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 619 - Statistical Machine Learning and Data Mining**

3. Course description (not more than 50 words): Techniques to deal with problems in science and technology where data streams are large scale, dynamic, and heterogeneous; classification; trees; bagging and boosting methods; neural networks; support vector machines; cluster analysis; and independent component analysis.

4. Prerequisite(s): **STAT 613**  
   Cross-listed with: **None**

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?  
   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, M.S. and Ph.D. in Statistics)

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

10. Prefix | Course # | Title (excluding punctuation)  
         | STAT 619 | STAT MACH LEARN & MINING

         | Lect | Lab | SCH | Subject Matter Content Code | Admin. Unit | Acad. Year | FICE Code |
         | 0 | 3 | 0 | 0 | 3 | 2 | 7 | 0 | 5 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |

Approval recommended by: **Michael Jordan**  
Head of Department  
Date: **2/19/09**

Chair, College Review Committee  
Date: **2/27/09**

Head of Department (if cross-listed course)  
Date: **2/27/09**

Dean of College  
Date: **2/27/09**

Submitted to Coordinating Board by:  
Date: **2/27/09**

Director of Academic Support Services  
Date: **2/27/09**

Effective Date: **2/27/09**

Questions regarding this form should be directed to Sandra Williams at 845.8836.  
OAR/AS - 04/07
Texas A&M University

Departmental Request for a New Course

- 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 613, Statistical Machine Learning and Data Mining.

3. Course description (not more than 50 words): This course will cover techniques to deal with problems in science and technology where data streams are large-scale, dynamic, and heterogeneous. Topics covered will include classification, trees, bagging, and boosting methods, neural networks, support vector machines, cluster analysis, and independent component analysis.

4. Prerequisite(s): STAT 613

5. Is this a variable credit course? ☑ Yes ☐ No

6. Is this a repeatable course? ☑ Yes ☐ No

7. Has this course been taught as a 289/489/689? ☑ Yes ☐ No

8. This course will be:
   a. required for students enrolled in the following degree program(s) (e.g., B.A. in history)

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

10. Prefix | Course # | Title (excluding punctuation) |
      | STAT 613 | STATISTICS | MACHINE LEARNING & MINING |

<table>
<thead>
<tr>
<th>Lect.</th>
<th>Lab</th>
<th>SCH</th>
<th>Subject Matter Content Code</th>
<th>Admin. Unit</th>
<th>Acad. Year</th>
<th>FICE Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>32705010001</td>
<td>18-11</td>
<td>003632</td>
<td>Level 0</td>
</tr>
</tbody>
</table>

Approval recommended by:

Head of Department: Michael Jorgensen, 2-19-2009

Chair, College Review Committee: Date 2/27/09

Dean of College: Date 2/27/09

Submitted to Coordinating Board by:

Date

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

OAR/AS - 04/07
STATISTICS 619 - Statistical Machine Learning and Data Mining

- **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
- **Office Hours:** TR 3:45-5:00 PM
- **Prerequisite:** STAT 613
- **Web page:** http://www.stat.tamu.edu/~genton/STAT617/STAT617.html
- **Text:** The Elements of Statistical Learning: Data Mining, Inference, and Prediction, by Hastie, Tibshirani, Friedman. Springer

- **Course Outline:**
  - Supervised learning: linear methods for regression/classification; additive models and trees (CART, PRIM, MARS, MART); bagging and boosting methods; neural networks; support vector machines
  - Unsupervised learning: association rules; market basket analysis; cluster analysis; 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.
  - 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.
  - The following letter grades and percentages will be applied:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100-90.0%</td>
</tr>
<tr>
<td>B</td>
<td>89.9-80.0%</td>
</tr>
<tr>
<td>C</td>
<td>79.9-70.0%</td>
</tr>
<tr>
<td>D</td>
<td>69.9-60.0%</td>
</tr>
<tr>
<td>F</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 126 of the Koldus Student Services Building. The phone number is 845-1637.

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."