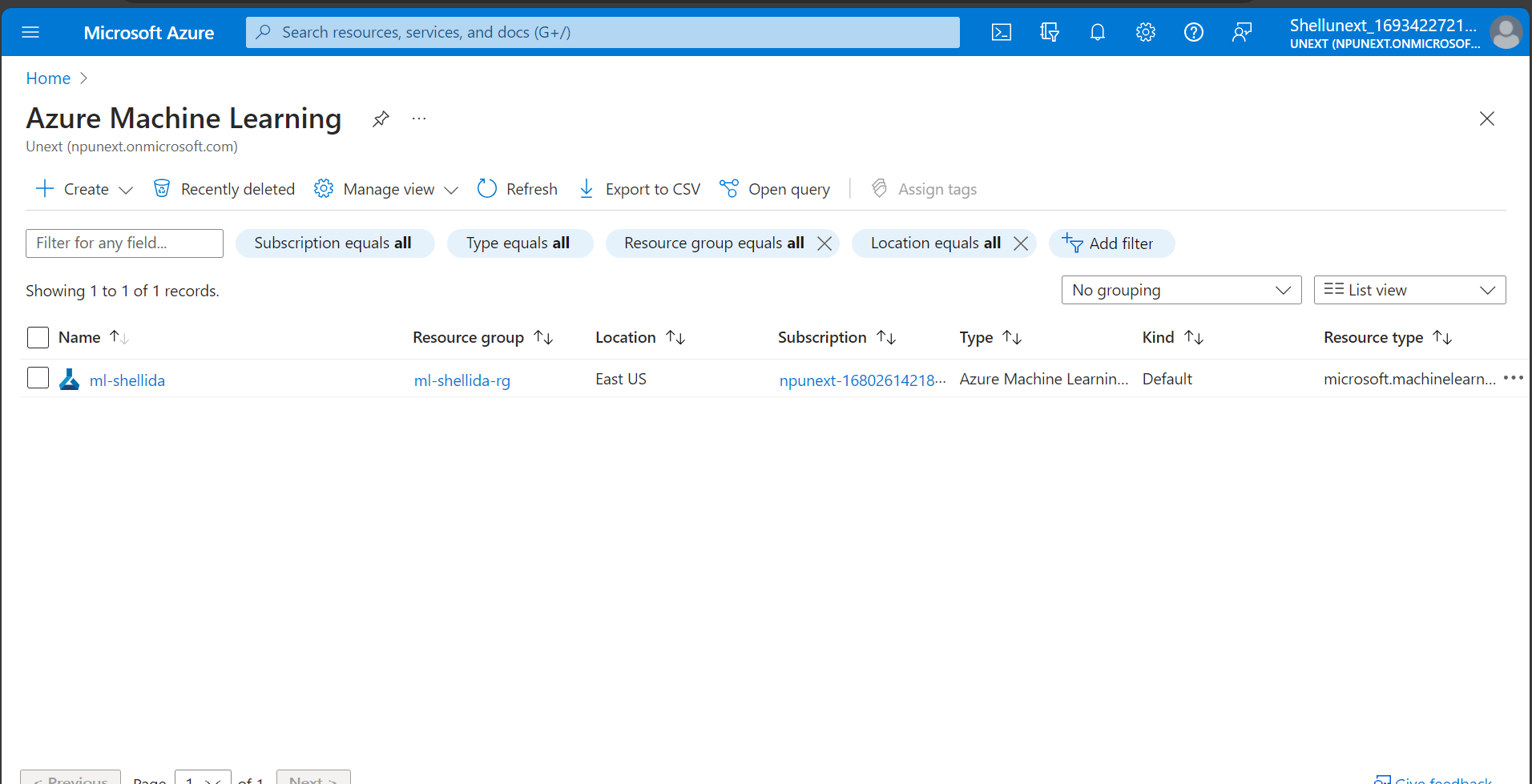
**Azure ML Studio - Hands-On Assessment**

