Data Mining and Machine Learning

Learning using Error Back-Prop

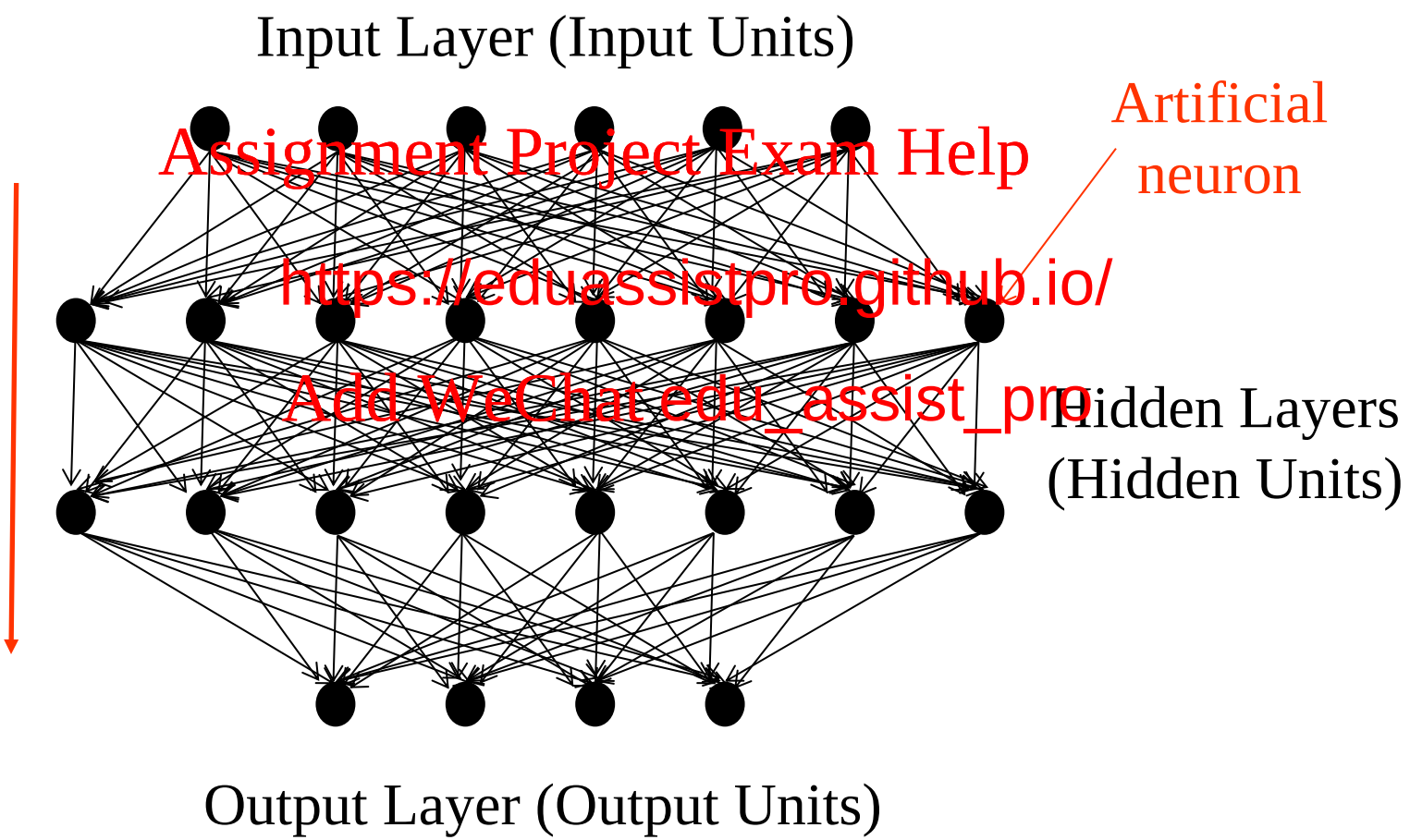
Peter Jančovič

Objectives

- Outline of the MLP training
  - The error function
  - Optimisation by gradient decent
- The Error B
  - Calculating the derivatives
  - Bringing everything togeth
  - Summary of the EBP algorithm
  - Practical considerations

Feed-forward Neural Networks

Multi-Layer Perceptron - Feed-Forward Neural Network



MLP training

- To define an MLP must decide:
  - Number of layers
  - Number of input units
  - Number of hidden units
  - Number of
- Choosing th rs and units is a combination of experience imentation
- Once these are defined, properties of the MLP are <u>completely</u> defined by the values of the weights
- How do we choose the weight values?

