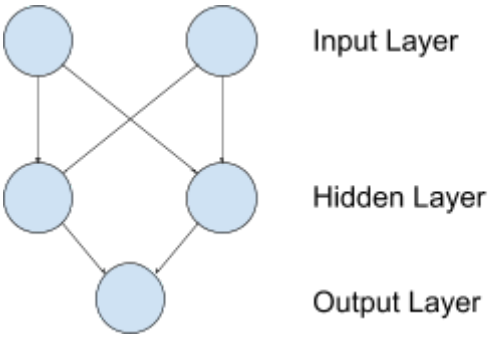
Everything on MLP

Structure

A certain type of neural network, known as multilayer perceptron MLP can learn a function between inputs and their outcome. It works by building the function out of many simple functions joined by weighted connections. The structure of a MLP consists of an input layer, a hidden layer and an output layer. All three layers can consist of multiple nodes. The following graph depicts the structure of a simple MLP



Learning

A neural network uses the data to modify the weighted connections between all of its functions, until it predicts the data accurately. This process is referred to as training the neural network.

Step-by-step MLP training

1) Prepare the data, so it contains predictors and predicted variables.

2) Spit the data into a test set and a training set

3) The MLP starts with random weights. The input data is propagated through each neuron and the output signal of each neuron is calculated. Eventually the output signal of the output neuron/s is calculated and a prediction is made

4) The output y is compared to the target output value z, found in the training set and the difference is called, error signal δ of the output layer neuron

$$\delta = z - y$$

5) The error signal δ is back propagated to all the neurons and the error signal for each neuron is computed

6) When the error signal for each neuron is computed the weights of each input node is modified

$$w'_{ij} = w_{ij} + \eta\delta_j \frac{df_j(e)}{de} x_i$$

7) Repeat until the error no longer reduces

Overfitting

A data mining predictor captures the structure of the data so well, that irrelevant details are picked up and used when they are not true

Advantages of MLPs

Very powerful predictors → Better than rule-based systems

Can cope with : Non-linear relationships
Multiple numeric and nominal variables

Can generalise to data, never seen before

Disadvantages of MLPs

Difficult to understand how predictions are made

No rules to look at

Big errors can occur if not trained properly