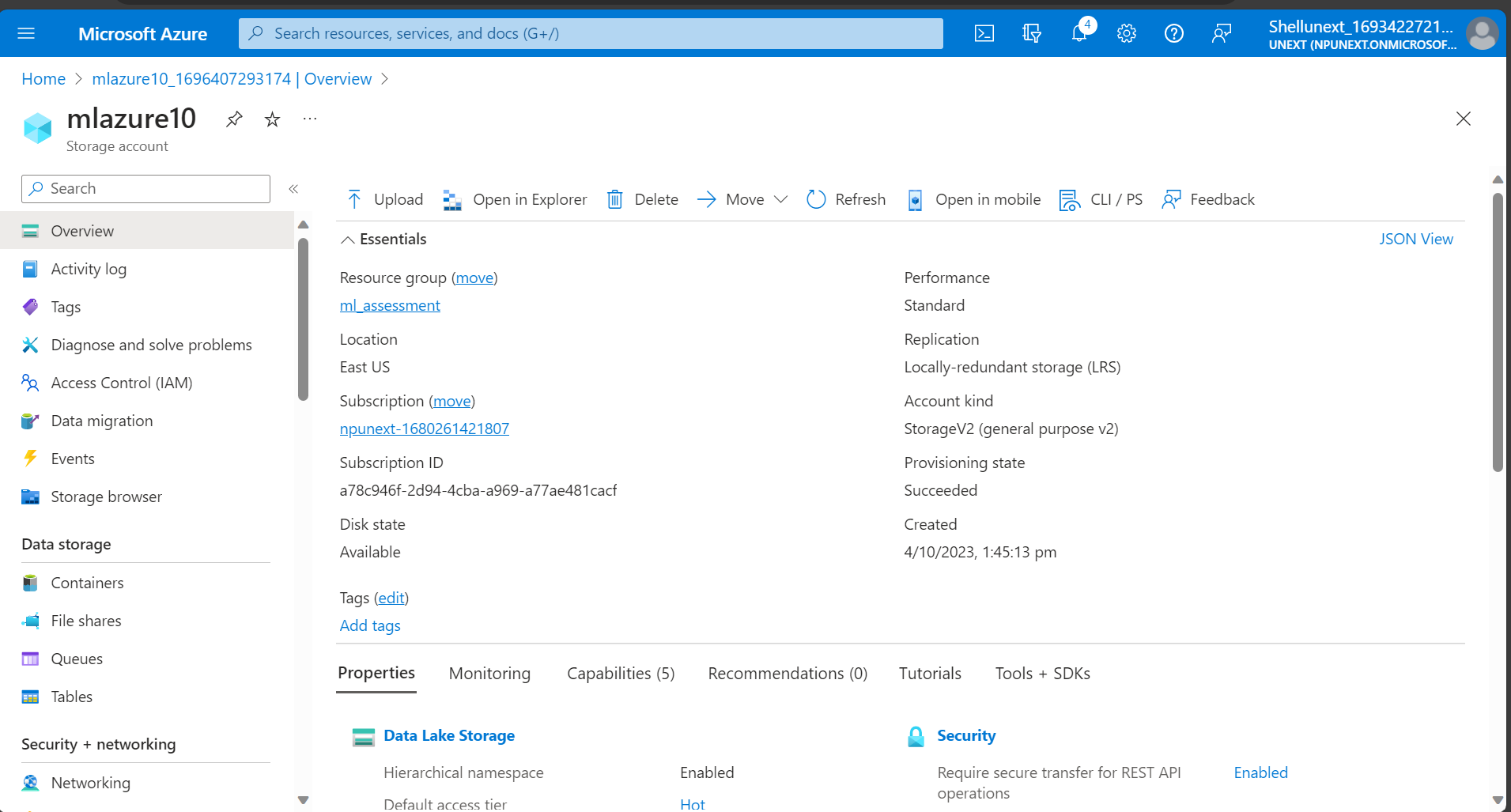
Akshita Verma

Shell IDA

Azure ML studio

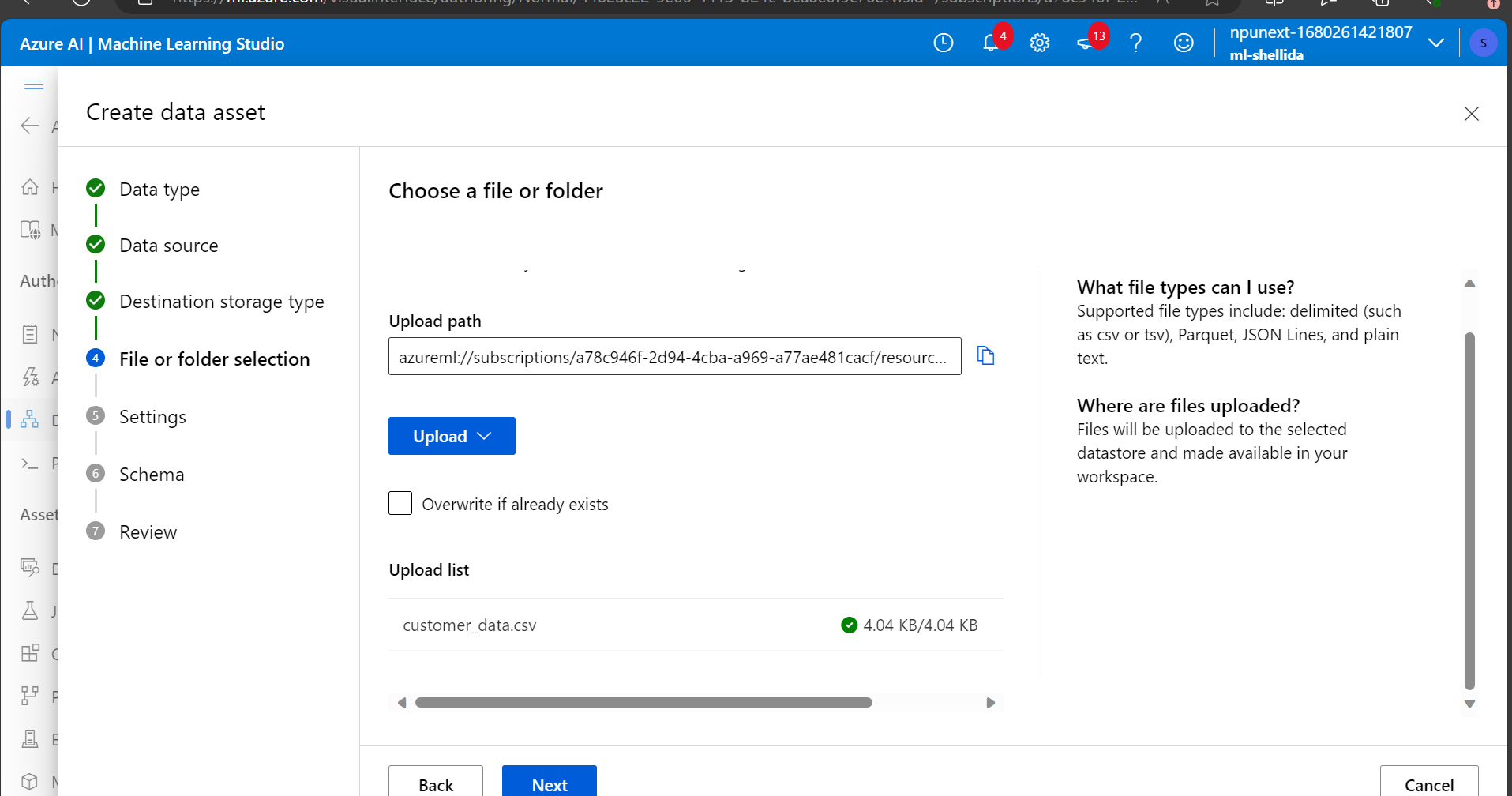


Azure storage account



A screenshot of a computer

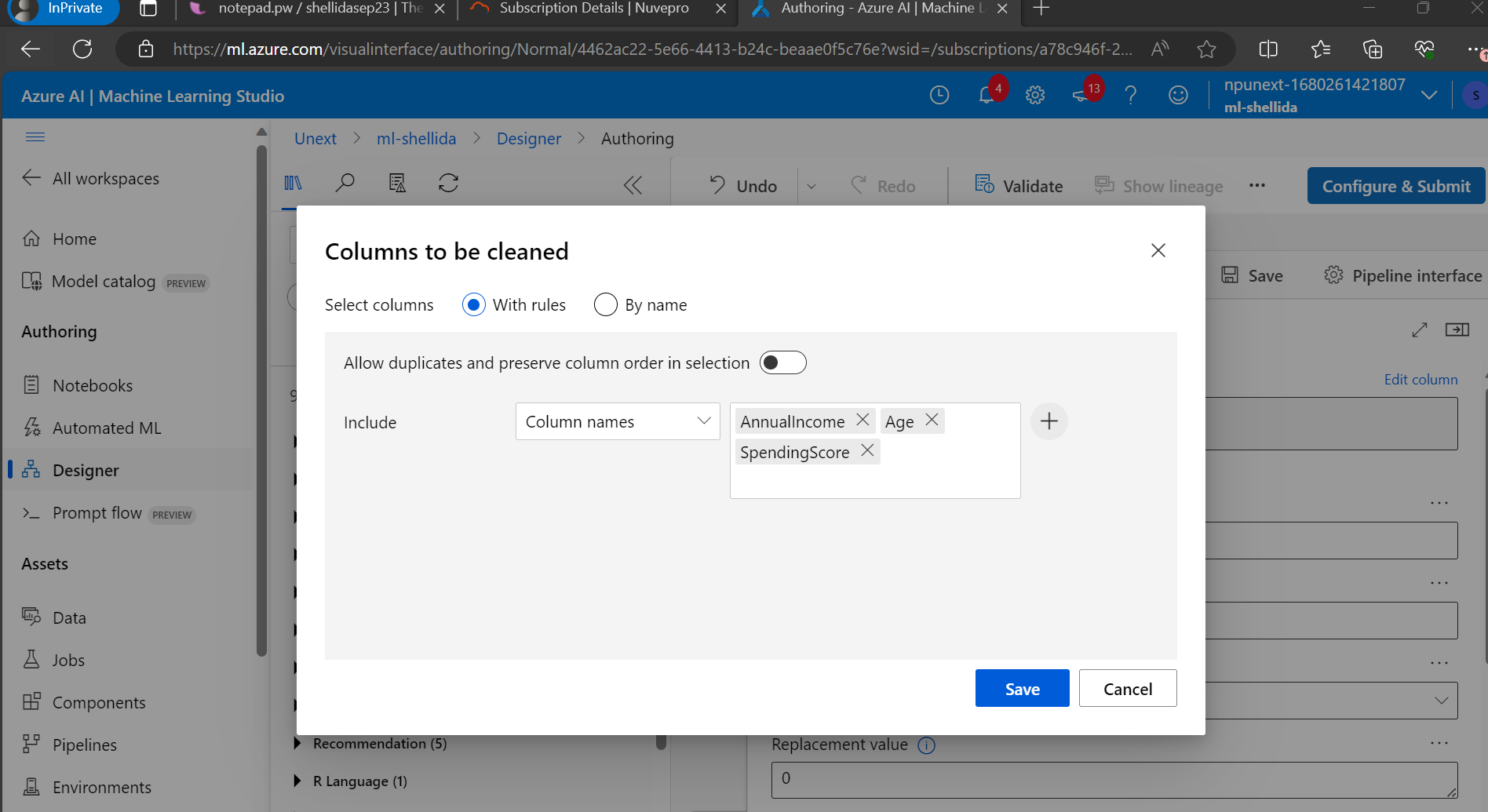
Description automatically generated



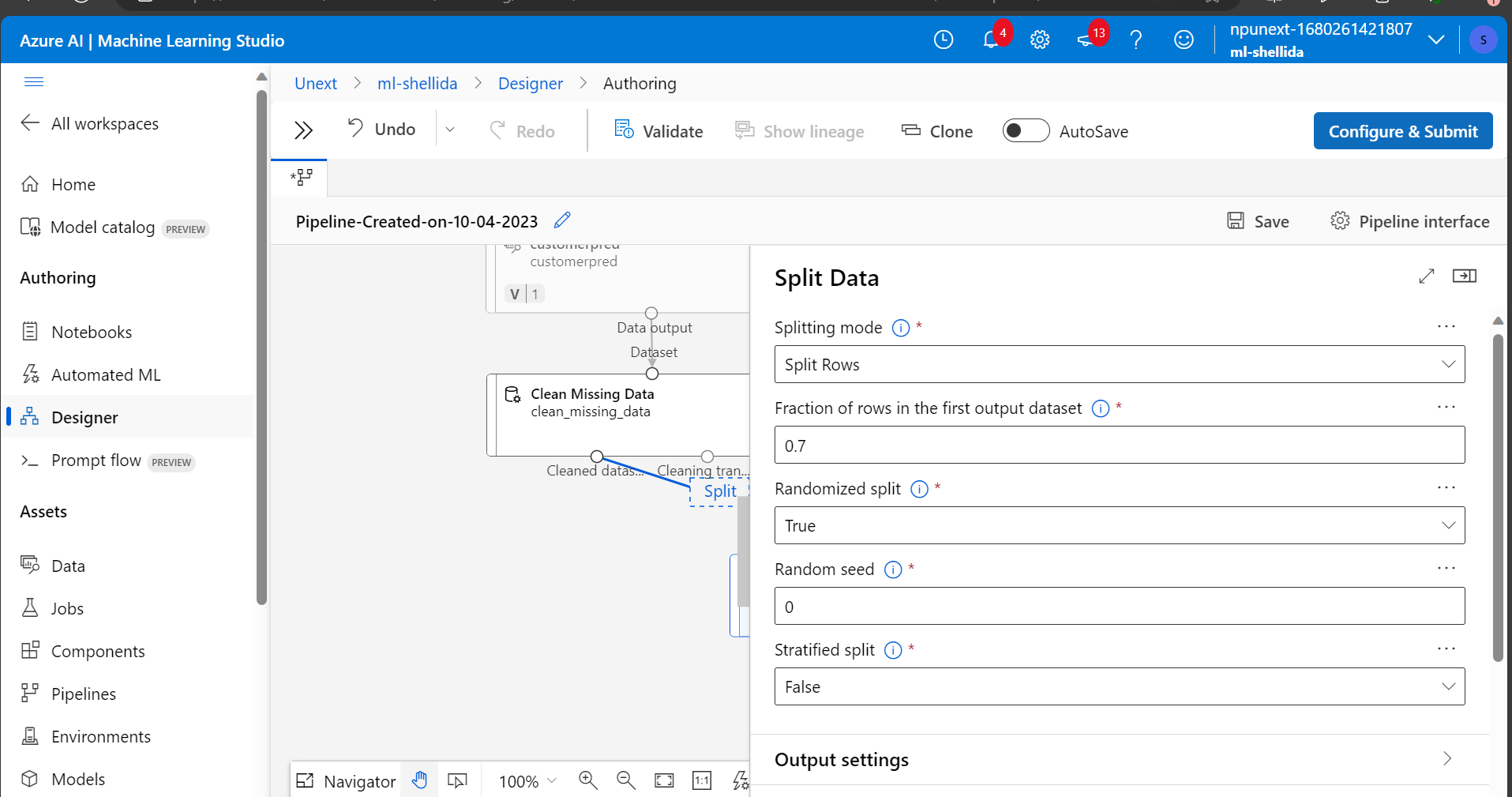
A screenshot of a computer

Description automatically generated

Clean data



Split data

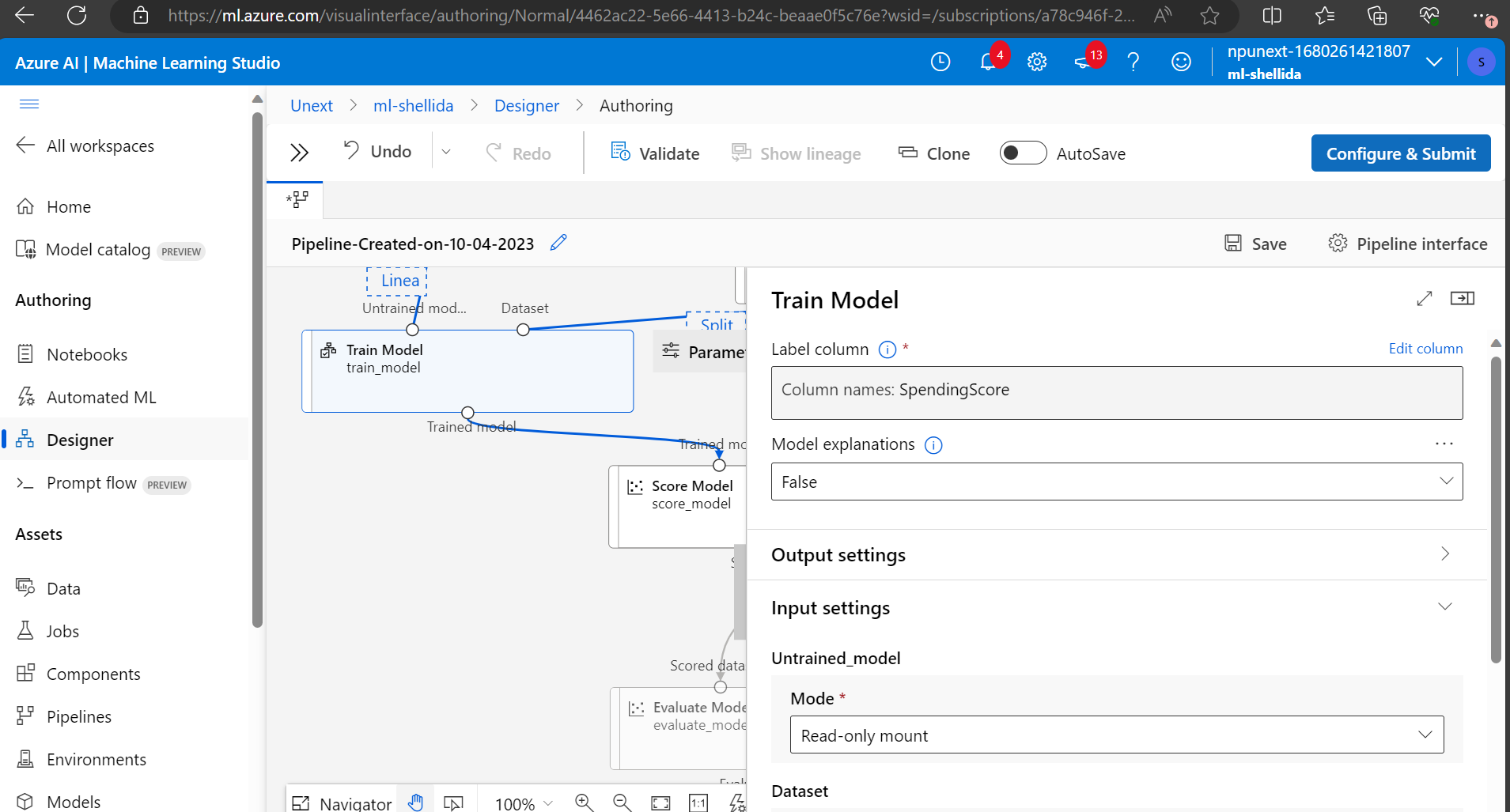


Linear regression model

