A neural network model for detecting intrusions or attacks on a computer network

Steven Smith

P13167208

Wednesday 11th May - 3pm

What is to come?

- Overview of the problem
- The training data
- The data transformation techniques used
- The method of construction
- The experiments
- The results

Network Intrusion Detection

- Network Security of great importance
  - Our lives are in the web
  - Risks – Litigation, Financial, Injury or Death
  - Hacked systems regularly reported in the press

- Network Intrusion Detection Systems
  - Good/Normal Connections
  - Bad/Attacking Connections

Neural Network based Intrusion Detection

- Supervised neural network to classify network activity
- Neural Network configuration
  - Difficult to determine ideal parameters
  - Trial and error
- Progress in evolutionary approaches
  - Particle swarm optimisation
  - Genetic algorithms

KDD Cup 1999 Data

- The Third International Knowledge Discovery and Data Mining Tools Competition
  - Build predictive model determine bad vs good connections
  - intrusions simulated in a military network environment
- CSV containing 10% of the data
    - 494021 rows
    - 41 Dimension
    - 23 classifications

More about the data

- 38 of 41 dimensions contain numeric data
  - 32 are Interval, Continuous
  - 6 are Discrete, 0 or 1
- 3 of 41 dimensions contain nominal data
  - Protocol_type, service, flag
- Result contains nominal data
  - 23 classes including normal and smurf.

Data Transformation

- 4 nominal columns
  - Lookup script (“if then else” mapping to ordinal value)
- Matlab data transformation functions
  - Each dimension treated equally
  - Mapminmax
  - Mapstd
  - RemoveConstantRows
  - PerformPCA

More Data Transformation

- Result column converted from single value to array
  - number 3 would be represented by [0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
- Maintain distribution in training datasets
  - 280790/494021 ≈ 57% - 'smurf' attacks
  - 97278/494021 ≈ 20% - 'normal' usage
- New generated smaller dataset
  - ≈ 5700/10000 ≈ 57% - 'smurf' attacks
  - ≈ 2000/10000 ≈ 20% - 'normal' usage
  - First x rows of the classification

Method of construction

- Limited time, too may combinations
  - Limited set of perturbable parameters
- Nprtool – Pattern recognition tool
  - ‘patternnet’ defaults
  - 'trainscg' - Scaled conjugate gradient BP
  - ‘crossentropy’ - performance
  - ‘softmax’ – output activation function
  - Random subsampling cross validation (10/90)

Parameters tested

- Network topology, hidden layers, number of neurons in each layer
- The size of the training and validation dataset.
- The normalizing functions - mapminmax and mapstd
- Principal Component analysis - on or off
- Training vs Validation sizes (10/90, 50/50, 90/10, 30/70, 70/30)
- Maximum number of epochs - 100 or 250
- The hidden layer activation functions - 15 functions (compet, elliotsig, hardlim, hardlims, logsig, netinv, poslin, purelin, radbas, radbasn, satlin, satlins, softmax, tansig, tribas).
- Number of folds for random subsampling
  - 10
  - maximum of 100 whilst performance continues to improve.

