



## Azure ML Classic Studio

### Predicting of blood data using Regression Model in AzureML Classic Studio.

This model (Pipeline) trains a linear regressor to predict blood data based on technical features such as collection, availability, Because you're trying to answer the question "How much?" this is called a regression problem. However, you can apply the same fundamental steps in this example to tackle any type of machine learning problem whether it be regression, classification, clustering, and so on.

#### Gallery Link:

Blood data amount prediction model[19K41A0590]

Blood data amount prediction using pre-available dataset and training the model using Linear Regression. Tags: Linear Regression, Automobile, Azure ML,



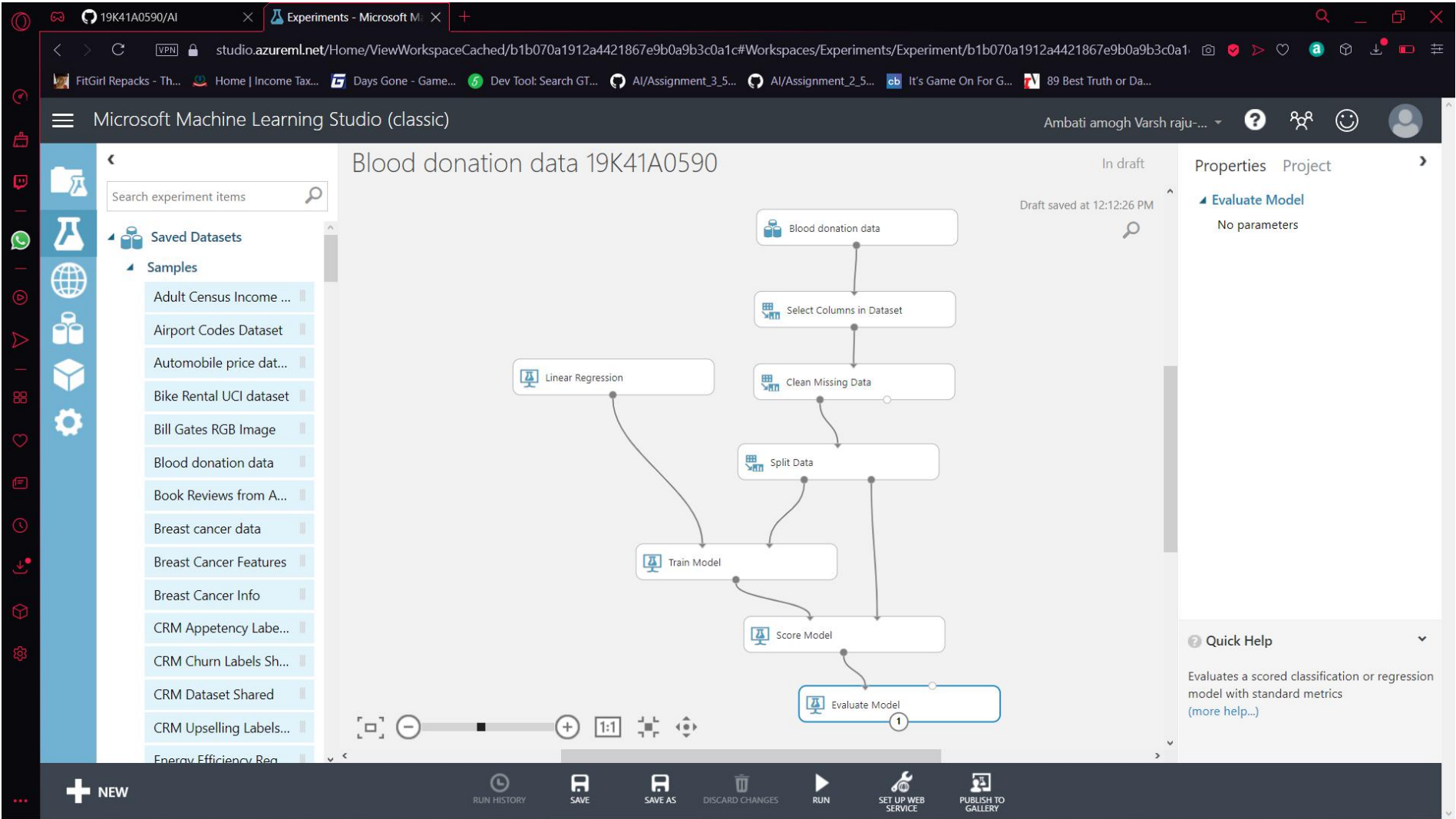
<https://studio.azureml.net/Home/ViewWorkspaceCached/b1b070a1912a4421867e9b0a9b3c0a1c#Workspaces/Experiments/Experiment/b1b070a1912a4421867e9b0a9b3c0a1c.f-id.fdfcce7270aa4ec9943feb132bd4c87d/ViewExperiment>