A screenshot of a computer

Description automatically generated

Train model

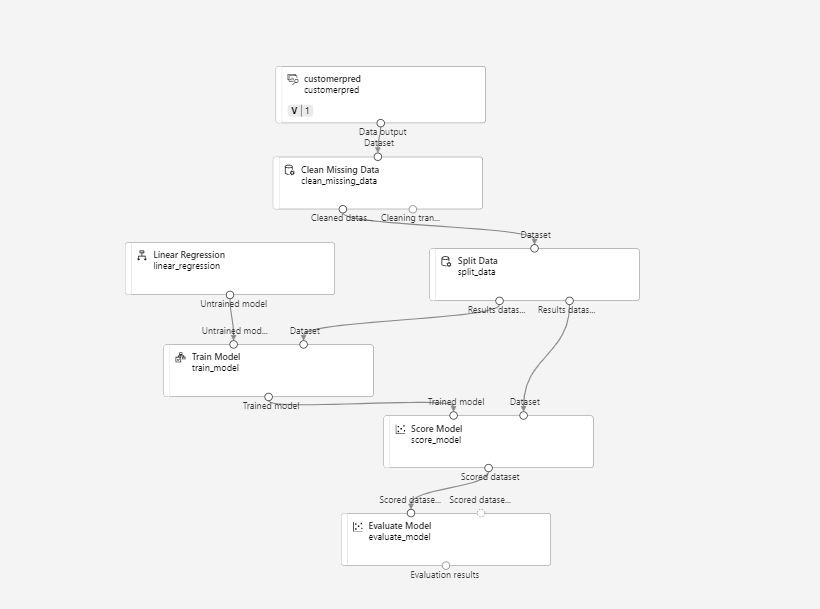


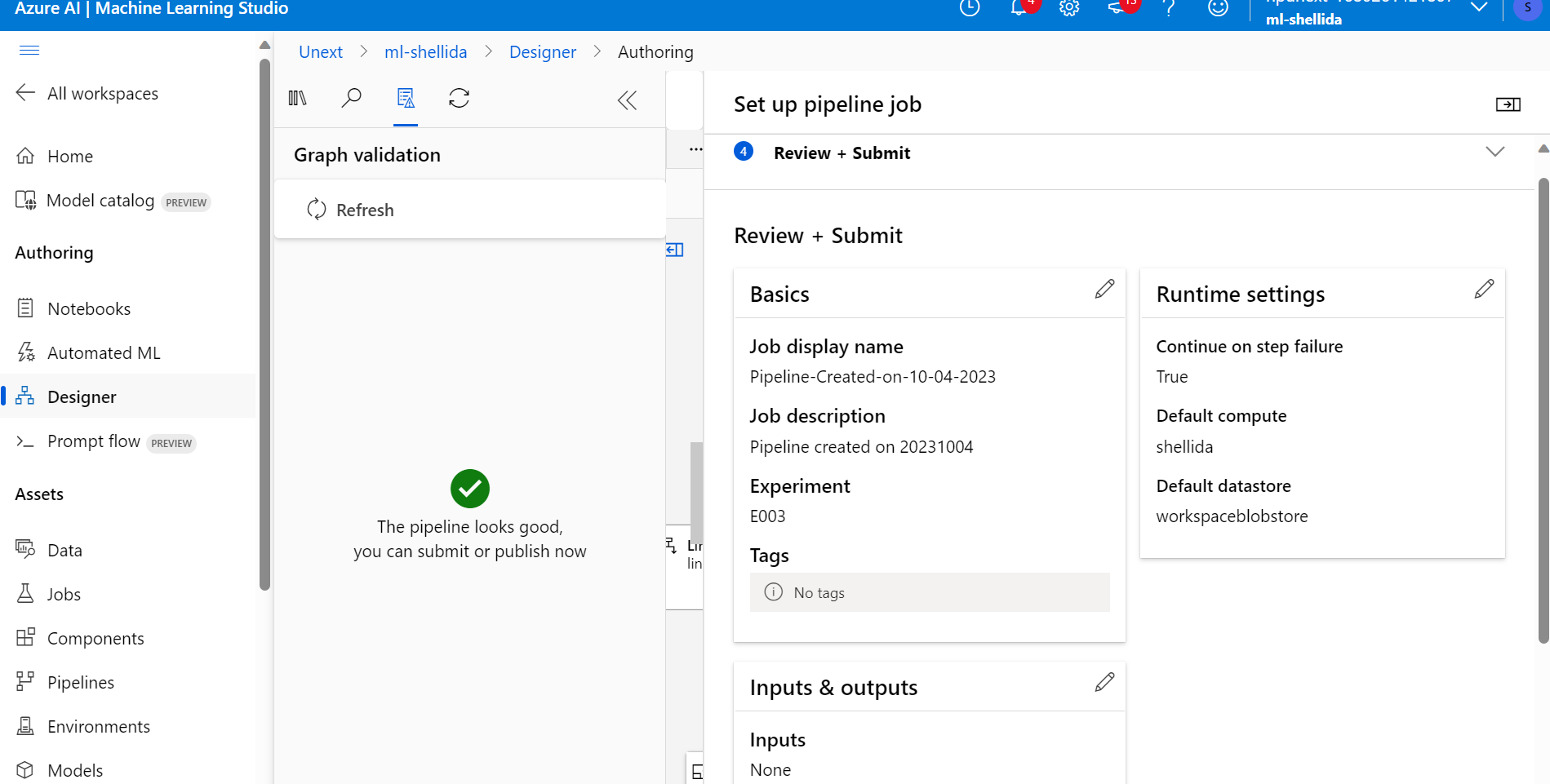
Score model

A screenshot of a computer

Description automatically generated

Pipeline





Result

A screenshot of a computer

Description automatically generated

1. What are the key steps involved in preparing the dataset for training a machine learning model using Azure Machine Learning? Briefly explain each step.

