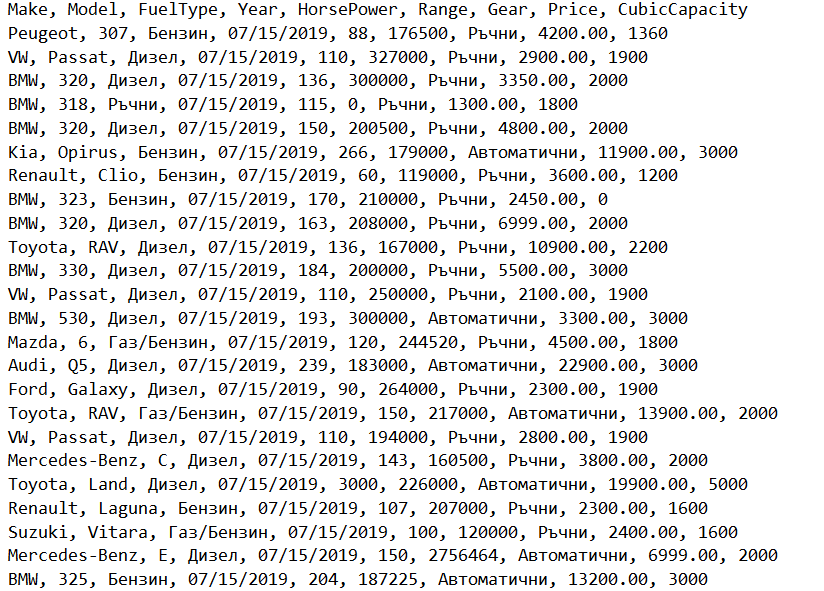
# Exercises: ML.NET

Problems for exercises and homework for the ["ASP.NET Core" course @ SoftUni"](https://softuni.bg/trainings/2244/csharp-oop-february-2019).

## Gather Data

You already have the data as resource in csv format (carsbg.csv), you can check it/explore it.



## Clean and Prepare Data

You can check the dataset and if you see data which is "dirty" you can clean or update it.

