Beauty Contest for a Line-up of Models

Edward Plowman & John Yick

Agenda

- Parade – meet the models
- Judging criteria
- Ball gown round
- Swimsuit round
- Crowning ceremony

Parade

- Selection of models from disciplines of:
  - Statistics
    - Tabular (one-way and multi-way) analysis
    - Linear model
    - Generalised linear model (GLM)
    - Generalised additive model (GAM)
  - Data mining
    - Decision tree
    - Neural network

- Models have been applied using a fairly mechanical process

Judging Criteria

- Tested on four real world insurance modelling problems:
  - A motor claim frequency analysis
  - A retention analysis for a personal lines portfolio
  - An average claim size analysis for CTP
  - A ‘return to work’ measure for workers’ compensation claims
- Data volumes are smallish but workable

Judging Criteria

- Models have been evaluated from two different perspectives
  - *Ball gown round* – measuring the 'elegance' of the model
  - *Swimsuit round* – measuring how well the model represents the underlying structure
- The first is necessarily a qualitative assessment
- The swimsuit round has been assessed using objective quantitative measures

Ball Gown Round: Criteria

- Ease of use
  - time, effort, expertise needed to set up model
- Output
  - form, interpretation, explanation, graphs/visual output, supporting statistics
  - portability, i.e. ease with which model can be implemented outside the analysis software
- Practicality
  - run-time, scalability, data volume requirements
- Structural
  - assumptions, extrapolation / trends, interactions and correlations

Ball Gown Round: 1-Way

- Easy and fast to carry out
- Easy to explain and interpret results
- Portable
- Small data volume requirements
- Only real downside is the risk of double-counting arising from correlations of exposure

Ball Gown Round: Multi-way

- Fairly easy to carry out
  - can be time-consuming if taking an exhaustive approach to factor combinations
- As dimensionality increases:
  - ease of interpretation rapidly diminishes
  - external implementation becomes more complex
  - data volume requirements increase rapidly
  - time required increases
- Can help to identify and deal with interaction effects between factors

Ball Gown Round: GLM

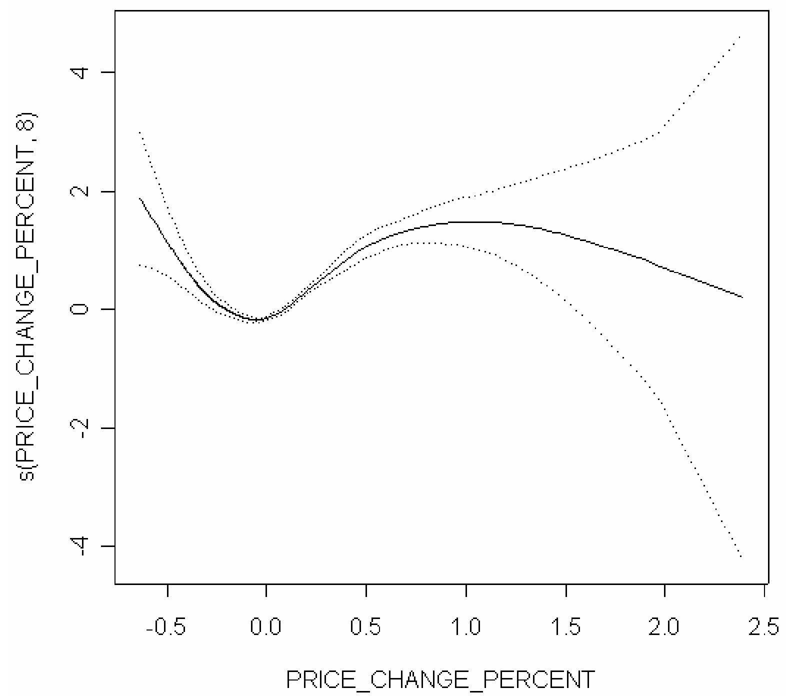
- Easy and fast with dedicated GLM software
- Needs some experience to interpret results effectively
- Easily portable
- Small data volume requirements
- Avoids risk of double-counting
- Makes structural assumptions

Ball Gown Round: GAM

- A generalisation of GLM
- Still allows 'link' functions, so the form of the model is not necessarily additive
- Essentially the same as GLM for categorical factors
- Allows non-linear functions ('scatterplot smoothers') to be fit to continuous factors
  - GLM: $\text{link}(Y) = X_1 + X_2 + \ldots + X_n$
  - GAM: $\text{link}(Y) = f(X_1) + f(X_2) + \ldots + f(X_n)$

