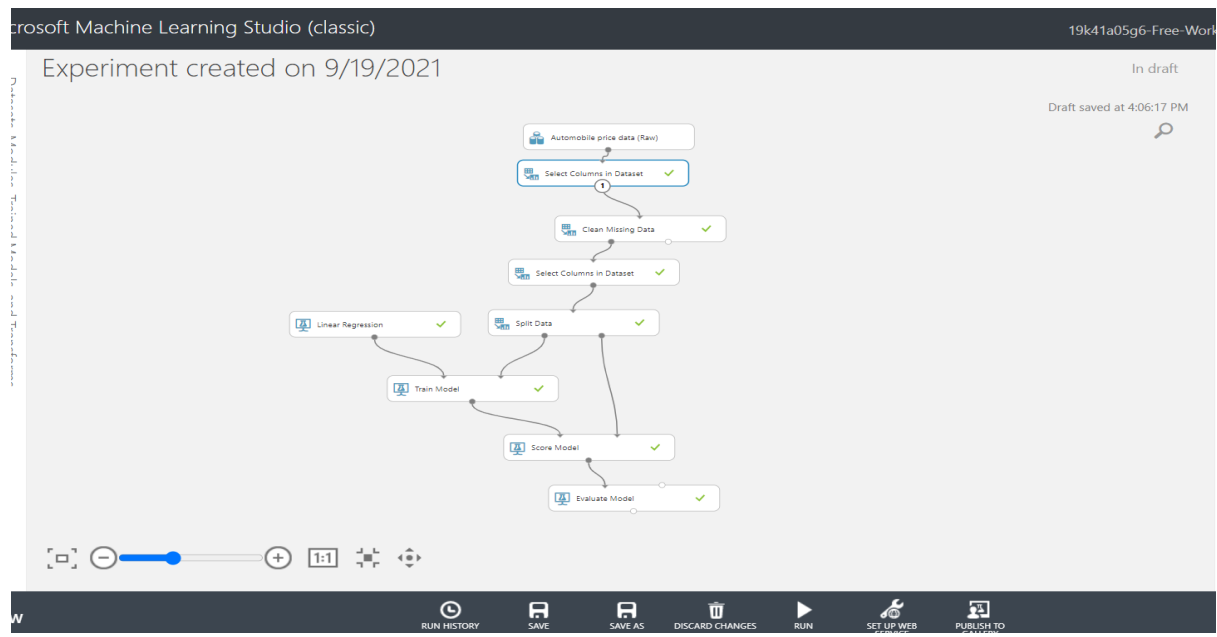


# AI Assignment

19K41A05G6

## Machine learning project workflow:

### Workflow:



### Explore Data:

Experiment created on 9/19/2021 › Automobile price data (Raw) › dataset

rows	columns	symboling	normalized-losses	make	fuel-type	aspiration	num-of-doors	body-style	drive-wheels	engine-location	wheel-base	length	width	height	curb-weight	engine-type	num-of-cylinders
205	26																
3				alfa-romero	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four
3				alfa-romero	gas	std	two	convertible	rwd	front	88.6	168.8	64.1	48.8	2548	dohc	four
1				alfa-romero	gas	std	two	hatchback	rwd	front	94.5	171.2	65.5	52.4	2823	ohcv	six
2	164			audi	gas	std	four	sedan	fwd	front	99.8	176.6	66.2	54.3	2337	ohc	four
2	164			audi	gas	std	four	sedan	4wd	front	99.4	176.6	66.4	54.3	2824	ohc	five
2				audi	gas	std	two	sedan	fwd	front	99.8	177.3	66.3	53.1	2507	ohc	five
1	158			audi	gas	std	four	sedan	fwd	front	105.8	192.7	71.4	55.7	2844	ohc	five
1				audi	gas	std	four	wagon	fwd	front	105.8	192.7	71.4	55.7	2954	ohc	five
1	158			audi	gas	std	four	sedan	fwd	front	105.8	192.7	71.4	55.7	2844	ohc	five

Microsoft Machine Learning Studio (classic)

Experiment created on 9/19/2021

In draft

Draft saved at 4:06:17 PM

Properties Project

Select Columns in Dataset

Select columns

Selected columns: All columns

Exclude column names: normalized-losses

Launch column selector

START TIME 9/19/20...

END TIME 9/19/20...

ELAPSED TIME 0:00:00.0...

STATUS CODE Finished

STATUS DETAILS Task output was present in output cache

Quick Help

Selects columns to include or exclude from a dataset in an operation. Formerly known as Project Columns. (more help...)