navigate to the link to see the Workflow and you can download the project as well.

### Machine Learning Project Workflow

1. Import Data
2. Explore Data (Missing values, outliers)
3. Preprocess data (Missing value imputation, outlier treatment, normalization)
4. Model Selection
5. Model Training
6. Model Testing
7. Model Deployment

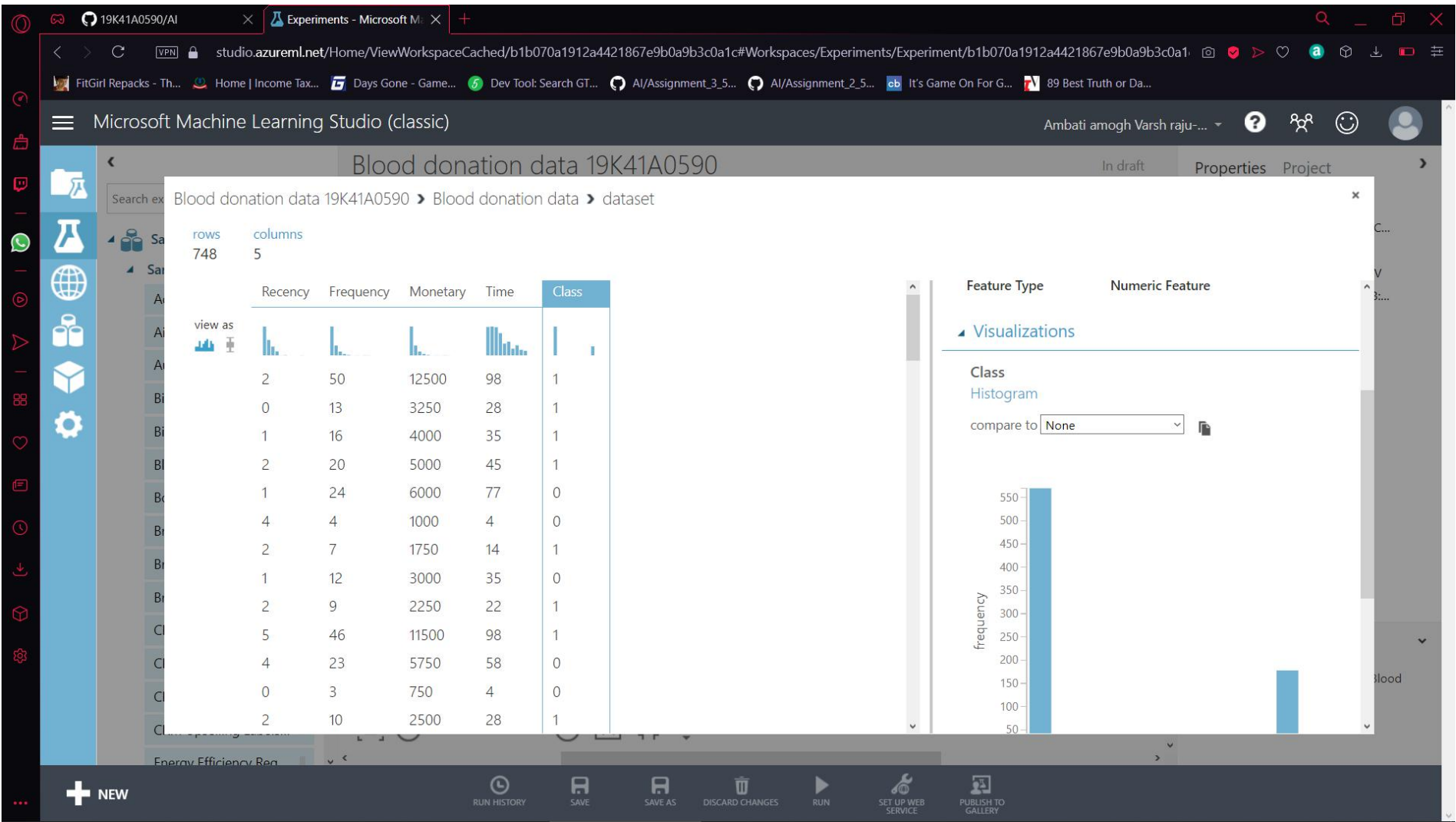
Workflow



Project Workflow

Import Data:

- importing the RAW dataset which is in CSV format.
- the dataset is pre-available in the Azure ML Classic Studio.



Automobile Price RAW dataset (CSV format)

## Explore Data

- this basically includes data visualization to search for any missing values in the Dataset.
- if any missing values are found, then they needs to be cleaned.
- selecting the required columns and clean the data using the Clean Missing Value module (Just Drag n' Drop )

Microsoft Machine Learning Studio (classic) interface showing a workflow for 'Blood donation data 19K41A0590'. The workflow includes modules: Blood donation data, Select Columns in Dataset, Clean Missing Data, Split Data, Train Model, Score Model, and Evaluate Model. The 'Select Columns in Dataset' module is highlighted, showing selected columns: Recency, Frequency, Monetary, Time, Class. The 'Clean Missing Data' module is also visible in the workflow.

