620-639 Data Mining

**Lecturer:** Hugh Miller, room 198B, Richard Berry Building. Consulting times will be 2pm-3pm on Tuesdays. The first two weeks will be taught by Yao-ban Chan, room 198, Richard Berry Building.

**Webpage:** Please visit [www.ms.unimelb.edu.au/~millerh/](http://www.ms.unimelb.edu.au/~millerh/) for content related to this course.

**Assessment:** 80% final exam, 20% for two assignments. Lecture times: Tuesdays 3:15pm-5:15pm in Richard Berry 215 (lecture), and Fridays 4:15pm-5:15pm in Richard Berry 212 (computer laboratory).

**Prerequisites:** No formal prerequisites, but statistical ideas will be used and R will be used for computer laboratories. Thus having taken other statistics courses beforehand will be an advantage.

**Topics:** The course will cover a range of techniques in data mining. We shall start with unsupervised learning (association rules and clustering), and then move on to supervised methods, including linear and logistic regression, kernel regression, splines and generalised additive models, decision trees, boosting, neural nets and support vector machines (subject to time). The computer labs will follow the lectures in content.

**References:** The lecture slides and material covered during lectures will form the formal content, so no external reference is required. However, the following references (the first in particular) are recommended for those interested in pursuing any topics in greater depth.