Automobile price data (Raw)

Select Columns in Dataset

Clean Missing Data

Select Columns in Dataset

Split Data

Linear Regression

Train Model

Score Model

Evaluate Model

NEW

RUN HISTORY

SAVE

SAVE AS

DISCARD CHANGES

RUN

SET UP WEB SERVICE

PUBLISH TO GALLERY

## Data Cleaning:

Microsoft Machine Learning Studio (classic)

Experiment created on 9/19/2021

In draft

Draft saved at 4:06:17 PM

Properties Project

Clean Missing Data

Columns to be cleaned

Selected columns: All columns

Launch column selector

Minimum missing value... 0

Maximum missing value... 1

Cleaning mode: Remove entire row

START TIME 9/19/20...

END TIME 9/19/20...

ELAPSED TIME 0:00:00.0...

STATUS CODE Finished

Quick Help

Specifies how to handle the values missing from a dataset (more help...)

Automobile price data (Raw)

Select Columns in Dataset

Clean Missing Data

Select Columns in Dataset

Split Data

Linear Regression

Train Model

Score Model

Evaluate Model

NEW

RUN HISTORY

SAVE

SAVE AS

DISCARD CHANGES

RUN

SET UP WEB SERVICE

PUBLISH TO GALLERY

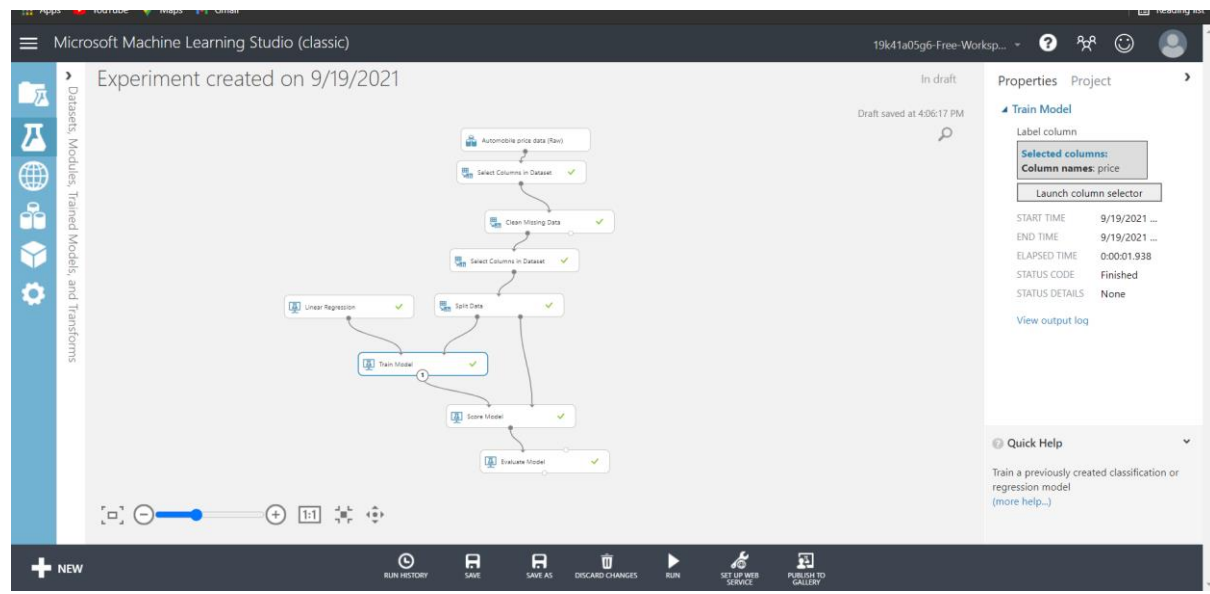
# Split Data:

The screenshot displays the Microsoft Machine Learning Studio (classic) interface. The main workspace shows a workflow diagram with the following steps: 'Automobile price data (Raw)' -> 'Select Columns in Dataset' -> 'Clean Missing Data' -> 'Select Columns in Dataset' -> 'Split Data' -> 'Linear Regression' -> 'Train Model' -> 'Score Model' -> 'Evaluate Model'. The 'Split Data' module is highlighted, and its properties are shown on the right. The 'Splitting mode' is set to 'Split Rows', 'Fraction of rows in the ...' is 0.70, 'Randomized split' is checked, and 'Random seed' is 0. The 'Stratified split' is set to 'False'. The 'Properties' pane also shows 'START TIME', 'END TIME', 'ELAPSED TIME', 'STATUS CODE', and 'STATUS DETAILS'. A 'Quick Help' section is visible at the bottom of the properties pane.

# Linear Regression:

The screenshot displays the Microsoft Machine Learning Studio (classic) interface. The main workspace shows a workflow diagram with the following steps: 'Automobile price data (Raw)' -> 'Select Columns in Dataset' -> 'Clean Missing Data' -> 'Select Columns in Dataset' -> 'Split Data' -> 'Linear Regression' -> 'Train Model' -> 'Score Model' -> 'Evaluate Model'. The 'Linear Regression' module is highlighted, and its properties are shown on the right. The 'Solution method' is set to 'Ordinary Least Squares', 'L2 regularization weight' is 0.001, 'Include intercept te...' is checked, and 'Allow unknown cat...' is checked. The 'Properties' pane also shows 'START TIME', 'END TIME', 'ELAPSED TIME', 'STATUS CODE', and 'STATUS DETAILS'. A 'Quick Help' section is visible at the bottom of the properties pane.