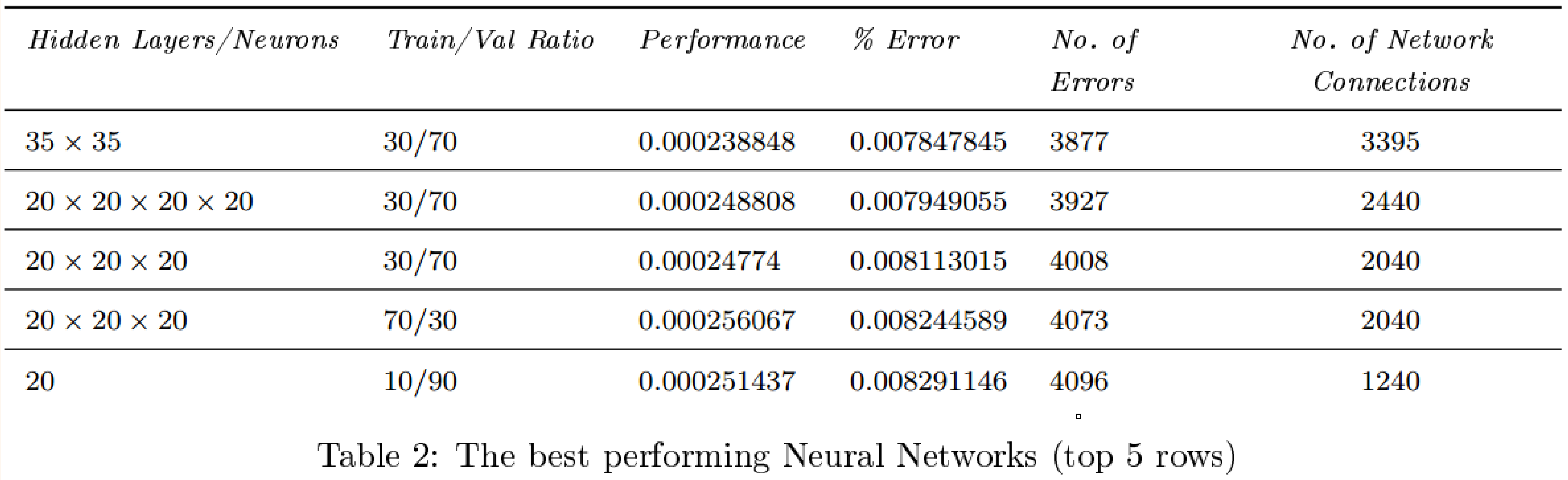
Experiments 1-9

- Test a combination of parameters
  - Aim to reduce options down to a single choice
  - 15 activation functions down to 1
  - 2 normalizing functions down to 1
- Generates a table of data
- Ranking results
  - Percentage Error
  - Cross Entropy Performance

Experiments 10 - 14

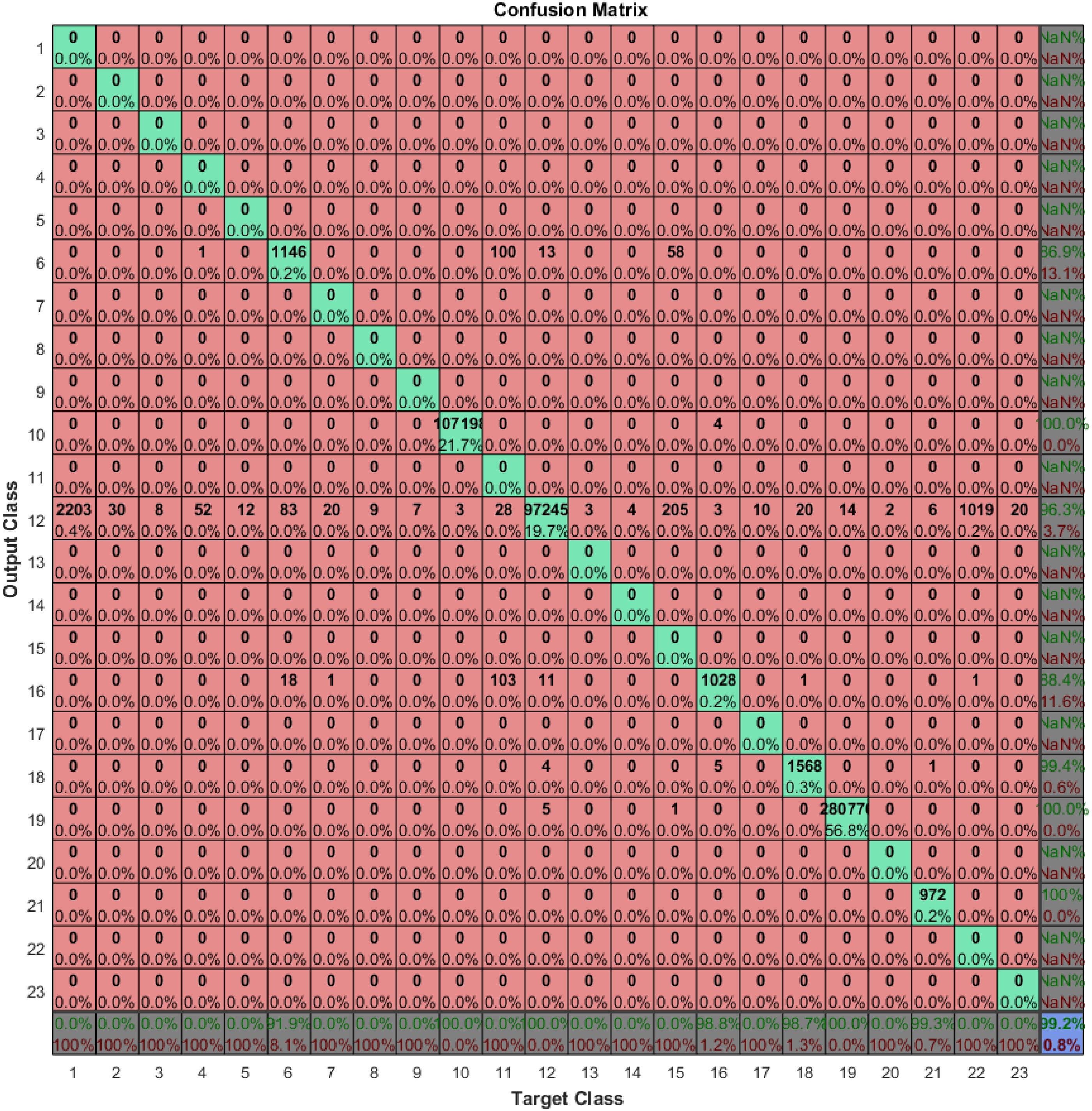
- Mapminmax, no PCA
- Activation function – radbas
- 5 Topologies
  - [35 35]; [20 20 20 20]; [20 20 20]; [15 15 15 15]; 20
- 5 training vs validation splits
  - 10/90, 50/50, 90/10, 30/70, 70/30
- maximum of 100 folds whilst performance continues to improve

