

## **DSC 433/533 – Information Analysis for Managerial Decisions – Study Guide**

The final exam will be in two parts: 15 multiple choice questions (each worth 10 points, totaling 150 points) and 10 short answer questions (worth a total of 200 points). We'll go through the practice exam (containing 5 multiple choice questions only) and preview some typical short answer questions in the last class of the quarter. This document is a 4-page study guide (including space to make notes) based on the short-answer section of the final exam. The exam itself is closed-book, closed-notes (other than the 4-page study guide).

### **Question topics on the short-answer section of the final exam**

1. Match the names of various data mining techniques (e.g., “association rules”) to short descriptions.
2. Make lift chart calculations for responses and net profits for a classification analysis with known costs of making an offer and expected revenue for a positive response. Scale-up the answers to calculate total net profits for a large-scale mailing based on the results.
3. Select appropriate words/phrases (from a list) to fill in blanks in a description of the steps involved in a classification analysis to determine potential customers.
4. Calculate response probabilities based on the results of a logistic regression analysis. Calculate the corresponding lift for responses and net profits using cost/benefit information. Scale-up the answers to calculate total net profits for a larger mailing based on the results.
5. Describe in words the decision rules resulting from a displayed classification tree. Briefly outline how the rules could be used. Draw a graphical representation of a classification tree using a set of rules described in words.
6. Label the parts of a neural network (shown graphically). Describe advantages and disadvantages of particular features of neural networks relative to other directed data mining techniques.
7. Use nearest neighbors results to estimate the probability of a positive response and to predict the amount of spending for a particular individual, for between 1 and 4 neighbors. Describe how to use such information to predict outcomes for other individuals.
8. Calculate confidence and lift for 3 association rules for a telephone service application. Describe how to use the results and also any limitations of such an analysis.
9. Consider why determining the number of clusters in a cluster analysis is critical; know how hierarchical cluster analysis and k-means cluster analysis segment data; and describe how cluster profiles can impact marketing.
10. Describe in detail the steps involved in a multiple linear regression analysis to predict lifetime value for customers at risk of churning.

**Space for notes**

**Space for notes**

**Space for notes**