Ans: The following are the key steps involved in preparing dataset for a ML model:

1. Uploading dataset on Azure blob storage: Azure blob storage is created, and a container made inside it. There one can upload the dataset.
2. Uploading dataset as component in the pipeline.
3. Cleaning the dataset and filling missing values
4. Feature engineering involves selecting columns that are to be used for training.
5. Data splitting: data is split into training and testing sets

2. Why is it important to split the dataset into training and testing sets when developing a machine learning model? How does this help in model evaluation?

Ans: Dataset is split into training and testing sets. The model observes and learns patterns from the training dataset. Once the training is done, it needs to be evaluated on unseen data so that we can know if its overfit or underfit. This unseen data is your test dataset. How the model scores on test data is the real metric as to how it will perform in the real world.

3. Describe a machine learning algorithm suitable for predicting customer purchasing behaviour in the given scenario. Explain why you chose this algorithm.

Ans: I chose the Linear Regression Algorithm to predict the customer purchasing power. In a linear regression model, the dependent variable is predicted using one or more independent variables. Here the dependent variable, spending score is predicted based on the age and annual income of a customer. We can get to know what age category (young, middle ,old) and section of people(poor ,middle, rich) spend the most and the least.

4. What is hyperparameter tuning, and why is it important in machine learning? Explain a technique used for hyperparameter tuning and its benefits.

Ans: Hyperparameter tuning is altering the parameters such as no. of epochs, batch size, learning rate etc to optimize the performance of the model. It is very important in machine learning as it helps in optimizing the model’s performance, avoiding underfitting and overfitting.

Easy to Implement: Grid search is straightforward to implement and is suitable for small to moderately sized hyperparameter search spaces.By using grid search, you can systematically find the best hyperparameters for your model and achieve better overall performance.