Final Results Table



Hidden Layers/Neurons	*Train/Val Ratio*	*Performance*	*% Error*	*No. of Errors*	*No. of Network Connections*
35 × 35	30/70	0.000238848	0.007847845	3877	3395
20 × 20 × 20 × 20	30/70	0.000248808	0.007949055	3927	2440
20 × 20 × 20	30/70	0.00024774	0.008113015	4008	2040
20 × 20 × 20	70/30	0.000256067	0.008244589	4073	2040
20	10/90	0.000251437	0.008291146	4096	1240

Table 2: The best performing Neural Networks (top 5 rows)



Confusion Matrix

Output Class

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
2	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
3	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
4	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
6	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	1146 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 0.0%	13 0.0%	0 0.0%	0 0.0%	58 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	86.9% 13.1%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
9	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
10	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	107198 21.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	4 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0% 0.0%
11	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
12	2203 0.4%	30 0.0%	8 0.0%	52 0.0%	12 0.0%	83 0.0%	20 0.0%	9 0.0%	7 0.0%	3 0.0%	28 0.0%	97245 19.7%	3 0.0%	4 0.0%	205 0.0%	3 0.0%	10 0.0%	20 0.0%	14 0.0%	2 0.0%	6 0.0%	1019 0.2%	20 0.0%	96.3% 3.7%
13	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
14	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
15	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
16	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	18 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	103 0.0%	11 0.0%	0 0.0%	0 0.0%	0 0.0%	1028 0.2%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	88.4% 11.6%
17	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
18	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	4 0.0%	0 0.0%	0 0.0%	0 0.0%	5 0.0%	0 0.0%	1568 0.3%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	99.4% 0.6%
19	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	280770 56.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0% 0.0%
20	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
21	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	972 0.2%	0 0.0%	0 0.0%	100% 0.0%
22	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
23	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	0.0% 100%	0.0% 100%	0.0% 100%	0.0% 100%	0.0% 100%	91.9% 8.1%	0.0% 100%	0.0% 100%	0.0% 100%	100.0% 0.0%	0.0% 100%	100.0% 0.0%	0.0% 100%	0.0% 100%	0.0% 100%	98.8% 1.2%	0.0% 100%	98.7% 1.3%	100.0% 0.0%	0.0% 100%	99.3% 0.7%	0.0% 100%	0.0% 100%	99.2% 0.8%

Target Class

The chosen neural network

- 'patternnet‘ defaults
- 39×20×23 using radbas and softmax
- Trainscg - scaled conjugate gradient BP
- removeconstantrows, mapminmax
- training/validation ratio of 10/90
- 96.3% accuracy classifying normal connections
- 99% accuracy classifying bad connections.

Any questions?