Ball Gown Round: GAM

- Scatterplot smoother output example
  - Cubic spline with 8 degrees of freedom

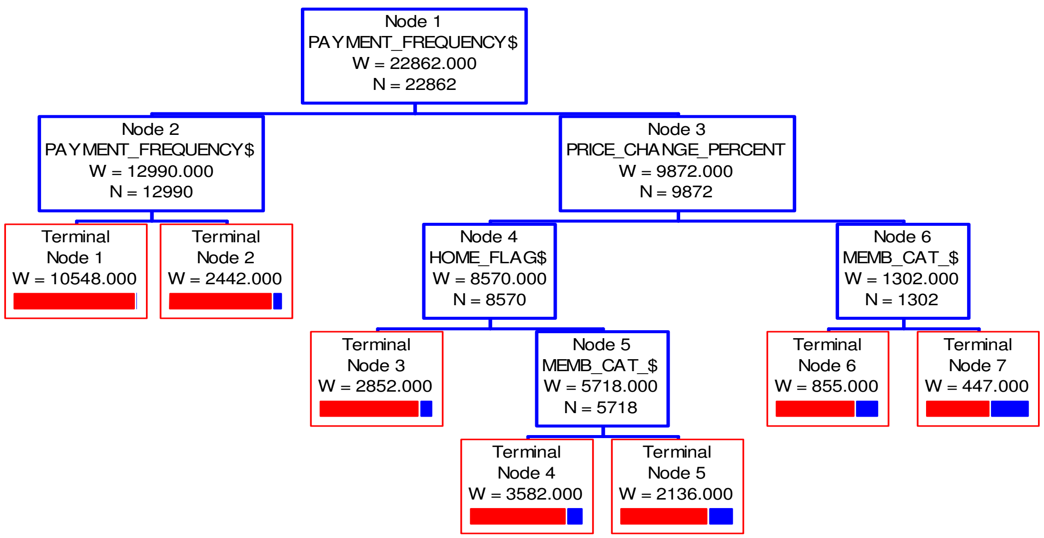


Ball Gown Round: GAM

- Output similar to GLM
- Results can be implemented externally but this is more complex than for GLM
- Data volume requirements as for GLM
- Only currently available in specialist statistical packages
  - not generally user-friendly
  - may not be suited to large datasets

Ball Gown Round: Decision Trees

- Several different algorithms for decision trees



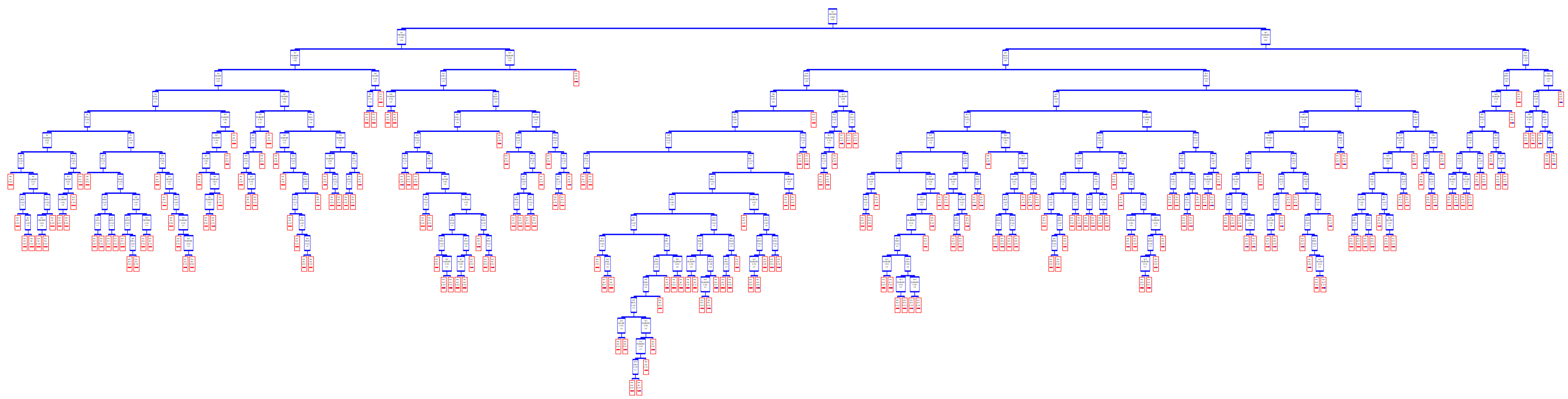
- We have used the CART algorithm

Ball Gown Round: CART

- Easy and fast with dedicated software
- Easily portable discrete segmentation
- Larger data volume requirements
- Visually appealing output
  - difficult to trace large trees

