

Machine Learning, Winter 2022
Practice Assignment 3

Exercise 3-1 Linear Regression
Coding Question

“Advertising.csv” shows a dataset comprised of advertising campaigns. It displays sales (in thousands of units) for a particular product as a function of advertising budgets (in thousands of dollars) for TV, radio, and newspaper media.

- a) Prepare the dataset for use
- b) Plot each *Spend* against *Sales* (individually)
- c) **Train** a sklearn Linear Regression model
- d) **Test** the model
- e) Evaluate the model’s predictions using
 - Mean absolute error
 - Mean squared error
 - Root mean squared error
- f) Using the predictions, produce a Residual Plot (hint: you will need to calculate the residuals beforehand)
- g) Create a new model that trains on the entire dataset and **save** it as “sales_model”
- h) Plot each *Spend* against *Sales* using the original labels as well as the predictions
- i) Plot the residuals for the new model
- j) **Load** the sales_model as “loaded_model”
- k) Show the coefficients of the loaded_model
- l) A campaign will have a total spend of 149k on TV, 22k on Radio, and 12k on Newspaper Ads, how many units could we expect to sell as a result of this?