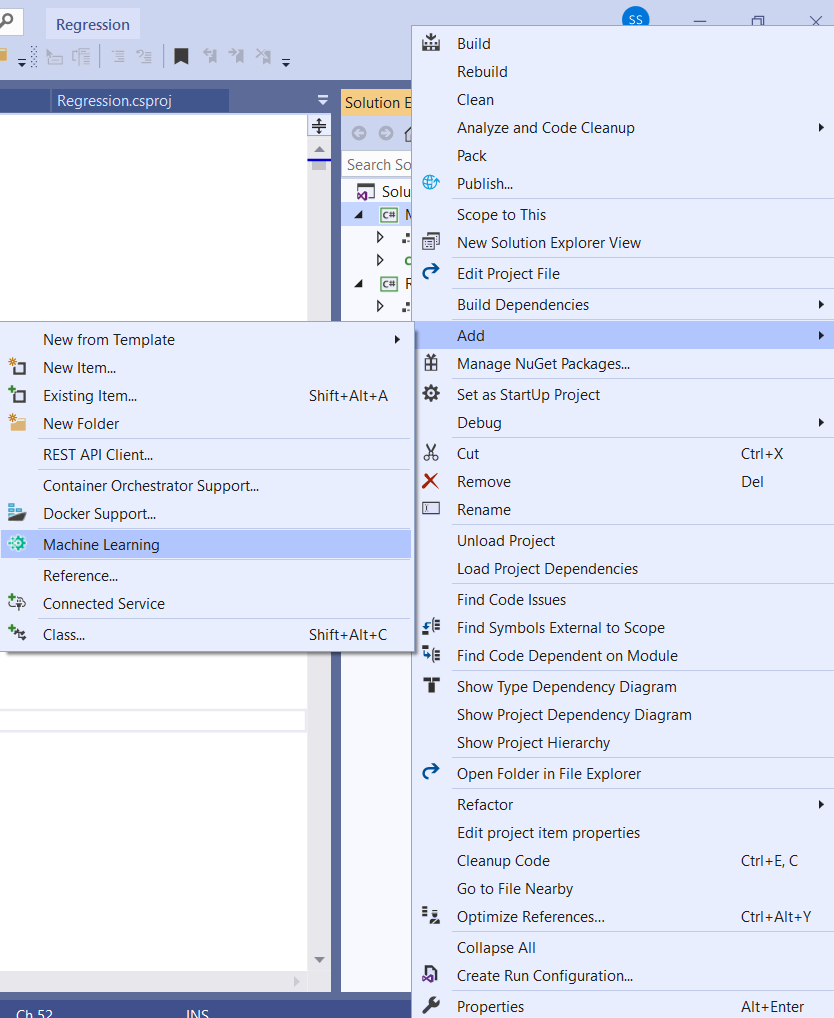
## Install ML.NET Model Builder

Link - <https://marketplace.visualstudio.com/items?itemName=MLNET.07>

NOTE: Restart Visual Studio

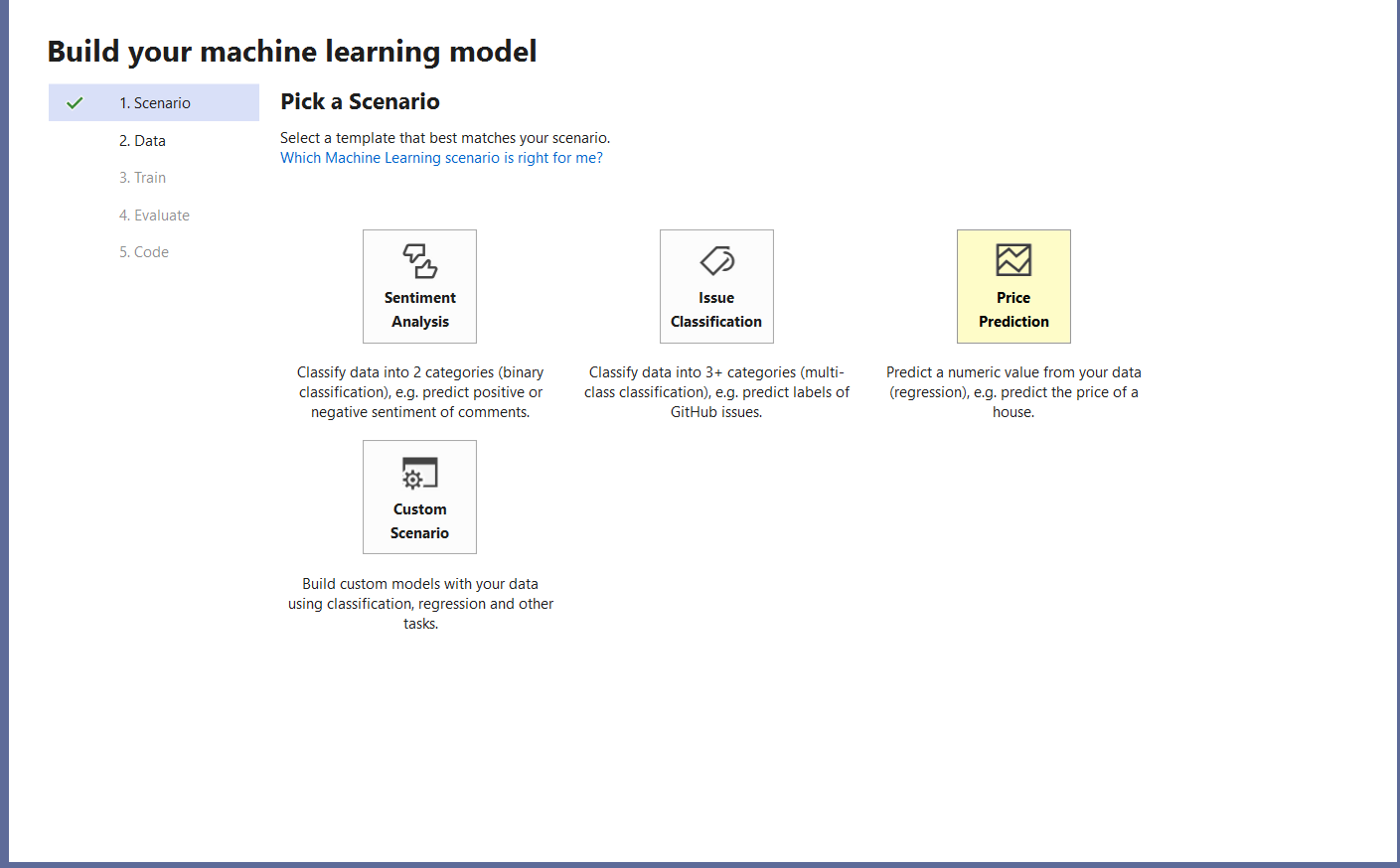
