CS623: Introduction to Computing with Neural Nets
(lecture-6)

Pushpak Bhattacharyya
Computer Science and Engineering
Department
IIT Bombay
Backpropagation algorithm

- Fully connected feed forward network
- Pure FF network (no jumping of connections over layers)
General Backpropagation Rule

- General weight updating rule:
  \[ \Delta w_{ji} = \eta \delta_j o_i \]

- Where
  \[ \delta_j = (t_j - o_j) o_j (1 - o_j) \]
  for outermost layer

  \[ = \sum_{k \in \text{next layer}} (w_{kj} \delta_k) o_j (1 - o_j) o_i \]
  for hidden layers
How does it work?

• Input propagation forward and error propagation backward (e.g. XOR)

\[ w_1 = 1 \]

\[ w_2 = 1 \]

\[ \theta = 0.5 \]
Local Minima

Due to the Greedy nature of BP, it can get stuck in local minimum $m$ and will never be able to reach the global minimum $g$ as the error can only decrease by weight change.

Figure- Getting Stuck in local minimum

$m$- local minima, $g$- global minima
Momentum factor

1. Introduce momentum factor.

\[(\Delta w_{ji})^{n\text{th - iteration}} = \eta \delta_j O_i + \beta (\Delta w_{ji})^{(n-1)\text{th - iteration}}\]

- Accelerates the movement out of the trough.
- Dampens oscillation inside the trough.
- Choosing \(\beta\): If \(\beta\) is large, we may jump over the minimum.
Symmetry breaking

- If mapping demands different weights, but we start with the same weights everywhere, then BP will never converge.

\[ \theta = 0.5 \]

XOR n/w: if we started with identical weight everywhere, BP will not converge
Example - Character Recognition

- Output layer – 26 neurons (all capital)
- First output neuron has the responsibility of detecting all forms of ‘A’
- Centralized representation of outputs
- In distributed representations, all output neurons participate in output
An application in Medical Domain
Expert System for Skin Diseases Diagnosis

• Bumpiness and scaliness of skin
• Mostly for symptom gathering and for developing diagnosis skills
• Not replacing doctor’s diagnosis
Architecture of the FF NN

- 96-20-10
- 96 input neurons, 20 hidden layer neurons, 10 output neurons
- Inputs: skin disease symptoms and their parameters
  - Location, distribution, shape, arrangement, pattern, number of lesions, presence of an active norder, amount of scale, elevation of papules, color, altered pigmentation, itching, pustules, lymphadenopathy, palmer thickening, results of microscopic examination, presence of herald patch, result of dermatology test called KOH
Output

- 10 neurons indicative of the diseases:
  - psoriasis, pityriasis rubra pilaris, lichen planus, pityriasis rosea, tinea versicolor, dermatophytosis, cutaneous T-cell lymphoma, secondary syphilis, chronic contact dermatitis, soberrheic dermatitis
Training data

• Input specs of 10 model diseases from 250 patients
• 0.5 is some specific symptom value is not known
• Trained using standard error backpropagation algorithm
Testing

- Previously unused symptom and disease data of 99 patients
- Result:
  - Correct diagnosis achieved for 70% of papulosquamous group skin diseases
  - Success rate above 80% for the remaining diseases except for psoriasis
  - Psoriasis diagnosed correctly only in 30% of the cases
  - Psoriasis resembles other diseases within the papulosquamous group of diseases, and is somewhat difficult even for specialists to recognise.
Explanation capability

• Rule based systems reveal the explicit path of reasoning through the textual statements

• Connectionist expert systems reach conclusions through complex, non linear and simultaneous interaction of many units

• Analysing the effect of a single input or a single group of inputs would be difficult and would yield incorrect results
Explanation contd.

• The hidden layer re-represents the data
• Outputs of hidden neurons are neither symptoms nor decisions
Figure: Explanation of dermatophytosis diagnosis using the DESKNET expert system.
Discussion

• Symptoms and parameters contributing to the diagnosis found from the n/w
• Standard deviation, mean and other tests of significance used to arrive at the importance of contributing parameters
• The n/w acts as apprentice to the expert
Exercise

• Find the weakest condition for symmetry breaking. It is not the case that only when ALL weights are equal, the network faces the symmetry problem.