Microsoft Machine Learning Studio (classic) interface showing the 'Clean Missing Data' module configuration. The 'Selected columns' are set to 'All columns'. The 'Minimum missing value ratio' is 0, and the 'Maximum missing value ratio' is 1. The 'Cleaning mode' is set to 'Custom substitution value', and the 'Replacement value' is 0. The 'Generate missing value indicator column' checkbox is unchecked.

## Data Cleaning



Split Data

Microsoft Machine Learning Studio (classic) interface showing a workflow for data splitting. The workflow includes: Blood donation data, Select Columns in Dataset, Clean Missing Data, Split Data (highlighted), Train Model, Score Model, and Evaluate Model. The Split Data module is selected, and its properties are shown on the right:

- Splitting mode: Split Rows
- Fraction of rows in the first output dataset: 0.5
- Randomized split: ☒
- Random seed: 0
- Stratified split: False
- START TIME: 9/21/2021 12:21:47 PM
- END TIME: 9/21/2021 12:21:47 PM
- ELAPSED TIME: 0:00:00.000
- STATUS CODE: Finished
- STATUS DETAILS: Task output was present in output cache

Quick Help: Split the rows of a dataset into two distinct sets ([more help...](#))

Data Splitting

Model Training and Algorithm

Microsoft Machine Learning Studio (classic) interface showing a workflow for model training. The workflow includes: Blood donation data, Select Columns in Dataset, Clean Missing Data, Split Data, Train Model (highlighted), Score Model, and Evaluate Model. The Train Model module is selected, and its properties are shown on the right:

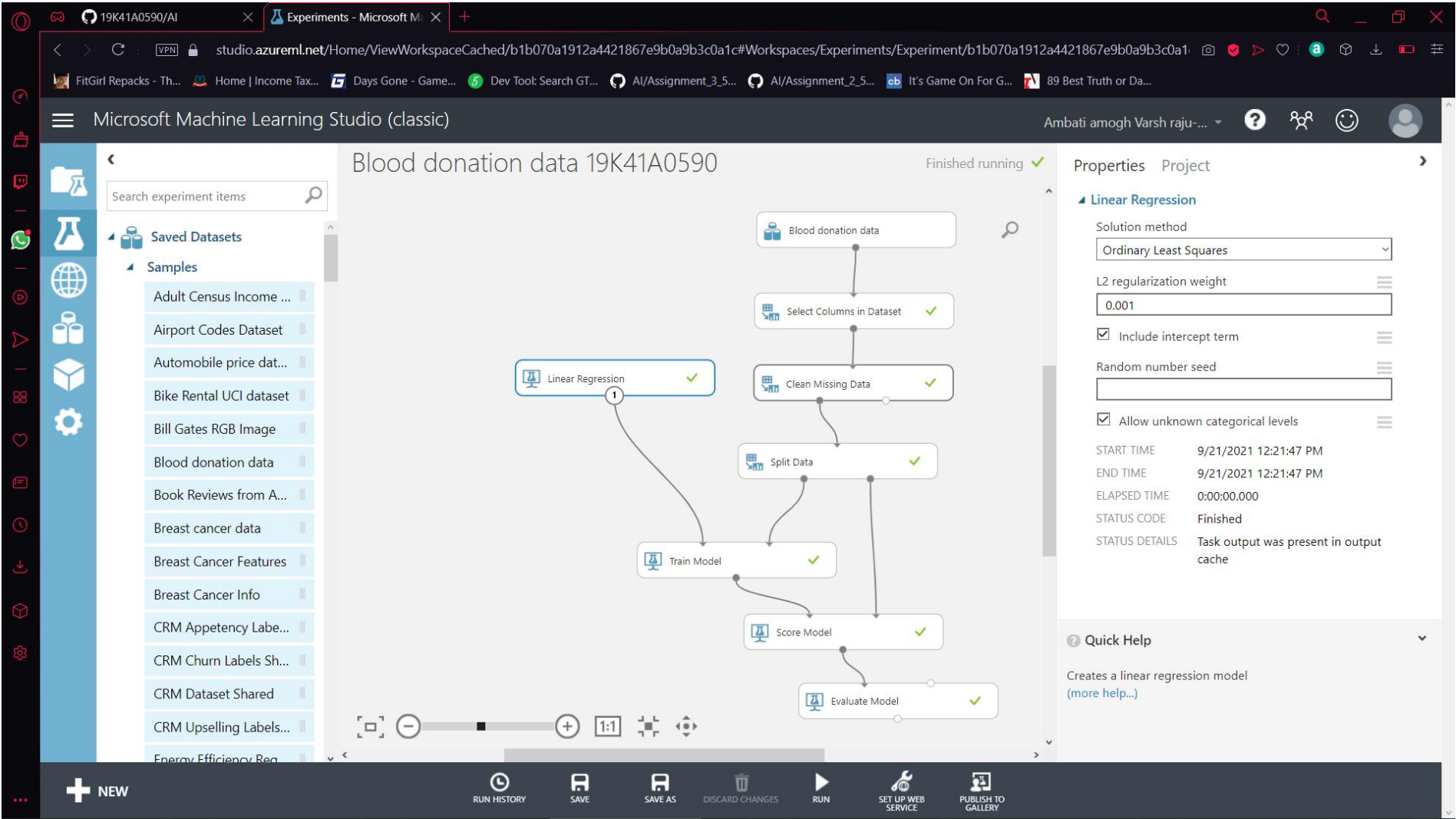
- Label column: Selected columns: Class
- Column names: Class
- Launch column selector
- START TIME: 9/21/2021 12:21:47 PM
- END TIME: 9/21/2021 12:21:47 PM
- ELAPSED TIME: 0:00:00.000
- STATUS CODE: Finished
- STATUS DETAILS: Task output was present in output cache

Quick Help: Train a previously created classification or regression model ([more help...](#))