# Model Training and Algorithm:

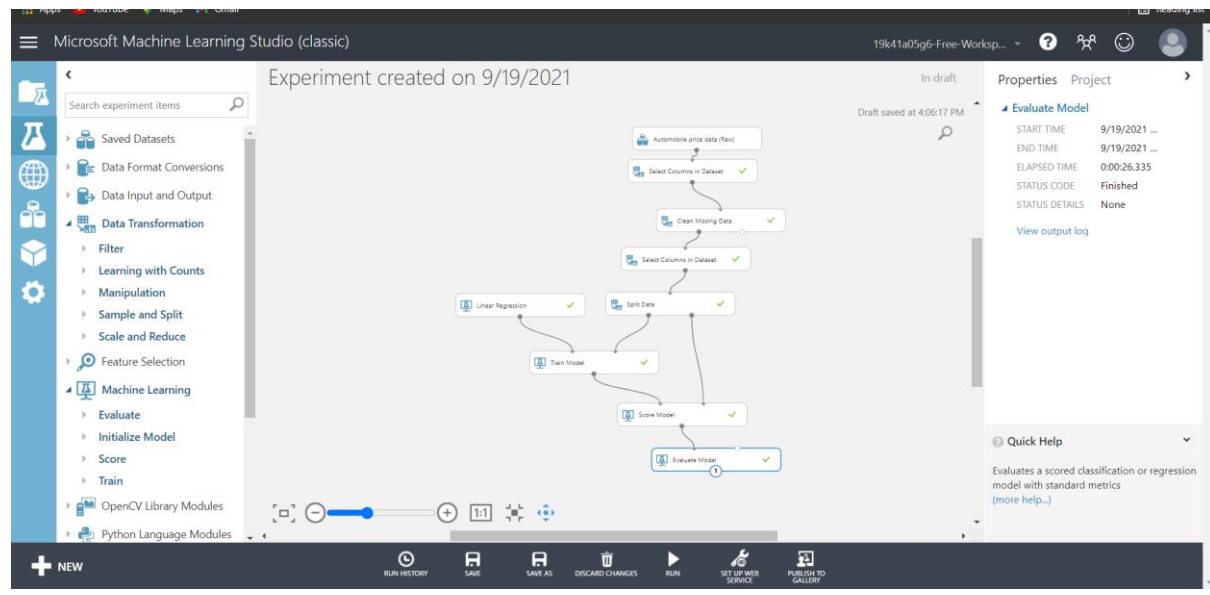


# Score Model and Evaluation:

Experiment created on 9/19/2021 > Score Model > Scored dataset

rows: 58, columns: 26

th	height	curb-weight	engine-type	num-of-cylinders	engine-size	fuel-system	bore	stroke	compression-ratio	horsepower	peak-rpm	city-mpg	highway-mpg	price	Scored Labels
5	56.1	2758	ohc	four	121	mpfi	3.54	3.07	9.3	110	5250	21	28	15510	13284.899402
3	50.8	2145	ohc	four	98	spdi	3.03	3.39	7.6	102	5500	24	30	7689	7658.867692
5	55.7	2212	ohc	four	109	mpfi	3.19	3.4	9	85	5250	27	34	8195	8863.220012
5	52	1874	ohc	four	90	2bbl	3.03	3.11	9.6	70	5400	38	43	6295	5905.454253
2	51.4	2734	ohc	four	119	spfi	3.43	3.23	9.2	90	5000	24	29	11048	10750.34708
7	56.5	3740	ohcv	eight	234	mpfi	3.46	3.1	8.3	155	4750	16	18	34184	38369.412133
3	54.5	1889	ohc	four	97	2bbl	3.15	3.29	9.4	69	5200	31	37	5499	5894.125916
3	50.8	2128	ohc	four	98	mpfi	3.03	3.39	7.6	102	5500	24	30	7957	9549.224752
9	53.7	3380	ohc	six	209	mpfi	3.62	3.39	8	182	5400	16	22	41315	30915.948926



# Results:

Experiment created on 9/19/2021 > Evaluate Model > Evaluation results

## Metrics

Mean Absolute Error	1605.514464
Root Mean Squared Error	2385.271889
Relative Absolute Error	0.266248
Relative Squared Error	0.083112
Coefficient of Determination	0.916888

## Error Histogram

Experiment created on 9/19/2021 > Evaluate Model > Evaluation results

## Metrics

## Error Histogram

