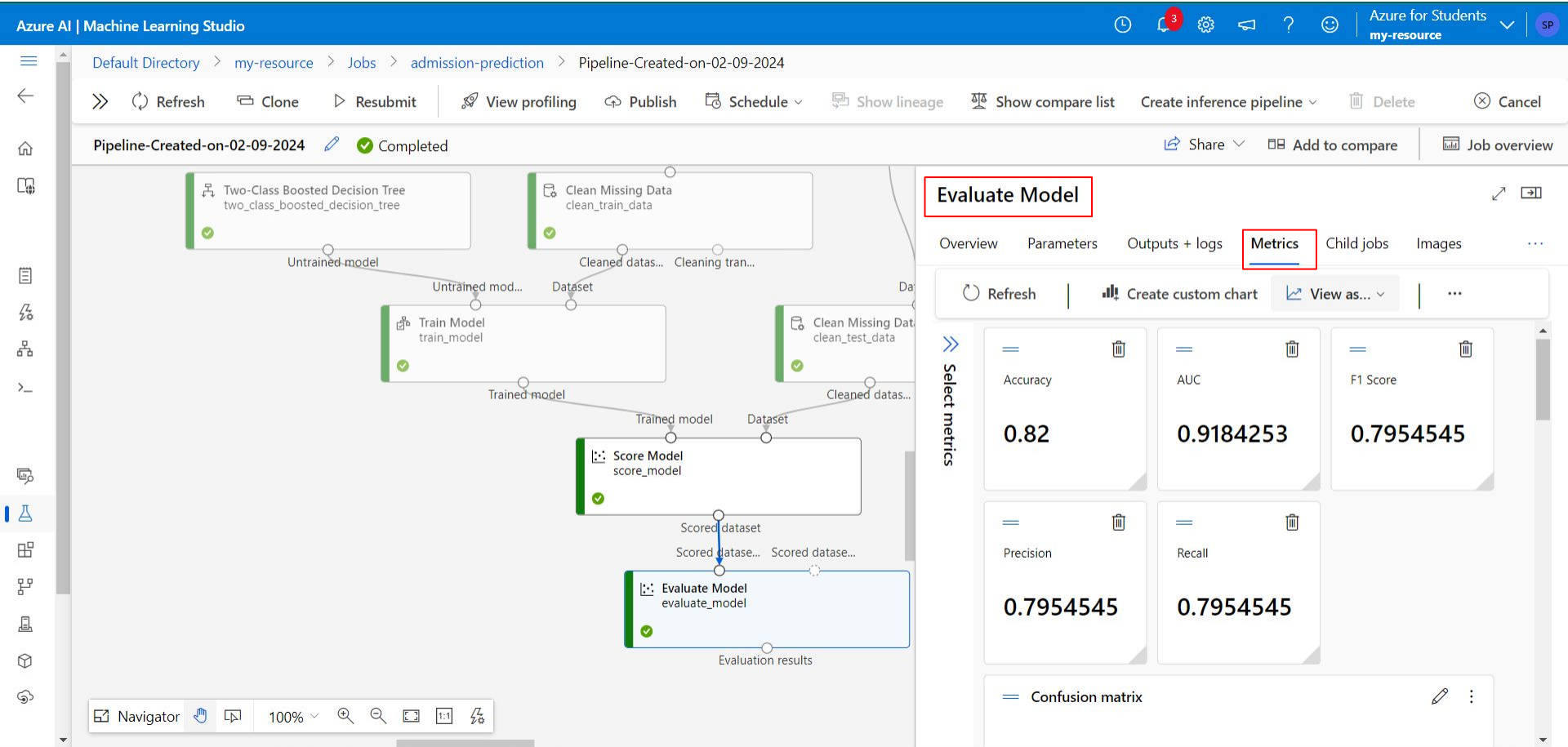
How to view the model predictions?



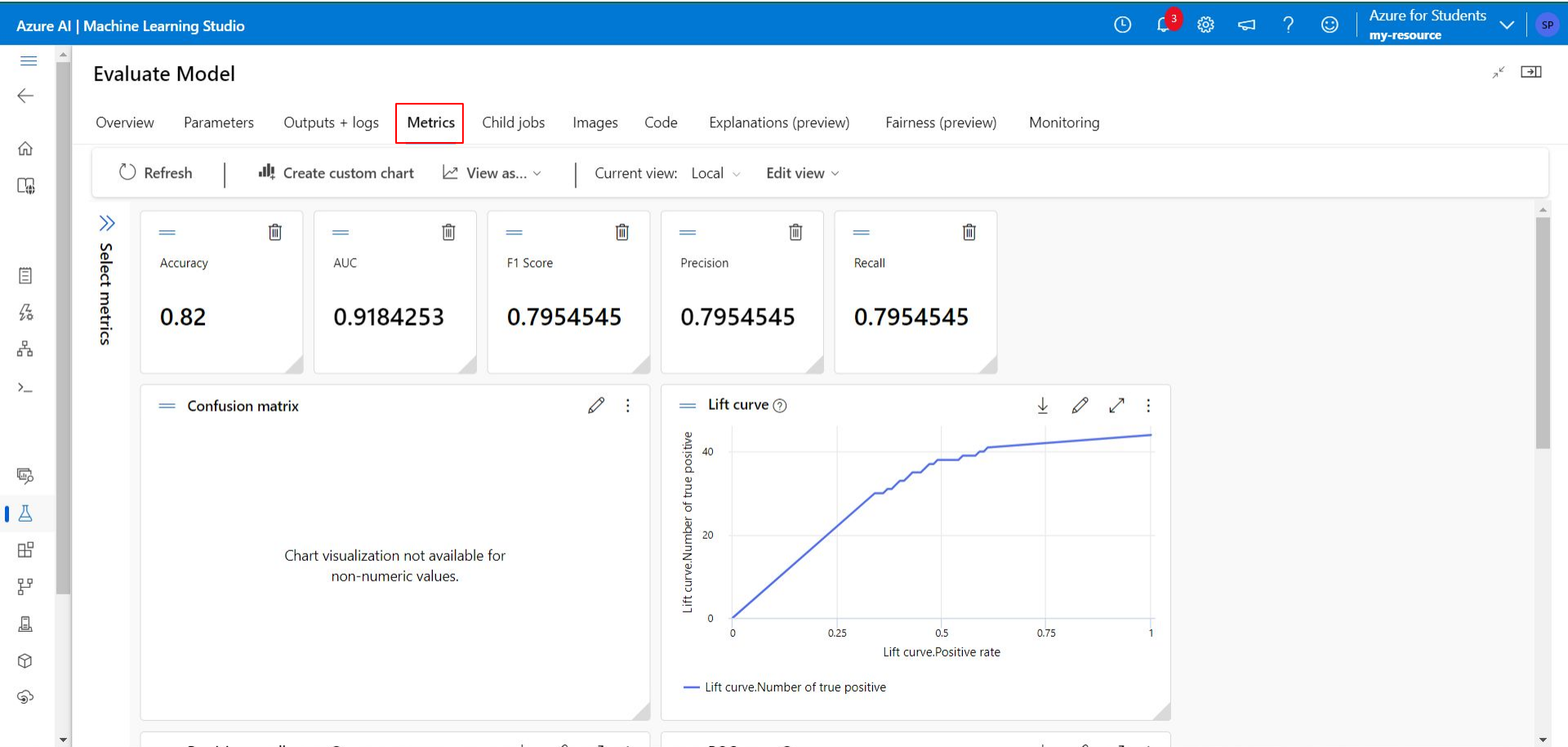
How to view the model metrics?