## Set up project

### Create a Project and run the "Model Builder"



### Choose a Scenario

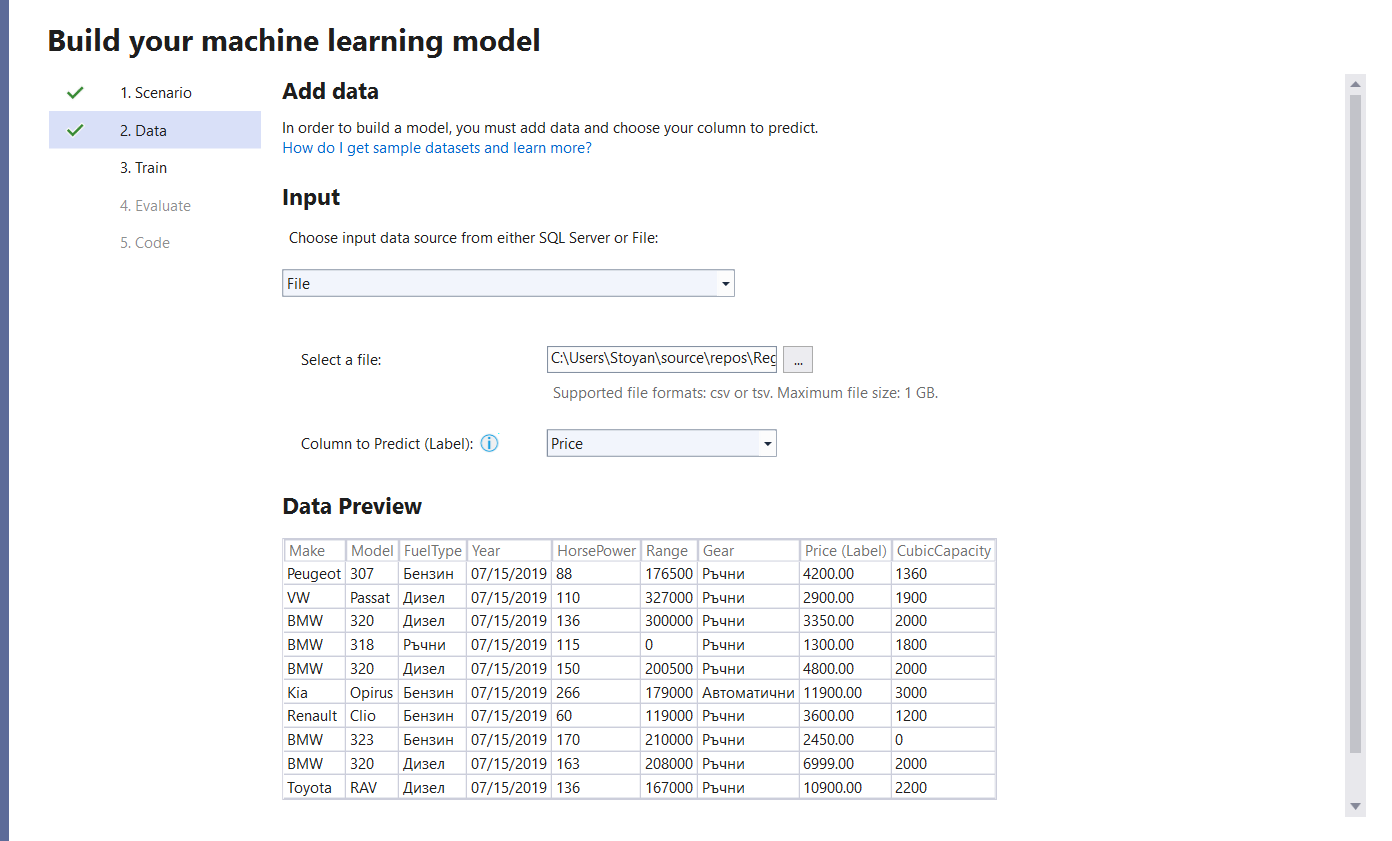
We want to predict the price of a car, therefore we will choose price prediction



