Homework_Regression 1_Used Car Sales

CS

❓ Assistance

Routine	Description
📋 importFiles	Import file(s) into H_2O
⊞ importSqlTable	Import SQL table into H_2O
⊞ getFrames	Get a list of frames in H_2O
✂ splitFrame	Split a frame into two or more frames
🔗 mergeFrames	Merge two frames into one
⊛ getModels	Get a list of models in H_2O
⦙⦙⦙ getGrids	Get a list of grid search results in H_2O
⚡ getPredictions	Get a list of predictions in H_2O
☰ getJobs	Get a list of jobs running in H_2O
🗚 runAutoML	Automatically train and tune many models
📦 buildModel	Build a model
📦 importModel	Import a saved model
⚡ predict	Make a prediction

📋 Import Files

Search: /Users/pearlpatelll/Desktop/dataset_regression_car_prices.csv   🔍

Search Results: (All files added)

Selected Files: 1 file selected: Clear All

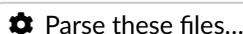✕ /Users/pearlpatelll/Desktop/dataset_regression_car_prices.csv

Actions: ☁ Import

☁ 1 / 1 files imported.

Files   🔑 /Users/pearlpatelll/Desktop/dataset_regression_car_prices.csv

Actions   ⚙ Parse these files...

⚙ Setup Parse

PARSE CONFIGURATION

Sources	🔑 nfs://Users/pearlpatelll/Desktop/dataset_regression_car_prices.csv
ID	dataset_regression_car_prices.hex
Parser	CSV ⬍
Separator	,: '044' ⬍
Escape Character	0
Column Headers	◯ Auto
	⦿ First row contains column names
	◯ First row contains data
Options	☐ Enable single quotes as a field quotation character
	☑ Delete on done

EDIT COLUMN NAMES AND TYPES

Search by column names

1	car_Name	Enum ⬍	ritz	sx4	ciaz	wagon r	swift	vitara brezza	ciaz	s c
2	year	Numeric ⬍	2014	2013	2017	2011	2014	2018	2015	2
3	selling_pri	Numeric ⬍	3.35	4.75	7.25	2.85	4.6	9.25	6.75	6
4	present_pri	Numeric ⬍	5.59	9.54	9.85	4.15	6.87	9.83	8.12	8
5	kms_driven	Numeric ⬍	27000	43000	6900	5200	42450	2071	18796	3
6	fuel_type	Enum ⬍	Petrol	Diesel	Petrol	Petrol	Diesel	Diesel	Petrol	D
7	seller_type	Enum ⬍	Dealer	Dealer	Dealer	Dealer	Dealer	Dealer	Dealer	D
8	transmissic	Enum ⬍	Manual	Manual	Manual	Manual	Manual	Manual	Manual	M
9	owner	Numeric ⬍	0	0	0	0	0	0	0	0

← Previous page   → Next page

⊟ Parse

⠿ Job

Run Time	00:00:00.459

Remaining Time	00:00:00.0
Type	Frame
Key	🔍 dataset_regression_car_prices.hex
Description	Parse
Status	DONE
Progress	100%
	Done.
Actions	🔍 View

⊞ dataset_regression_car_prices.hex

Actions: | ▦ View Data | ✖ Split | ◈ Build Model | ⛀ Run AutoML | | 🗑 Delete |
⚡ Predict | ☁ Download | 🖫 Export |

Rows	Columns	Compressed Size
301	9	16KB

▼ C O L U M N   S U M M A R I E S

label	type	Missing	Zeros	+Inf	-Inf	min	max	mean	sigma	ca
car_Name	enum	0	1	0	0	0	97.0	·	·	
year	int	0	0	0	0	2003.0	2018.0	2013.6279	2.8916	
selling_price	real	0	0	0	0	0.1000	35.0	4.6613	5.0828	
present_price	real	0	0	0	0	0.3200	92.6000	7.6285	8.6441	
kms_driven	int	0	0	0	0	500.0	500000.0	36947.2060	38886.8839	
fuel_type	enum	0	2	0	0	0	2.0	·	·	
seller_type	enum	0	195	0	0	0	1.0	0.3522	0.4784	
transmission	enum	0	40	0	0	0	1.0	0.8671	0.3400	
owner	int	0	290	0	0	0	3.0	0.0432	0.2479	

← Previous 20 Columns   → Next 20 Columns

▸ CHUNK COMPRESSION SUMMARY

▸ FRAME DISTRIBUTION SUMMARY

🔀 Run AutoML

PARAMETERS

project_name

distribution   AUTO ⬍

*training_frame**   dataset_regression_car_prices.hex

*response_column**   selling_price ⬍

validation_frame   (Choose...)

blending_frame   (Choose...)

leaderboard_frame   (Choose...)

A D V A N C E D

nfolds -1

balance_classes ☐

custom_metric_func _____

fold_column [(Choose...) ▲▼]

weights_column [(Choose...) ▲▼]

ignored_columns

☑ car_Name	ENUM(98)
☐ year	INT
☐ selling_price	REAL
☐ present_price	REAL
☐ kms_driven	INT
☐ fuel_type	ENUM(3)
☐ seller_type	ENUM(2)
☐ transmission	ENUM(2)
☐ owner	INT

(☑ All)   (☐ None)

Only show columns with more than _____ % missing values.

