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 656 - Applied Analytics Using SAS Enterprise Miner

3. Course description (not more than 50 words): Introduction to data mining and will demonstrate the procedures, optimal prediction decisions; comparing and deploying predictive models; neural networks; constructing and adjusting tree models; the construction and evaluation of multi-stage models.

4. Prerequisite(s) STAT 657 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. Spring 2009 - 25 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 these departments. Attach approval letters.

10. Prefix Course # Title (excluding punctuation) STAT 656 APPLIED ANALYTICS

<table>
<thead>
<tr>
<th>Lect</th>
<th>Lab</th>
<th>SCH</th>
<th>Subject Matter Code</th>
<th>Content Code</th>
<th>Admin. Unit</th>
<th>Acad. Year</th>
<th>FICE Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>00</td>
<td>327050100012740</td>
<td>10 - 11</td>
<td></td>
<td></td>
<td>003632</td>
</tr>
</tbody>
</table>

Level

Approval recommended by:

Michael Long
Head of Department 2-19-2009

Chair, College Review Committee 2/7/09

Dean of College 2/27/09

Head of Department (if cross-listed course) Date

Submitted to Coordinating Board by:

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
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 656 - Applied Analytics Using SAS Enterprise Miner**

3. Course description (not more than 50 words): This course is an introduction to data mining and will demonstrate the procedures on data sets. Topics covered will include optimal prediction decisions; comparing and deploying predictive models; neural networks; constructing and adjusting tree models; the construction and evaluation of multi-stage models.

4. Prerequisite(s) **STAT 657**

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? 1 Indicate the number of students enrolled for each academic period it was taught. **Spring 2009 - 25 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 these departments. **Attach approval letters.**

10. Prefix | Course # | Title (excluding punctuation) | Lect. | Lab | SCH | Subject Matter Content Code | Admin. Unit | Acad. Year | FICE Code | Level |
      ---------|---------|--------------------------------|-------|-----|-----|-----------------------------|------------|-----------|-----------|-------|
      STAT   | 656     | APPLIED ANALYTICS              | 0     | 3   | 0   | 032705010001274008 - 09    | 003632     |

Approval recommended by:

**Michael Fargnold** 2-19-2009
Head of Department Date

Chair, College Review Committee 2-27-09

Dean of College 4-21-09

Submitted to Coordinating Board by:

Date

Effective Date

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

OAR/AS – 04/07
Stat 656

Applied Analytics Using SAS Enterprise Miner

This course covers the knowledge and skills required to assemble analysis flow diagrams using the rich tool set of SAS Enterprise Miner for both pattern discovery (segmentation, association, and sequence analyses) and predictive modeling (decision tree, regression, and neural network models).

Prerequisites Stat 657

Grading:

<table>
<thead>
<tr>
<th>Component</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>10%</td>
</tr>
<tr>
<td>Exam 1</td>
<td>20%</td>
</tr>
<tr>
<td>Exam 2</td>
<td>30%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>40%</td>
</tr>
</tbody>
</table>

Course Contents

Introduction

- touring SAS Enterprise Miner 5.2
- placing SAS Enterprise Miner in the analysis workflow
- application examples and case studies
- pattern discovery and predictive modeling

Accessing and Assaying Prepared Data

- SAS Enterprise Miner system architecture
- defining a SAS Enterprise Miner 5.2 project
- defining a data source
- validating source data

Introduction to Pattern Discovery

- clustering and segmenting data
- using the Transform Variables node
- using the Clustering node for k-means cluster analysis
- applying association and sequence discovery
- using the Associations node in a consumer banking example
- quantifying the associations among items
- exploring sequences among items
Introduction to Predictive Modeling with Decision Trees

- defining a modeling data source
- partitioning data for model development
- growing a decision tree with the Desktop Tree Application
- running the Decision Tree node
- using Decision Tree node options
- understanding predictive modeling results data

Introduction to Predictive Modeling with Regressions

- comparing linear and logistic regression
- using the Regression node
- imputing missing values with the Impute node
- replacing data values with the Replacement node
- performing input selection
- understanding regression modeling output
- extending regression models with polynomial and interaction terms

Introduction to Predictive Modeling with Neural Networks and Other Modeling Tools

- introduction to neural network (multilayer perceptron) models
- using the Neural Network node
- performing model selection with the Auto Neural node
- other SAS Enterprise Miner 5.2 modeling tools

Model Assessment

- defining a prior vector
- effects of prior vectors on model development
- changing model selection criteria
- defining a profit matrix
- comparing models with the Model Comparison node
- introduction to model assessment statistics

Model Implementation

- defining a Score data set
- scoring a data set with the Score node
- using a SAS Code node to export scored data
- generating and using score code

Special Topics
• using the Variable Selection node
• combining models with the Ensemble node
• consolidating categorical inputs with the Decision Tree node
• explaining complex models with the Decision Tree node

Case Studies

• segmenting bank customer transaction histories
• association analysis on Web services data
• creating a simple risk model from consumer loan data
• predicting university enrollment
• creating a response model for an insurance product

Software Addressed
This course addresses the following software product(s): SAS Enterprise Miner.

Know the Code:
"An Aggie does not lie, cheat or steal, or tolerate those who do."

Information about the Aggie honor code can be found: http://www.tamu.edu/aggiehonor

Policies and Rules:

As a student, you are governed by a set of policies and rules. These can be found: http://student-rules.tamu.edu/

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, in-class materials, review sheets, and additional materials. Because these materials are copyrighted, you do not have the right to copy them unless I expressly grant permission.

As commonly defined, plagiarism consists of passing off as one’s own ideas, words, writings, 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 current issue of the Texas A&M Student Rules, under the section, "Scholastic Dishonesty."

Americans with Disabilities Act (ADA) Policy Statement:
The Americans with Disabilities Act (ADA) is a federal antidiscrimination 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 of their disabilities. If you believe you have a disability requiring an accommodation, please contact Disability Services in Cain Hall, Rm. B118, or call 845-1637.