Model Training

using Linear regression to train the model

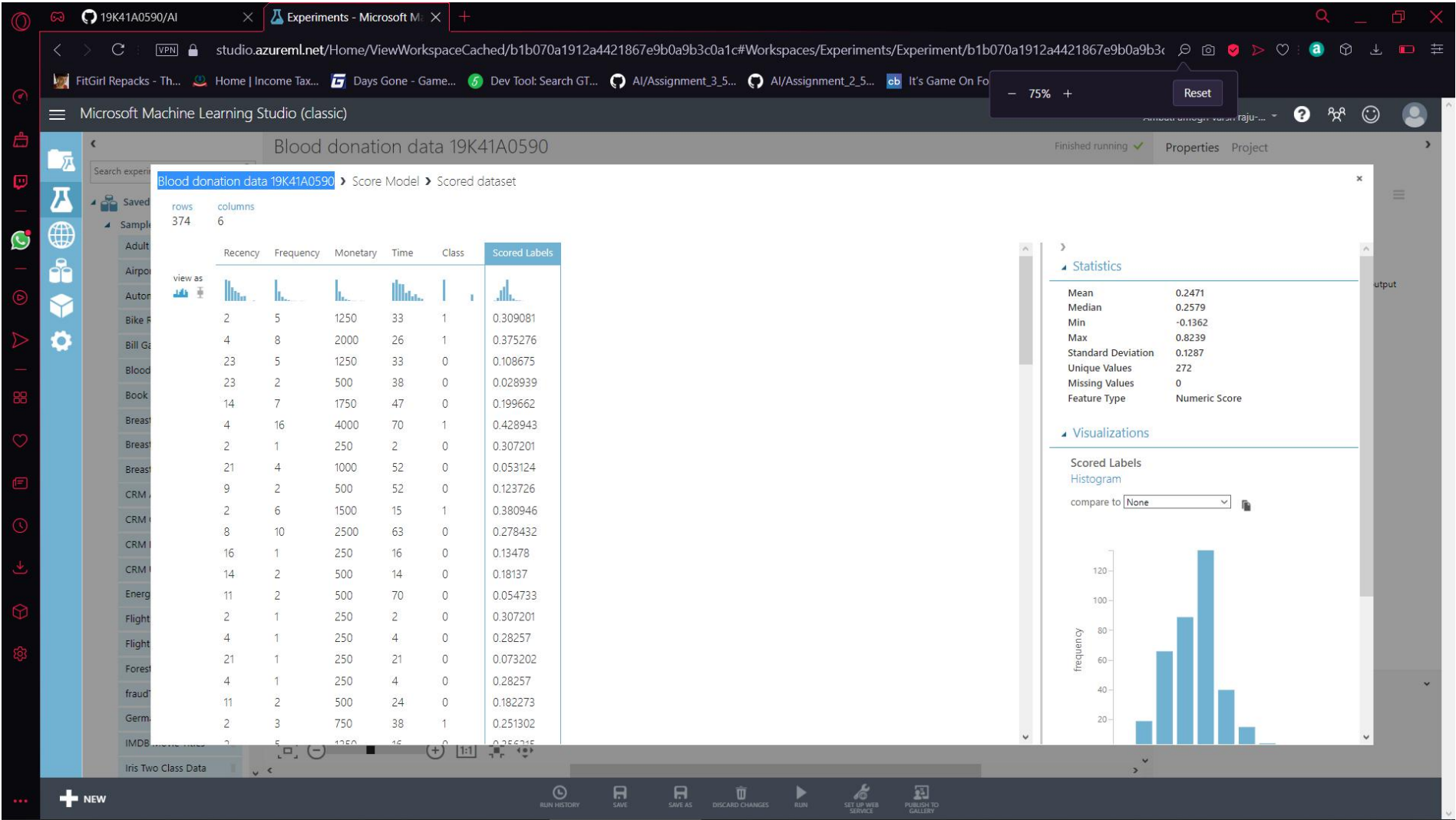
- Since the goal of this sample is to predict automobile prices, and because the label column (price) is continuous data, a regression model can be a good choice. We use Linear Regression for this pipeline.



