Tutorial 8

- Current best learning
- Decision Trees
- Naive Bayes Model

Current Best Learning

- Recognising true and false positives and negatives
  - False positive – hypothesis predicts positive but actually negative
  - False negative – hypothesis predicts negative but actually positive
  - True positive – hypothesis predicts positive and actually positive
  - True negative – hypothesis predicts negative and actually negative

Current Best Learning

- How to change the hypothesis when the example is:
  - False positive -> specialise
  - False negative -> generalise
  - True positive ->no change
  - True negative -> no change

Decision Trees

- Determine the ‘most important’ attribute using information theory
- Recommended steps:
  - State the positive and negative examples for each attribute
  - Calculate remainder using equation, showing all working (if you use the graph, state that you have done so)
  - Calculate gain using equation (may not be needed)
  - Determine ‘most important’ attributed given remainders or gains that have been calculated

Decision Trees

- Completing the tree
- In examples with small data sets, information theory should be used to determine the top attribute for the tree, but the remaining branches may be simple enough to determine by looking at the data and finding an attribute that classifies the remaining examples
- Nodes should be attributes
- Leaf nodes should be outcomes
- Branches should be labelled with the attribute value

Naive Bayes

- For each test pattern, determine the relative probabilities for each of the possible outcomes
  - State the equation being used
  - State the conditional probabilities being used
  - Calculate the relative probabilities
- If all you need is the result of the model, state the outcome predicted by the model (which outcome has a higher relative probability)
- If you also need the normalised probabilities, the relative probabilities should be adjusted so that they add up to 1

Naive Bayes

Calculating Conditional Probabilities

- P(Hungry = true | Will Wait = false) =
  P(Hungry = true and Will Wait = false) /
  P(Will wait = false) = (2/10) / (5/10) = 2/5

	Fri/Sat	Hungry	Patrons	Type	Will wait?
1	FALSE	TRUE	Some	French	TRUE
2	FALSE	TRUE	Full	Thai	FALSE
3	FALSE	FALSE	Some	Burger	TRUE
4	TRUE	TRUE	Full	Thai	TRUE
5	TRUE	FALSE	Full	French	FALSE
6	FALSE	TRUE	Some	Italian	TRUE
7	FALSE	FALSE	None	Burger	FALSE
8	FALSE	TRUE	Some	Thai	TRUE
9	TRUE	FALSE	Full	Burger	FALSE
10	TRUE	TRUE	Full	Italian	FALSE

Naive Bayes
Calculating Conditional Probabilities

- P(Hungry = true | Will Wait = false) = 2 / 5

	Fri/Sat	Hungry	Patrons	Type	Will wait?
2	FALSE	TRUE	Full	Thai	FALSE
5	TRUE	FALSE	Full	French	FALSE
7	FALSE	FALSE	None	Burger	FALSE
9	TRUE	FALSE	Full	Burger	FALSE
10	TRUE	TRUE	Full	Italian	FALSE