[ML’ 25] Simple Linear Regression Assignment

- In this assignment, you will use linear regression tools in Python to predict the expected daily revenue of a cafe based on an exploration of a number of factors.
- **Dataset Header**:

RevenuePerDay	NCustomersPerDay	AverageOrderValue	WorkingHoursPerDay	NEmployees	MarketingSpendPerDay	LocationFootTraffic
1547.81	152	6.74	14	4	106.62	97
2084.68	485	4.5	12	8	57.83	744
3118.39	398	9.09	6	6	91.76	636
2912.2	320	8.48	17	4	462.63	770
1663.42	156	7.44	17	2	412.52	232
1155.18	121	8.88	6	9	183.49	484
2179.13	238	9	11	4	331.35	156

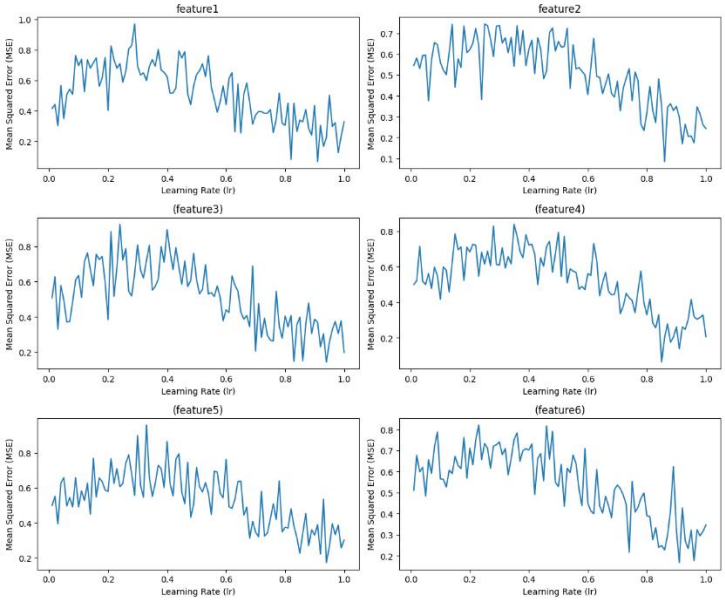
- **Assignment deadline** is **Friday 7/3/2025 11: 59 PM**.
- Apply simple linear regression to predict the revenue per day of a cafe.
- Find the best variable (X) that can be used to predict the performance index (Y).
- This variable (X) can be provided by the dataset as is (for example: use NCustomersPerDay as “X” to predict RevenuePerDay “Y”) **or** it can be a calculated variable (for example: use a new variable, for example: AverageOrderValue+NEmployees as “X” to predict performance index “Y”).

- **Your code must show at least all six basic trials (for the six columns).**
- Use simple linear regression from scratch as discussed in Lab 2 and try adjusting the learning rate and number of epochs as needed.
- You must show your learning rate and epochs trials in the code, Otherwise, it will not count.
- Plot a graph that shows the relation between each learning rate you tried and the mse loss. Plot a similar graph for epochs. Your plots should be included in the code.

❖ You should represent this for each model that you train (i.e., at least 6 models). DO NOT create 6 separate plots for lr/mse and 6 separate plots for epochs/mse. You can either have one plot for lr/mse that has 6 subplots and another for epochs/mse,, or you can have one plot that has 6 lines for lr and another for epochs. If you have more models, you can have more subplots/lines than 6 but do not create 12+ separate plots.

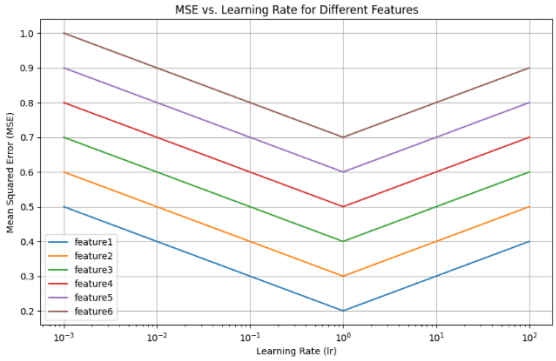
Example:

A single plot that has 6 subplots:



OR

A single plot that has 6 lines with legend:



Either way, we need two plots: one for lr/mse and one for epochs/mse

The values of lr and mse in the above plots are a dummy example, and the values are random, you should figure out what values to use for lr.

- DO NOT manually adjust initializations of 'm' and 'c'. They must be initialized by zeros.
- Cheating Detection will be applied and it will not be tolerated.
- **You are expected to deliver the assignment using** a google form that will be announced soon **according to the following instructions**:

  1- Upload a single "YourStudentID.py" code file containing the code for the six models as well as the trials and prints the MSE for each model.

  2- Upload a PDF file (report) containing:

  - Your name, id, department and UMS Level
  - A table that summarizes the mean square error of each model
  - All the plots you generated.
  - Your conclusion the best variable to use for this task.

Take Note that the code file name must be YourStudentID.py (example:20191700123.py) or you will be penalized.

Take Note that the report file name must be YourStudentID.pdf (example: 20191700123.pdf) or you will be penalized.