Linear Regression

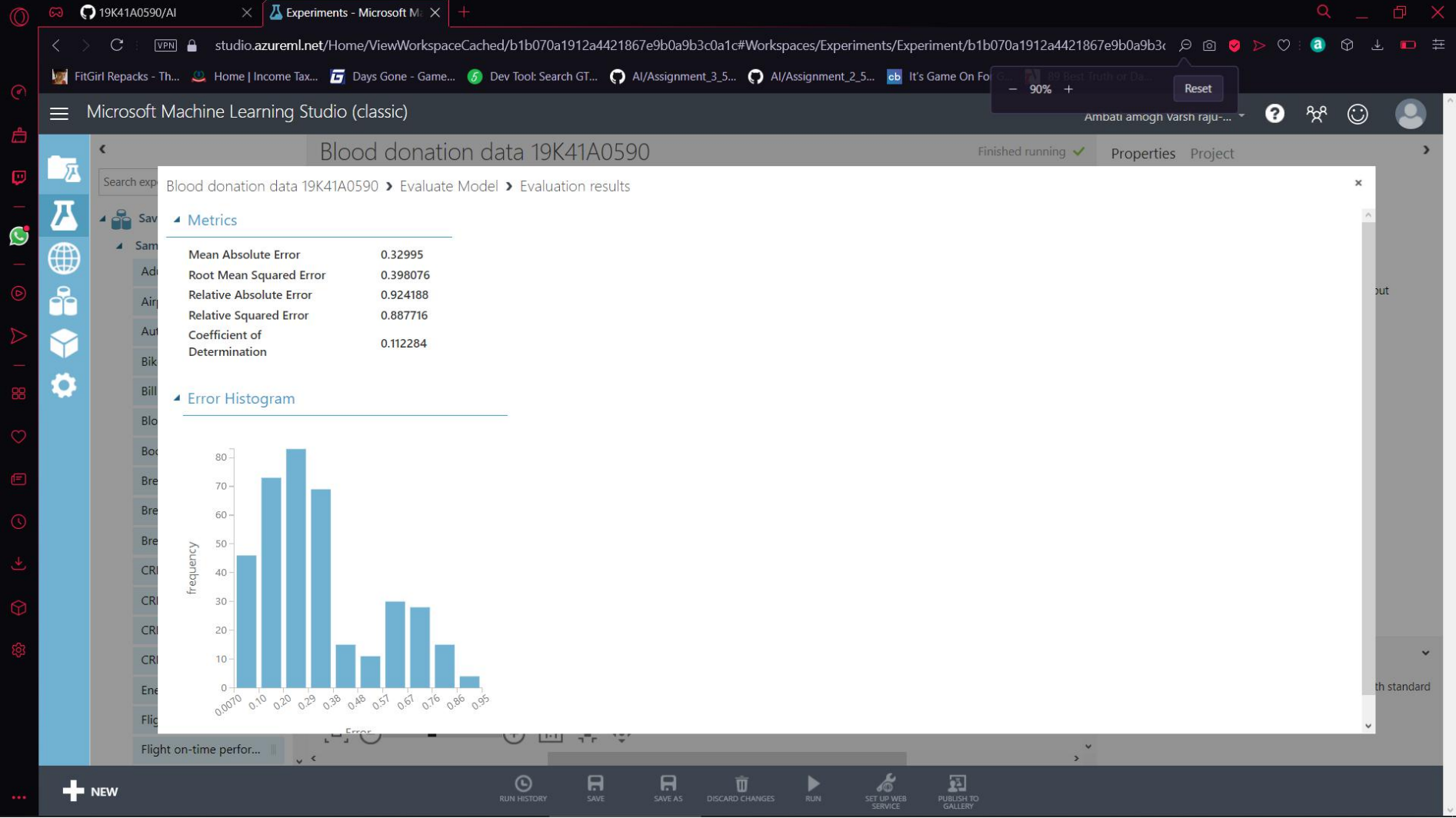
Score Model and Evaluate Model

- After the model is trained, we can use the Score Model and Evaluate Model modules to generate predicted results and evaluate the models.



Score Labels

# Evaluation Results



Model Evaluation Results