7: Neural Networks – Exercises with AISpace

CSC 4510 – Machine Learning

Dr. Mary-Angela Papalaskari
Department of Computing Sciences
Villanova University

Course website: www.csc.villanova.edu/~map/4510/
Neural Nets tools in AI Space

• Sample Data and Graph: Mail
  – Examples
  – Set properties
  – Initialize the parameters
  – Solve
  – How do we use it? Calculate output
A small example: “Which class to take”

• Inputs?
• Outputs?
• Sample data
Some Examples

• **Example 1:**
  – 3 inputs, 1 output, all binary

• **Example 2:**
  – same inputs, output inverted
Getting the right inputs

- Example 3
  - Same inputs as 1 and 2
  - Same output as 1
  - Outcomes reversed for half the cases
Getting the right inputs

• Example 3
  – Same inputs as 1 and 2
  – Same output as 1
  – Outcomes reversed for half the cases

• Network is not converging

• The output here cannot be predicted from these inputs.

• Whatever is determining whether to take the class, we haven’t captured it
Representing non-numeric values

• **Example 4**
  – Required is represented as “yes” and “no”
Representing non-numeric values

• Example 4
  – Required is represented as “yes” and “no”

• Actual model still uses 1 and 0; transformation is done by applet.
More non-numeric values

• **Example 5**

• Workload is low, med, high: text values but they can be ordered.
More non-numeric values

• Example 5
• Workload is low, med, high: text values but they can be ordered.

• Applet asks us to assign values. 1, 0.5, 0 is typical.
Unordered values

- **Example 6**
  - Input variables here include professor
  - Non-numeric, can’t be ordered.
Unordered values

• Example 6
  – Input variables here include professor
  – Non-numeric, can’t be ordered.

• Still need numeric values
• Solution is to treat $n$ possible values as $n$ separate binary values
• Again, applet does this for us
Variables with more values

• **Example 7**
  – GPA and number of classes taken are integer values

• Takes considerably longer to solve
• Looks for a while like it’s not converging
And Reals

- **Example 8**
  - GPA is a real.

- Takes about 20,000 steps to converge
And multiple outputs

- Small Car database from Alspace
- For any given input case, you will get a value for each possible outcome.
- Typical for, for instance, character recognition.
Training and Test Cases

• The basic training approach will fit the training data as closely as possible.

• But we really want something that will generalize to other cases.

• This is why we have test cases.
  – The training cases are used to compute the weights
  – The test cases tell us how well they generalize

• Both training and test cases should represent the overall population as well as possible.
Representative Training Cases

• **Example 9**
  – Training cases and test cases are similar
Representative Training Cases

• Example 9
  – Training cases and test cases are similar
  – (actually identical...)

• Training error and test error are comparable
Non-Representative Training Cases

• **Example 10**
  – Training cases and test cases represent different circumstances
  – We’ve missed including any cases involving Lee in the training

• Training error goes down, but test error goes up.
• In reality these are probably bad training AND test cases; neither seems representative.
So:

• Getting a good ANN still involves understanding your domain and capturing knowledge about it
  – choosing the right inputs and outputs
  – choosing representative training and test sets
  – Beware “convenient” training sets
• You can represent any kind of variable: numeric or not, ordered or not.
• Not every set of variables and training cases will produce something that can be trained.
Once it’s trained...

• When your ANN is trained, you can feed it a specific set of inputs and get one or more outputs.

• These outputs are typically interpreted as some decision:
  – take the class
  – this is probably a “5”

• The network itself is black box.

• If the situation changes the ANN should be retrained
  – new variables
  – new values for some variables
  – new patterns of cases
One last note

• These have all been simple cases, as examples
• Most of these examples could in fact be predicted much more easily and cleanly with a decision tree (or even a couple of IF statements!)
• A more typical use for any neural net system has many more inputs and many more training cases
And one last example:

• Our **class data example** (apples or oranges?)

• **Homework:** Use what you learned today with this example, write up your comments and include some screenshots showing your work.

• Due Tuesday 3/20/12