MLNN Laboratory Session

8 February 2023

You are given data relating sales (gross revenue) at a medium-sized company to advertising spend (all in thousands of GBP per month, adjusted for inflation).

The data is in `inputdata.csv`, an ASCII file that can be inspected with any basic text editor. Each line corresponds to one month. The left and middle columns in this file show the amount spent on advertising, the right column shows the level of sales (revenue).

Split the data into training and test sets (approximately 7:3). Use linear regression to create models of the form

$$S = u + k_1 A_1$$

$$S = u + k_2 A_2$$

and

$$S = u + k_1 A_1 + k_2 A_2$$

where S is the revenue, A_1 and A_2 are amounts spent on online and TV advertising, respectively.

Use the test set to evaluate the rms error of each model and comment on the quality of each model.