Neural Networks in the Real World
Functional Programming and Intelligent Algorithms

Prof Hans Georg Schaathun

Høgskolen i Ålesund

12th February 2015

Outline

Six step of neural network application

1. Input and output
2. Normalise input
3. Split the data
4. Network Architecture
5. Train the Network
6. Test the Network

Input and output

1. Feature vector
   - What does your input look like?
2. Output values
   - Discrete or continuous?
   - How do you code class labels?

Normalise input

- Features should be in the ± 1 range
- Scaling may be required
  - use a spreadsheet

Split the data

- Training (e.g. 50%)
- Validation (e.g. 25%)
- Testing (e.g. 25%)

Network Architecture

- How many hidden layers?
- How many nodes per layer?

Train the Network

- Train with different configurations and parameters
  - number of epochs
  - different learning rates (η)
  - different architectures
  - different β in the sigmoid?
- Test each configuration on the *validation set*

Test the Network

- Choose *one* design based on tests with the validation set
- Use the *test set* once only
- Final assessment of your chosen configuration

Outline

HØGSKOLEN
I ÅLESUND
Aalesund University College

Prof Hans Georg Schaathun   Neural Networks in the Real World   12th February 2015   10 / 13

Local and global minima

- Common challenge in all optimisation problems
- Gradient decents finds a minimum
    - some ditch surrounded by higher points
    - local minimum
- Many ditches may exist
    - some lower than others
- Global minimum is hard to find
- Possible remedy: momentum

Batch and sequential training

1. We have implemented sequential training
   - consider one training item at a time
   - update weights for every individual item
2. Batch training considers the entire data set
   - Weight updates are added together
   - Weights are updated when all changes are aggregated

What are the advantages and disadvantages of each approach?

Advantages and disadvantages
Batch training

1. Batch training is more accurate
2. Batch training is more likely to get stuck in a local minimum
3. In sequential training, to get full benefits
   - randomise the order of the training items
4. In batch training, order does not matter

- Compromise: minibatches
  - take a subset of the training set for batch training
  - move on to another (disjoint) subset