



数据科学基础 I (Matlab)

—— 东北大学 ——





回归工具箱



操作步骤



导入数据

```
load carbig  
cartable = table(Acceleration, Cylinders, Displacement,...  
Horsepower, Model_Year, Weight, Origin, MPG);
```

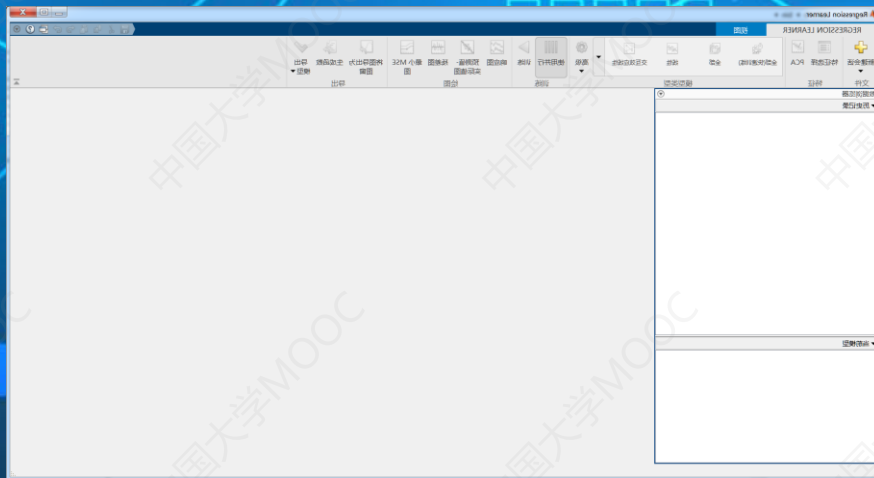



回归工具箱



操作步骤

在MATLAB的APP菜单中打开“regression learner”工具箱





回归工具箱



操作步骤



新建session设置输入属性，输入标注，验证方式。

The 'New Session' dialog box is shown with the following settings:

- Data set:**
 - Workspace Variable: cartable (406x8 table)
 - Response: MPG (double, 9 - 46.6)
- Predictors:**

	Name	Type	Range
<input checked="" type="checkbox"/>	Acceleration	double	8 - 24.8
<input checked="" type="checkbox"/>	Cylinders	double	3 - 8
<input checked="" type="checkbox"/>	Displacement	double	68 - 455
<input checked="" type="checkbox"/>	Horsepower	double	46 - 230
<input checked="" type="checkbox"/>	Model_Year	double	70 - 82
<input checked="" type="checkbox"/>	Weight	double	1613 - 5140
<input checked="" type="checkbox"/>	Origin	char	7 unique
<input type="checkbox"/>	MPG	double	9 - 46.6
- Validation:**
 - ☒ Cross-Validation: Protects against overfitting by partitioning the data set into folds and estimating accuracy on each fold. Cross-validation folds: 5 folds.
 - ☐ Holdout Validation: Recommended for large data sets. Percent held out: 25%.
 - ☐ No Validation: No protection against overfitting.

Buttons: Add All, Remove All, How to prepare data, Start Session, Cancel.



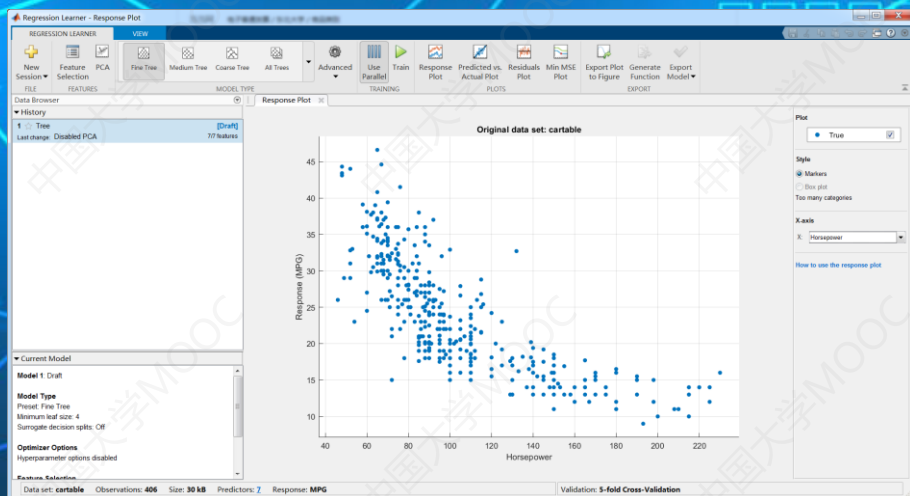
回归工具箱



操作步骤



开始session。通过散点图分析数据特征。





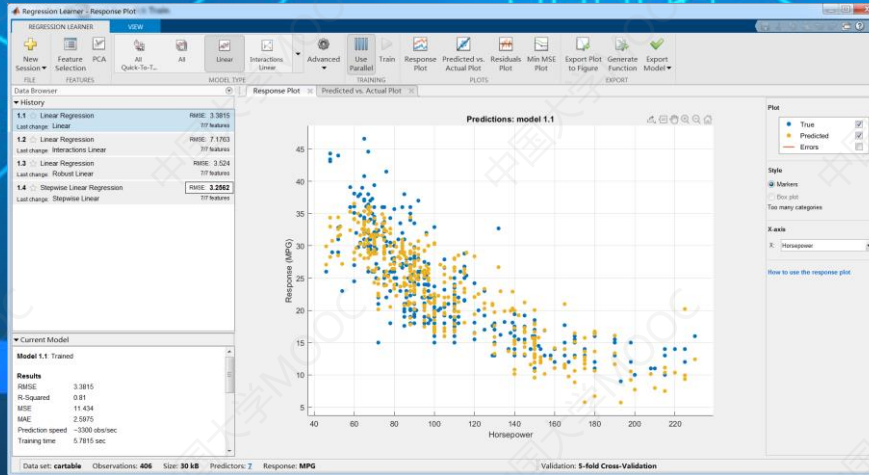
回归工具箱



操作步骤



选择训练算法（例如All Linear），点击“Train”。





回归工具箱



操作步骤



选择不同的算法查看结果，可输出Model进一步分析和预测。

