Supervised Learning Notation

This sheet has some useful notation which you will see throughout the course.

Individuals/Examples

- The number of individuals we have is n. Therefore the individuals are $1, ..., n$
- If we want to talk about a *representative individual* we can use the notation i

Variables/Features

- A variable/features: x
- When there is more than one we use **subscript notation**. $x_1, x_2, ..., x_k$
- The target variable/feature (also called 'class'): y
- We use superscript notation to denote a specific realisation of a variable/feature $x_1^{(i)}, x_2^{(i)}, ..., x_k^{(i)}$ and $y^{(i)}$

Datasets

- A dataset has n individuals, many input features and possible a target variable. We write datasets with the following notation: $\{(x^{(1)}, y^{(1)}), ..., (x^{(n)}, y^{(n)})\}$ where x is a vector of all input features. If you didn't have a target features then the dataset would be $\{(x^{(1)}), ..., (x^{(n)})\}$.

Supervised Learning Models

- Predicted values are given hats whereas real values are not. $\hat{y}^{(i)}$ vs $y^{(i)}$
- A model takes input features $x_1^{(i)}, x_2^{(i)}, ..., x_k^{(i)}$ and maps them to a predicted output $\hat{y}^{(i)}$
- A model has parameters θ where θ is a set or vector of all parameters.
- A model can be written as a function $f_\theta(x)$ where x is all of the features. This function maps the parameters to the predicted output.
- Loss functions are generally written $J(\theta)$