Institute of Actuaries of Australia

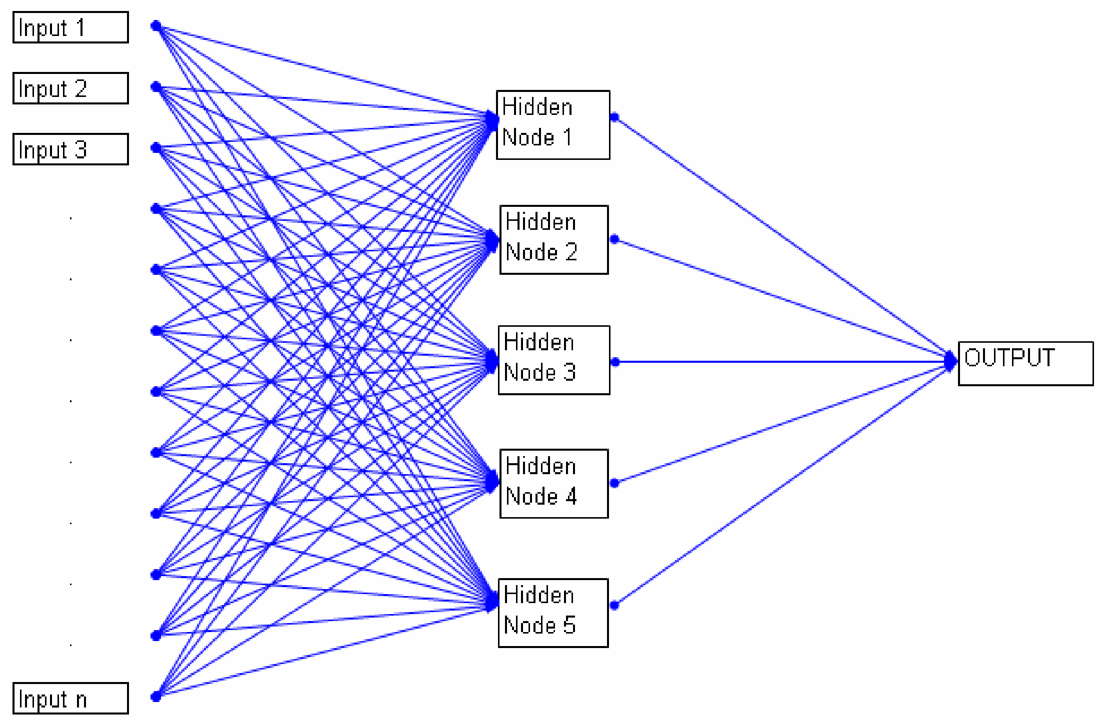
Ball Gown Round: CART



Ball Gown Round: Neural Nets

- There are many classes of neural net
- We have used a simple feed-forward network with a single hidden layer
- This consists of:
  - A set of inputs – i.e. the values of the explanatory factors
  - A ‘hidden’ layer of neurons – these generate sub-outputs using an activation function applied to a weighted sum of the inputs
  - A single output – i.e. the predicted value of the model, calculated as a linear weighted sum of the sub-outputs

Ball Gown Round: Neural Nets



- The neural net algorithms fit the weights linking inputs to neurons and neurons to the output
  - One for each blue arrow

Ball Gown Round: Neural Nets

- Easy to set up with dedicated software
  - Difficult to know an appropriate number of hidden neurons to specify
- Output largely meaningless apart from the prediction itself
- Complex to implement externally
- Unpredictable at extrapolating results (e.g. to allow for time trends)
- Long run-times even on small datasets, and converges on local optima (if at all), so may need to run several times
- Intermediate data volume requirements

Ball Gown Round: Results

- In traditional reverse order

Neural Net

Decision Tree

2-Way

GAM

1-Way

GLM

Swimsuit Round: Criteria

- Data split into:
  - Training
  - Testing (random sample and time)
- For each data subsets measure:
  - Area under gains chart
  - Misclassification matrix and kappa statistics
  - Sum of squared residuals

Swimsuit Round – Motor claims frequency

Swimsuit Round – Area Under Curve

Area Under Curve (AUC)

		Motor - Retention	Motor - Accident Claim Frequency	Workers - Return to work probability	CTP - Claim Severity
Predicting	One-way	72.5%	58.8%	-	-
	Two-way	72.4%	60.1%	-	-
	Simple Linear Model	72.9%	65.1%	-	-
	GLM	73.2%	65.3%	60.4%	56.2%
	GAM	73.4%	65.3%	60.5%	56.8%
	CART	71.3%	64.5%	59.3%	56.2%
	Neural Nets	73.2%	-	58.4%	56.3%
Training	One-way	73.2%	60.8%	-	-
	Two-way	73.6%	61.5%	-	-
	Simple Linear Model	73.9%	67.5%	-	-
	GLM	74.2%	67.5%	61.8%	57.5%
	GAM	74.4%	67.5%	62.0%	58.0%
	CART	72.6%	67.1%	59.4%	56.9%
	Neural Nets	74.6%	-	64.5%	58.9%

Swimsuit Round

- NN often gives best fit against training data, but is less good against testing data (i.e. tends to overfit)
- Other models exhibit similar performance on training and testing datasets
- GAM gives a better fit than GLM; improvements modest unless there are some key continuous factors
- Two-way offers only modest improvement over one-way, and is consistently worse than GLM
- Decision tree generally weak on these tests – low data volumes perhaps work against it

Swimsuit Round: Results

- In traditional reverse order

1-Way

2-Way

Decision Tree

Neural Net

GLM

GAM

Crowning Ceremony

- The ‘best’ model will of course depend on the application and the desired nature of the output
- However, generally the important thing is that the model is predictive
- We have therefore given double points for the swimsuit round in our assessment
- The final scores…

MODEL	Ball Gown	Swimsuit	Overall
1-way	4	0	Joint 5th
2-way	2	2	Joint 5th
GLM	5	8	Joint 1st
GAM	3	10	Joint 1st
CART	1	4	4th
NN	0	6	3rd