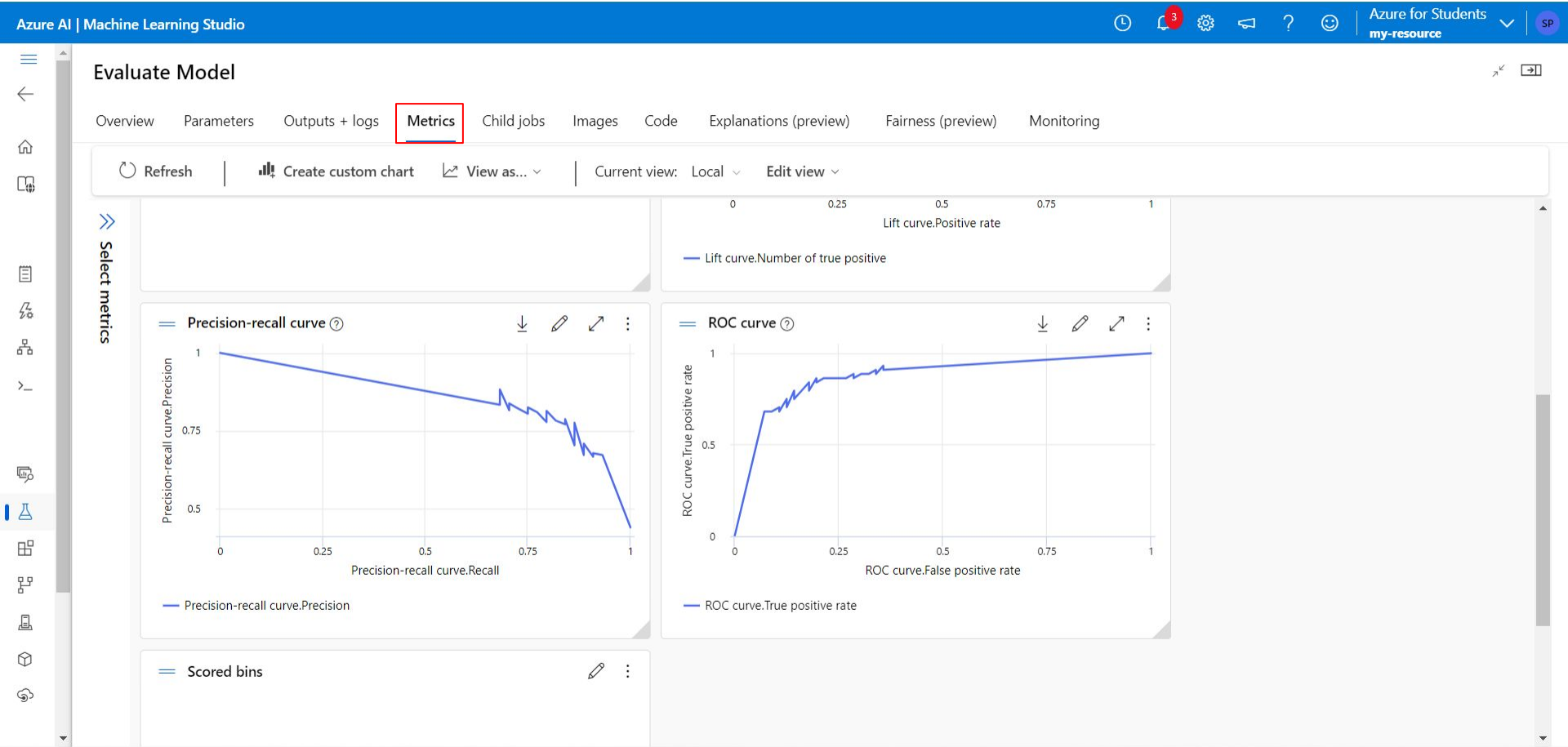
How to view the model metrics?



How to view the model metrics?



How to view the model metrics?



Closing the project - resource deletion