sort_metric [RMSE ▲▼]

seed 1234

max_models 0

max_runtime_secs 0

max_runtime_secs_per_model 0

stopping_rounds 3

stopping_metric AUTO

stopping_tolerance -1

exclude_algos

☐ GLM

☐ DRF

☐ GBM

☐ DeepLearning

☐ StackedEnsemble

☐ XGBoost

(☑ All)   (☐ None)

exploitation_ratio   -1

monotone_constraints   [Choose... ⇕] ↔ [⇕]   **+**

EXPERT

keep_cross_validation_predictions ☑

keep_cross_validation_models ☑

keep_cross_validation_fold_assignment ☐

export_checkpoints_dir

⊹⊟ **Build Models**

⋮☰ Job

Run Time	01:00:01.897
Remaining Time	00:00:00.0
Type	AutoML
Key	🔍 AutoML_2_20260207_73027@@selling_price
Description	AutoML build
Status	DONE
Progress	100%
	Done.
Actions	🔍 View

⋮☰ Leaderboard

🔄 **Monitor Live**

▼ **M O D E L S**

models sorted in order of rmse, best first

	model_id	rmse	mse
0	DeepLearning_grid_1_AutoML_2_20260207_73027_model_11	0.858368057346554	0.73679
1	DeepLearning_grid_1_AutoML_2_20260207_73027_model_26	0.9029458983491231	0.8153
2	StackedEnsemble_BestOfFamily_4_AutoML_2_20260207_73027	0.9040401261231578	0.8172
3	DeepLearning_grid_2_AutoML_2_20260207_73027_model_4	0.9157486710550706	0.8385
4	DeepLearning_grid_1_AutoML_2_20260207_73027_model_9	0.9269266047355292	0.8591
5	DeepLearning_grid_3_AutoML_2_20260207_73027_model_4	0.9421102899878816	0.8875
6	DeepLearning_grid_1_AutoML_2_20260207_73027_model_21	0.9768917574385103	0.9543
7	DeepLearning_grid_1_AutoML_2_20260207_73027_model_17	0.9777523889625872	0.9559
8	DeepLearning_grid_1_AutoML_2_20260207_73027_model_20	0.981975134577053	0.9642
9	DeepLearning_grid_1_AutoML_2_20260207_73027_model_2	1.0145300286118157	1.0292
10	DeepLearning_grid_1_AutoML_2_20260207_73027_model_22	1.0209428360288055	1.0423
11	DeepLearning_grid_1_AutoML_2_20260207_73027_model_5	1.0491695202789144	1.1007
12	DeepLearning_grid_1_AutoML_2_20260207_73027_model_4	1.0525405388154125	1.1078
13	DeepLearning_grid_1_AutoML_2_20260207_73027_model_16	1.056491670770874	1.1161
14	DeepLearning_grid_1_AutoML_2_20260207_73027_model_24	1.098766190247261	1.2072

15	DeepLearning_grid_1_AutoML_2_20260207_73027_model_1	1.0995634987523184	1.2090?
16	DeepLearning_grid_1_AutoML_2_20260207_73027_model_23	1.1034189370483054	1.2175?
17	DeepLearning_grid_1_AutoML_2_20260207_73027_model_10	1.1107228686028463	1.2337(
18	DeepLearning_grid_2_AutoML_2_20260207_73027_model_1	1.1243900471157542	1.2642!
19	DeepLearning_grid_1_AutoML_2_20260207_73027_model_19	1.1264052285774329	1.2687?
20	DeepLearning_grid_1_AutoML_2_20260207_73027_model_18	1.138161335976187	1.2954,
21	DeepLearning_grid_1_AutoML_2_20260207_73027_model_25	1.143434821752001	1.3074·
22	DeepLearning_grid_1_AutoML_2_20260207_73027_model_15	1.1472093546947448	1.3160{
23	DeepLearning_grid_1_AutoML_2_20260207_73027_model_8	1.1484582367994756	1.3189!

▼ E V E N T L O G

Actions taken and discoveries made by AutoML

	timestamp	level	stage	message
0	07:30:27.913	INFO	Workflow	Project: AutoML_2_20260207_73027
1	07:30:27.917	INFO	Validation	5-fold cross-validation will be used.
2	07:30:27.918	INFO	Validation	User didn't set any runtime constraints (max r using default 1h time limit
3	07:30:27.920	INFO	Validation	Setting stopping tolerance adaptively based on 0.05
4	07:30:27.920	INFO	Validation	Build control seed: 1234
5	07:30:27.925	INFO	DataImport	training frame: Frame key: AutoML_2_20260207_73027_training_dataset_regre cols: 9 rows: 301 chunks: 1 size: 16217 checks
6	07:30:27.925	INFO	DataImport	validation frame: NULL
7	07:30:27.925	INFO	DataImport	leaderboard frame: NULL
8	07:30:27.925	INFO	DataImport	blending frame: NULL
9	07:30:27.925	INFO	DataImport	response column: selling_price
10	07:30:27.925	INFO	DataImport	fold column: null
11	07:30:27.925	INFO	DataImport	weights column: null
12	07:30:27.950	INFO	Workflow	Loading execution steps: [{XGBoost : [def_2 (1 10w), def_3 (3g, 10w), grid_1 (4g, 90w), lr_se : [def_1 (1g, 10w)]}, {DRF : [def_1 (2g, 10w), : [def_5 (1g, 10w), def_2 (2g, 10w), def_3 (2g 10w), def_1 (3g, 10w), grid_1 (4g, 60w), lr_an {DeepLearning : [def_1 (3g, 10w), grid_1 (4g, grid_3 (5g, 30w)]}, {completion : [resume_best {StackedEnsemble : [best_of_family_1 (1g, 5w), 5w), best_of_family_3 (3g, 5w), best_of_family best_of_family_5 (5g, 5w), all_2 (2g, 10w), al (4g, 10w), all_5 (5g, 10w), monotonic (6g, 10w