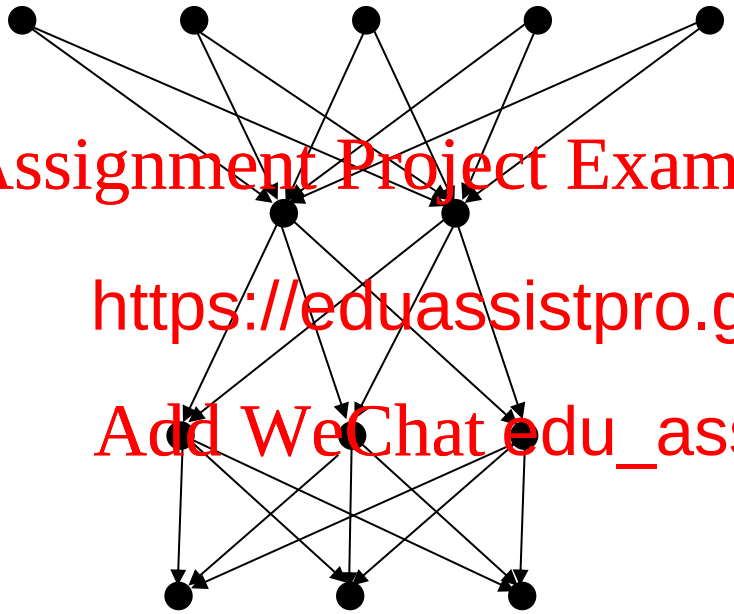
MLP training (continued)

- MLP training needs a set of input vectors ***i*** with corresponding target output vectors ***t*(*i*)**
- Each input vector ***i*** is propagated through the network to p
- The error E is the differen the actual output ***o*(*i*)** and the target o
- Objective of training is to learn the weights which minimise the average error over the training set

$$E = \sum_i (o(i) - t(i))^2$$

Error Back-Propagation

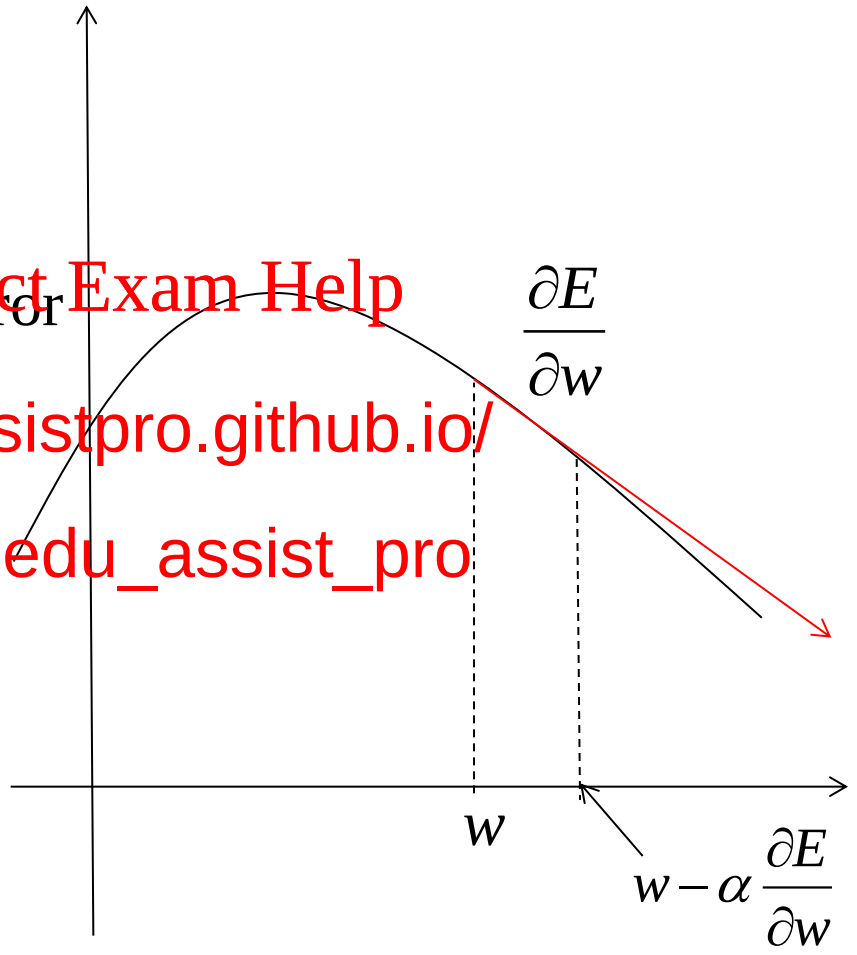
$i = i_1,\ldots,i_5$



$t(i) = t(i)_1,\ldots, t(i)_3$

MLP training (continued)

- MLP training uses gradient descent
- For each weight w calculate $\frac{\partial E}{\partial w}$
- Subtract a proportion of $\frac{\partial E}{\partial w}$ from w



MLP training (continued)

- MLP weights learnt automatically from training data
- Training uses an iterative computational technique called Error-Back-Propagation (EBP)
- There are m

Error-back propagation (EBP)

1. Choose initial values for the weights
2. Propagate each training sample ***i*** through the network to obtain ***o*(*i*)**. Set $\boldsymbol{E = |t(i) - o(i)|}$
3. EBP calculates $\frac{\partial E}{\partial w}$ for each weight w by propagating [illegible] ugh the network
4. When all training pattern [illegible] n seen, w is changed by an amount proportional to the average value of $\frac{\partial E}{\partial w}$
5. Repeat until the change in error falls below a threshold

w_{jk}

Summary

- MLP training
- Error Back P