School of Mathematics

463S (ST463) - Data Mining 2008-09
(SS Mathematics)

Lecturer: Dr. M. O'Regan

Requirements/prerequisites: ST370, permission of the Lecturer

Duration: 18 weeks

Number of lectures per week: 2 lectures and 1 laboratory hour per week.

Assessment: Two assessments which require student to carry out an analysis on a dataset and write a report. (40%).

End-of-year Examination: 3-hour end of year examination (counting remaining 60%). (Students required to answer a compulsory part consisting of 13 short questions and two out of three other questions.)

Description:
Aims To introduce students to a set of 'data mining' techniques enabling them to carry out analysis of data using these techniques. The course also stresses the comparison of these techniques to the classical statistical techniques described in ST370.
Learning Outcomes When students have successfully completed this module they should:

- Understand the theory underlying the topics given in the next section
- Construct models using these techniques and explain the results to a client
- Compare these methods with the methods covered in ST370

Syllabus: Specific topics addressed in this module include:

- Classification and regression trees
- Evaluation of Models
- Neural networks
- Overview of Support vector methods
- Association Rules
- Combining Classifiers

Bibliography

1. Ayres, I. Supercrunchers, How anything can be predicted, John Murray, 2007.
2. Berry M. J, A., & Linoff, G. Data Mining Techniques, John Wiley & sons, 1997
3. Bishop, Christopher, Pattern Recognition and Machine Learning, Springer Science, 2006.

4. Bishop, Christopher, Neural Networks for Pattern Recognition, Oxford: Clarendon Press, 1995

5. Bethold M. & Hand, D.J. Intelligent Data Analysis, Springer, 199.

6. Breiman, L., Friedman, J. H. Olshen, R. A. & Stone, C. J. Classification and regression Trees, Chapman and Hall,1984

7. Davenport, T.H. Harris, J.G. Competing on Analytics, The New Science of Winning, Harvard Business School Press, 2007.

8. Garson, G. D. Neural Networks An Introductory Guide for Social Scientists, Sage Publications, 1998

9. Giudici, P. Applied Data Mining : Statistical Methods for Business and Industry , Wiley, 2003.

10. Hand, D. Construction and Assessment of Classifiation Rules, Wiley, 1997.

11. Hand, D., Mannila, H. & Smyth P. Principles of Data Mining, MIT Press, 2001.

12. Hastie Trevor, Tibshirani, R., Friedman, J. The Elements of Statistical Learning, Springer Series, 2001

13. Haykin Simon , Neural Networks 2nd Edition, London, Prentice Hall, 1999

14. Ripley, B. D. Pattern recognition and Neural Networks, Cambridge University Press, 1996

15. Rud, Olivia Parr, Data Mining Cookbook, John Wiley & Sons , 2001

16. Tan, Pang-Ning Steinbach, M. Kumar, V. Introduction to Data Mining, Pearson, 2006

17. Thomas, Lyn, C., Edelman, D.B., & Crook, J. N. Credit Scoring and Its Applications, Monographs on Mathematical Modeling and Computation, SIAM 2002.

18. Webb, Andrew, Statistical Pattern Recognition 2nd Edition, Wiley, 2002.

Students are also encouraged to look for articles on the net.

December 11, 2008