ISM 209: Data Mining and Business Analytics
In Knowledge Services

Location
UCSC Silicon Valley Center at NASA (Mountain View): Room 2077
UCSC Main campus: Engineering 2 Room 194

Time
Lecture: Wednesdays 6:00—9:30 p.m (including mid-class break)
Boot Camp: Tentatively Mondays, 6-9 pm
TA sessions: Wednesdays 5:00-6:00 p.m.
Office Hours (Akella): Wednesdays 5:00-6:00 (by appointment)

Course Objectives
The intent of the course is to focus on several Knowledge Service areas in Business Management, describe the critical challenges and issues, and develop fundamental techniques and methods to solve these problems.

Knowledge Services comprises of two aspects or elements: 1. Analysis of data and information to produce knowledge which is provided to the user to assist in understanding, knowledge discovery, and alerts for decision making purposes, such as a potential failure of an electronic system, or an ad opportunity alert, and 2. Decision making to deliver actual physical or other services, such as in a call or service center.

Our focus in this course will be on the former aspects of knowledge computation and alert; the decision making for achieving service delivery will be considered in subsequent courses. We will specifically consider:

- Application areas: Online marketing and computational advertising, support services for operations and innovation, financial services, entertainment, and healthcare
- Specific topics: Recommenders, fraud detection, anomaly detection, online marketing and purchase probabilities, social networks, support services and centers, healthcare prognostics and diagnostics, financial prediction, text mining, information retrieval, incorporation of human user feedback
- Techniques: Machine learning, Prediction - Linear Regression, logistic regression, constrained optimization, recommenders, relational learning, time series, discriminant analysis and classification, text mining and information retrieval
- Software (Data Mining and Statistical) Tools: S and Enterprise Miner (SAS) and or R (Open Source), Matlab, (Plus Weka – open source – for Machine learning, and XLMiner, an Excel add-on for data mining),

This course is introductory and will develop the fundamental statistical and machine learning models and techniques for data mining and business analytics progressively, in the context of real world applications and examples. We will develop the techniques systematically and sequentially, but will move back-and-forth the different real world
applications and domain areas. Industry speakers and class presentations will provide a broader exposure to topics and leading-edge research and industry practice results. The subsequent courses, such as ISM 250 and 251, will expand on the techniques and domain areas such as web mining, computational advertising and online marketing, social networks and relational learning, reinforcement learning, constrained optimization, and also explore significant projects, including possibly with industry.

Prerequisites: Students are expected to be mathematically mature, and to have had prior exposure to undergraduate linear algebra at the level of MATH21 or AMS10 and probability/statistics at the level of AMS 131 or MPE 107. We will provide a refresher in the form of a “bootcamp” early in the course, to enable students to relearn basics required for the course.

Grading
There will be two quizzes on Oct 29nd and December 3rd. Each will count towards 10% of the grade. The midterm exam will be on Nov 12th and will be 25% of the grade. There will be assignments with two parts: one will be analytical, and the other will be more implementation-oriented.

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Quizzes and midterm</td>
<td>45%</td>
</tr>
<tr>
<td>Miniprojects</td>
<td>35%</td>
</tr>
</tbody>
</table>

Required Text Book

Other Reference Books: