Texas A&M University
Departmental Request for a New Course
Undergraduate + Graduate + Professional
• Submit original form and attach a course syllabus.

1. This request is submitted by the Department of **STATISTICS**

2. Course prefix, number and complete title of course: **STAT 618 - Statistical Aspects of Machine Learning and Data Mining**

3. Course description (not more than 50 words): This course will examine the statistical aspects of techniques used to examine data streams which are large scale, dynamic, and heterogeneous. This course will examine the underlying statistical properties of classification; trees; bagging and boosting methods; neural networks; support vector machines; cluster analysis; and independent component analysis;

4. Prerequisite(s): **STAT 610, STAT 611, STAT 613**

5. Is this a variable credit course? □ Yes  ☑ No  If yes, from _______ to _______.

6. Is this a repeatable course? □ Yes  ☑ No  If yes, this course may be taken ______ times. Will the course be repeated within the same semester/term? □ Yes  ☑ No

7. Has this course been taught as a 289/489/689?  ☑ Yes  □ No  If yes, how many times? 2  Indicate the number of students enrolled for each academic period it was taught. Fall 2006 - 13 students, Spring 2005 - 16 students

8. This course will be:
   a. required for students enrolled in the following degree program(s) (e.g., B.A. in history)
   b. an elective for students enrolled in the following degree program(s) (e.g., M.S., Ph.D. in geography)

9. If other departments are teaching or are responsible for related subject matter, the course must be coordinated with these departments. **Attach approval letters.**

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Course #</th>
<th>Title (excluding punctuation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT 618</td>
<td>STAT MACHINE LEARN &amp; MINING</td>
<td></td>
</tr>
<tr>
<td>Lect.</td>
<td>Lab</td>
<td>SCH</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

Approval recommended by:  
Head of Department  
Date  
Chair, College Review Committee  
Date  
Dean of College  
Date  
Director of Academic Support Services  
Date  
Effective Date

Questions regarding this form should be directed to Sandra Williams at 845-8836.

OAR/AS – 04/07
STATISTICS 618
Statistical Aspects of Machine Learning and Data Mining - FALL 2009

• Time and Place: TR 2:20 p.m. - 3:35 p.m., BLOCKER 411
• Instructor: Dr. Marc Genton
• Office: 502B Blocker Building
• Phone: 848-0889 (Office) Email: genton@stat.tamu.edu
• Web page: http://www.stat.tamu.edu/genton/STAT618/STAT618.html
• Office Hours: TBA
• Prerequisites: STAT 610, 611, 613
• Text: Modern Multivariate Statistical Methods, by Alan Izenman. Springer

• Course Outline:
  A. Supervised learning:
     1. linear methods for regression/classification
     2. additive models and trees (GAM, CART, PRIM, MARS, MART)
     3. bagging and boosting methods
     4. neural networks
     5. support vector machines
  B. Unsupervised learning:
     1. association rules
     2. market basket analysis
     3. cluster analysis
     4. independent component analysis

• Grade Determination: Your grade for the course will be based on (relative weights given in %):
  • Homework assignments (50%): Bi-weekly homework assignments will be given throughout the semester. Assignments will be collected at the START of class on the due date. Late assignments will not be accepted unless a university approved excuse is provided.
  • Project (50 %): A project, done individually, will be due near the end of the semester, and presented in class (20 minutes). More details will be given as the semester progresses. Late projects will not be accepted unless a university approved excuse is provided.
• The following letter grades and percentages will be applied:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100-90.0%</td>
<td>89.9-80.0%</td>
<td>79.9-70.0%</td>
<td>69.9-60.0%</td>
<td>59.9-0%</td>
</tr>
</tbody>
</table>

• Notes:
  
  • Check website frequently for updates!
  
  • I would like to know about any particular academic difficulties or personal problems that are affecting a student's performance.

STATEMENT ON DISABILITIES: The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation for their disabilities. If you believe you have a disability requiring an accommodation, please contact the Office of Support Services for Students with Disabilities in Room B118, Cain Hall. The phone number is 845-1637. For additional information visit http://disability.tamu.edu

STATEMENT ON PLAGIARISM: The handouts used in this course are copyrighted. By "handouts," I mean all materials generated for this class, which include but are not limited to syllabi, quizzes, exams, lab problems, in-class materials, review sheets, and additional problem sets. Because these materials are copyrighted, you do not have the right to copy the handouts, unless I expressly grant permission. As commonly defined, plagiarism consists of passing off as one's own ideas, words, writing, etc., which belong to another. In accordance with this definition, you are committing plagiarism if you copy the work of another person and turn it in as your own, even if you should have the permission of that person. Plagiarism is one of the worst academic sins, for the plagiarist destroys the trust among colleagues without which research cannot be safely communicated. If you have any questions regarding plagiarism, please consult the latest issue of the Texas A&M University Student Rules, under the section "Scholastic Dishonesty." Also, more information is available at the website: www.tamu.edu/aggiehonor.
MEMORANDUM:

TO: Simon J. Sheather  
Professor and Department Head  
Department of Statistics

FROM: Valerie E. Taylor  
Royce E. Wisenbaker Professor and Department Head  
Department of Computer Science and Engineering

DATE: July 2, 2009

SUBJECT: Statistical Aspects of Machine Learning and Data Mining Course

The Department of Computer Science and Engineering has reviewed the materials sent forth for course STAT 618 – Statistical Aspects of Machine Learning and Data Mining and supports the course as described.

Should you have any questions, please feel free to contact me.