MLNN Laboratory Session

Practical ML Intro – Lab 1

Problem 1

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 etc etc). The data is in `inputdata1.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, online and on TV, respectively. The right column shows the level of sales (revenue).

Use linear regression to create models of the form
$S=u+k_1A_1$
$S=u+k_2A_2$
and
$S=u+k_1A_1+k_2A_2$
where S is the revenue, A_1 and A_2 are amounts spent on online and TV advertising, respectively.

Evaluate MSE and comment on the quality of each model.

Problem 2

You are given data relating one independent variable, X, with one dependent variable, Y. The data is in `inputdata2.csv`. Create the learning curve to find the best polynomial fit for this data.