### Load Data

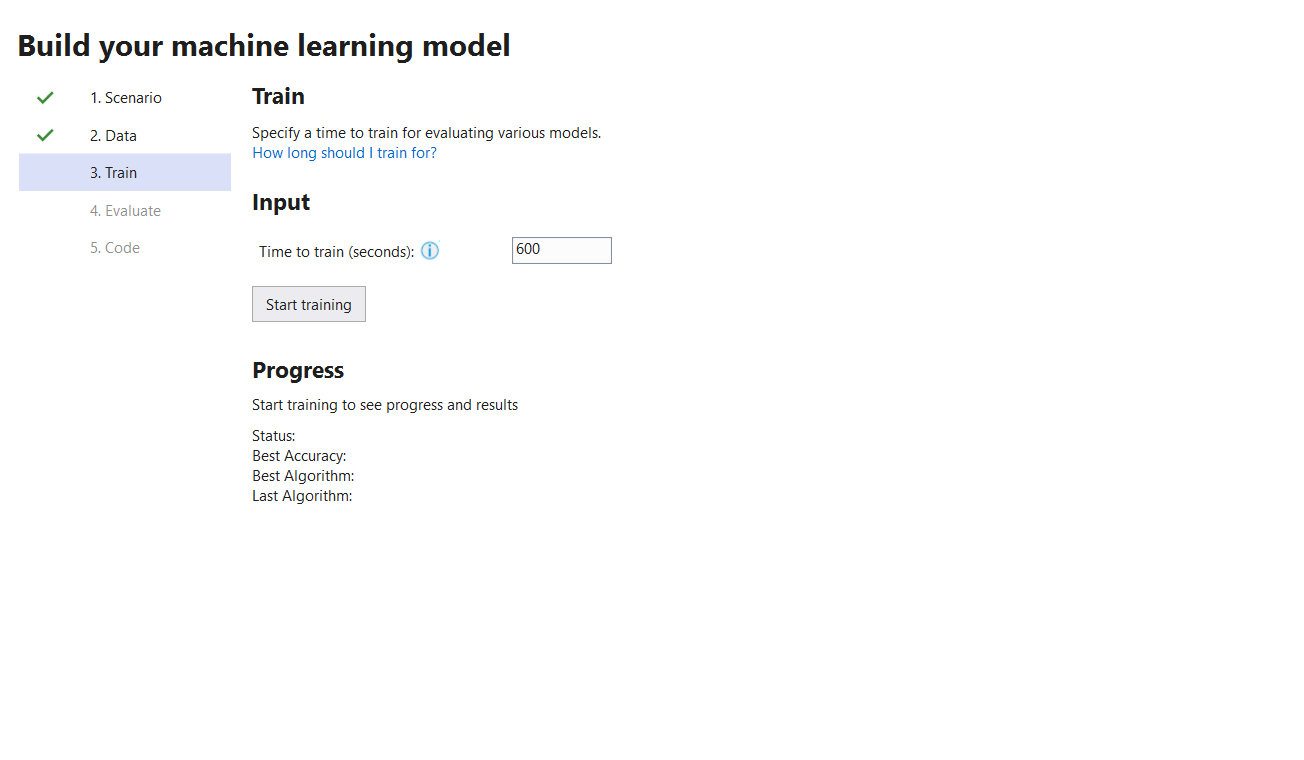
Find the carsb.csv file

Chose the column which you want to predict. (In our case the label will be "Price")

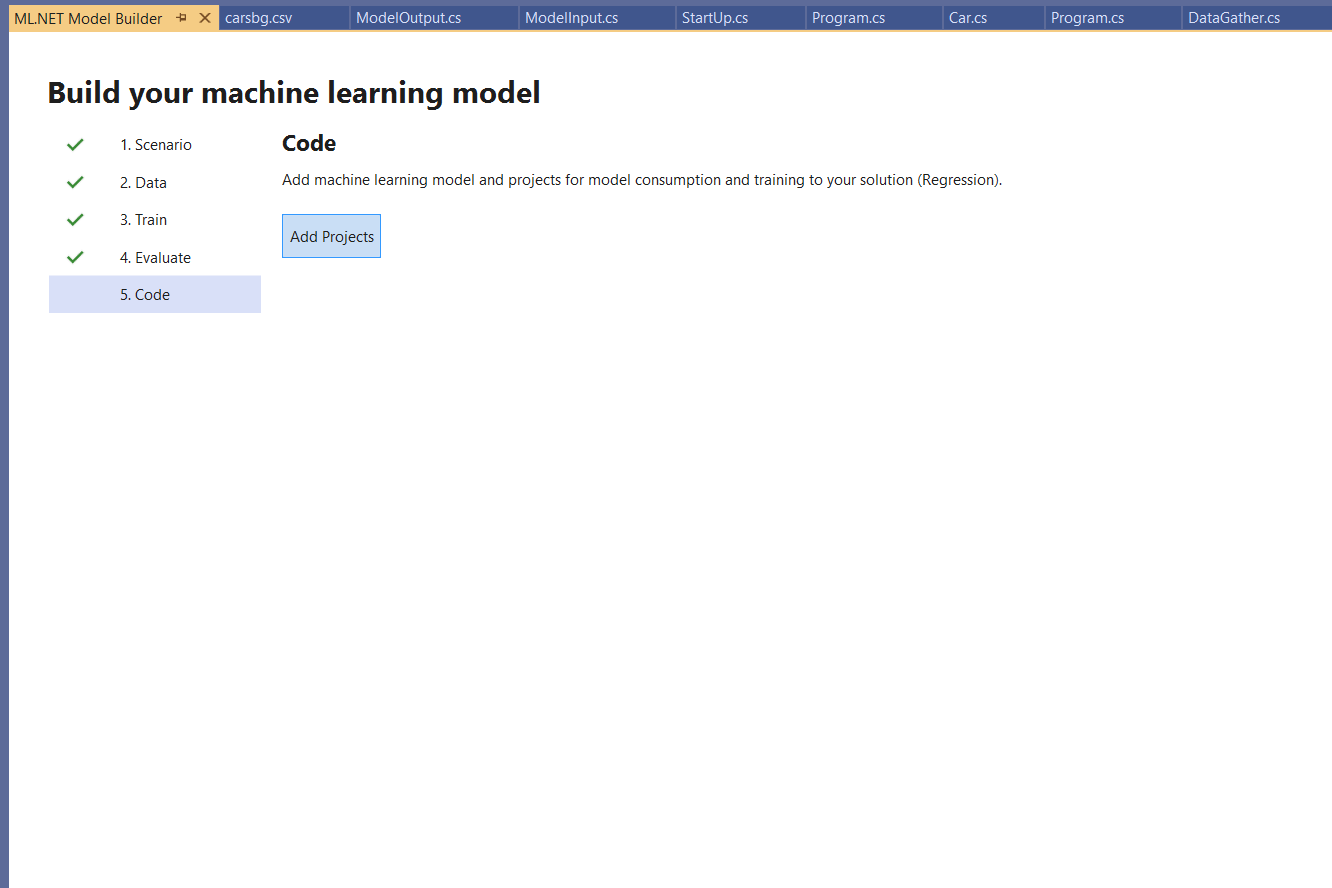


### Train the model

Set time to train to 600 seconds (10 minutes).



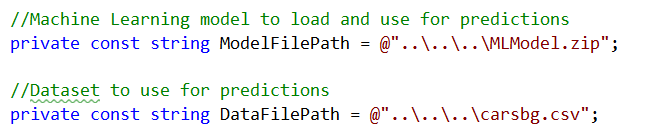
### Evaluate and add projects

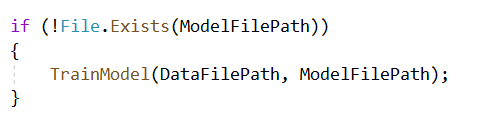


## Refactor Time

### File paths

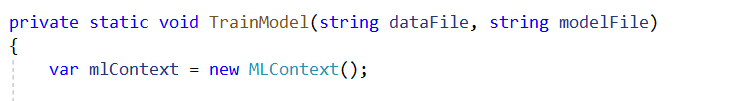
Leave two constants which will keep the Model and File data path.





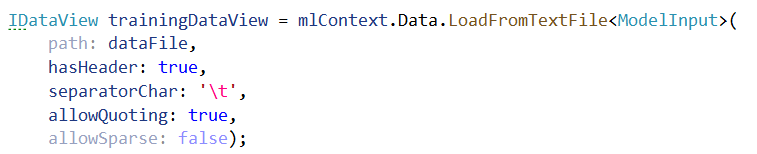
### Train the Model

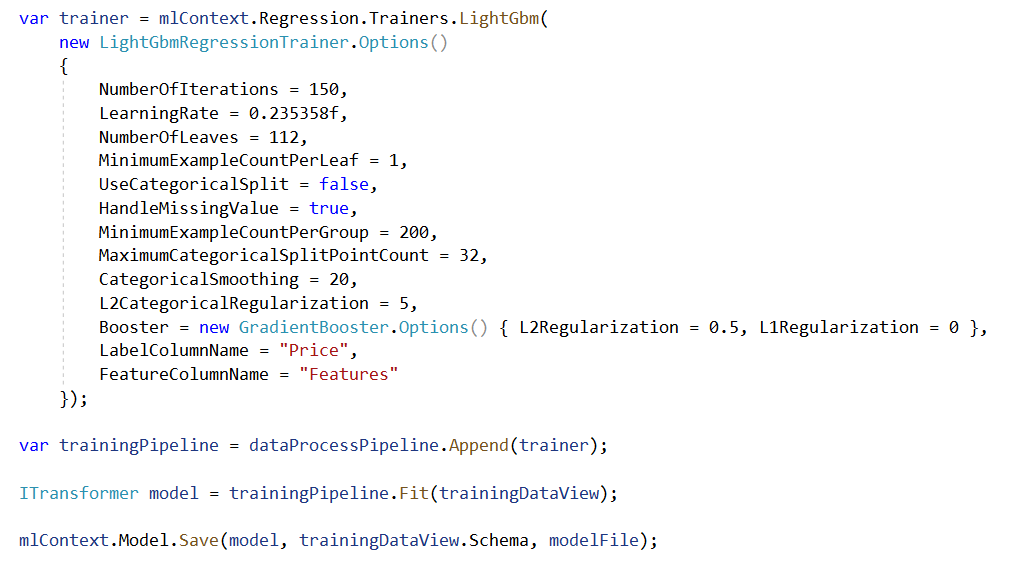
Create a new instance of the MLContext (You can think of MLContext as something DbContext)



### Load the data

Since we use C#, everything is self-explanatory

### What is "OneHotEncoding"? One hot encoding is a process by which categorical variables are converted into a form that could be provided to ML algorithms to do a better job in prediction.

### What is "Features"?

The Label is what you are predicting, the rest of the columns are Features, which are attributes that help predict the Label. Limiting to most relevant attributes or features will result in providing you the best performance for your models.

### What is "LightGbm"?

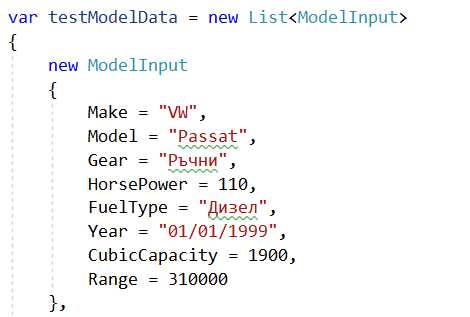
A fast, distributed, high performance gradient boosting (GBT, GBDT, GBRT, GBM or MART) framework based on decision tree algorithms, used for ranking, classification and many other machine learning tasks.

### What does "Fit" do?

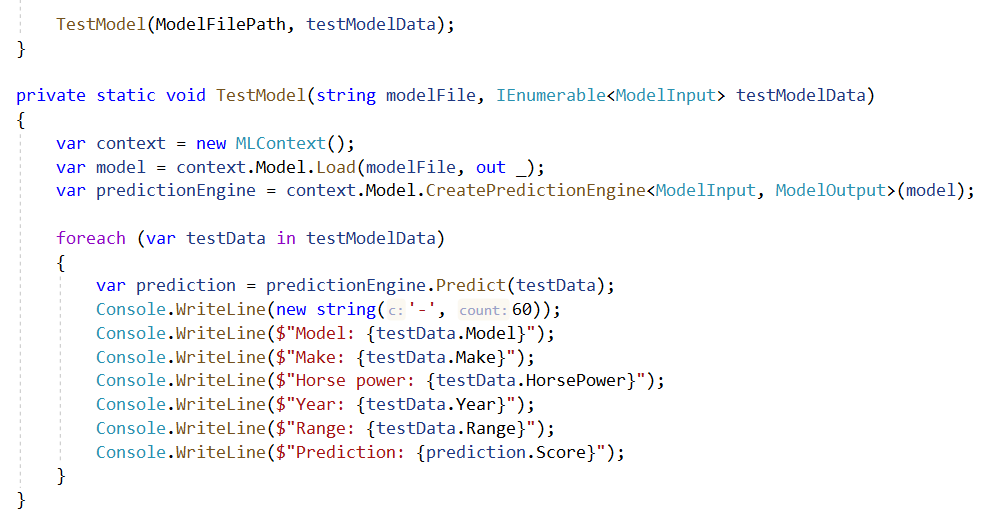
Calling Fit() uses the input training data to estimate the parameters of the model. This is known as training the model.

## Test Model

### Create a collection of the "ModelInput":



### Create a "TestModel" method:



## Result

