New Course Request

Indiana University

Check Appropriate Boxes: Undergraduate credit □ Graduate credit □ Professional credit □

1. School/Division: Science, Biostatistics
2. Academic Subject Code: BIOS

3. Course Number: S612 (must be cleared with University Enrollment Services)
4. Instructor: Hayezlak

5. Course Title: Modern Statistical Learning Methods

Recommended Abbreviation (Optional): (Limited to 32 Characters including spaces)

6. First time this course is to be offered (Semester/Year): Fall 2011

7. Credit Hours: Fixed at 3 or Variable from _________ to _________

8. Is this course to be graded S-F (only)? Yes ___ No ___ X

9. Is variable title approval being requested? Yes ___ No ___ X

10. Course description (not to exceed 50 words) for Bulletin publication: P: STAT 52500

   This course will cover the topics pertaining to the modern methods of high-dimensional data analysis.

11. Lecture Contact Hours: Fixed at 3 or Variable from _________ to _________

12. Non-Lecture Contact Hours: Fixed at N/A or Variable from _________ to _________

13. Estimated enrollment: 15 of which 100 percent are expected to be graduate students.

14. Frequency of scheduling: annually Will this course be required for majors? yes

15. Justification for new course: Course for new Biostatistics program

16. Are the necessary reading materials currently available in the appropriate library? yes

17. Please append a complete outline of the proposed course, and indicate instructor (if known), textbooks, and other materials.

18. If this course overlaps with existing courses, please explain with which courses it overlaps and whether this overlap is necessary, desirable, or unimportant.

19. A copy of every new course proposal must be submitted to departments, schools, or divisions in which there may be overlap of the new course with existing courses or areas of strong concern, with instructions that they send comments directly to the originating Curriculum Committee. Please append a list of departments, schools, or divisions thus consulted.

Submitted by: ___________________________ Date: 10-5-2009

Department Chairman/Division Director

Approved by: ___________________________ Date: 10-30-2009

Dean

Dean of Graduate School (when required)

Chancellor/Vice-President

University Enrollment Services

After School/Division approval, forward the last copy (without attachments) to University Enrollment Services for initial processing, and the remaining four copies and attachments to the Campus Chancellor or Vice-President.

UPS 724

University Enrollment Services Final—White; Chancellor/Vice-President—Blue; School/Division—Yellow; Department/Division—Pink; University Enrollment Services Advance—White
BIOS 598: Modern Statistical Learning Methods
SPRING 2009

Instructor:
Jaroslaw Harezlak, PhD, Assistant Professor

Contact information:
Jaroslaw Harezlak, PhD
Division of Biostatistics
Indiana University School of Medicine
410 West 10th Street, Suite 3000
Indianapolis, IN 46202

Office: (317) 274-2682
Email: harezlak@iupui.edu

Prerequisites: STAT 525 and familiarity with regression analysis.

Course Description:
Recent years have brought rapid growth in the amount and complexity of data in biostatistical applications. Among others, data collected in the genomic, proteomic and other “-omic” areas call for new statistical techniques in both predictive and descriptive learning. This course will cover the topics pertaining to the modern methods of high-dimensional data analysis. The specific topics will include penalized regression methods, linear classification methods, basis expansions and regularization, kernel methods, model assessment and selection, model inference and averaging, boosting and additive trees, neural networks, support vector machines, nearest-neighbor methods, and unsupervised clustering. Class presentations and homework assignments will cover both the methodological issues and the applications of the methods to data analysis.

Educational objectives:
At the end of the course students will be able to:
- use modern statistical methods for the analysis of high dimensional data
- utilize statistical software R to conduct data analysis of such data

Meeting Times: Lectures twice per week
# COURSE OUTLINE

<table>
<thead>
<tr>
<th>Lectures 1-3.</th>
<th>Introduction; Supervised Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lectures 4-5</td>
<td>Linear Methods for Regression</td>
</tr>
<tr>
<td>Lecture 6.</td>
<td>Linear Methods for Classification</td>
</tr>
<tr>
<td>Lectures 7-8</td>
<td>Nonparametric Methods</td>
</tr>
<tr>
<td>Lectures 9-10</td>
<td>Basis Expansions and Regularization</td>
</tr>
<tr>
<td>Lectures 11-12</td>
<td>Kernel Methods</td>
</tr>
<tr>
<td>Lectures 13-14</td>
<td>Model Assessment and Selection</td>
</tr>
<tr>
<td>Lectures 15-16</td>
<td>Model Inference and Averaging</td>
</tr>
<tr>
<td>Lectures 17-19</td>
<td>Additive Models, Trees, and Related Methods</td>
</tr>
<tr>
<td>Lectures 20-21</td>
<td>Boosting and Additive Trees</td>
</tr>
<tr>
<td>Lectures 22-23</td>
<td>Neural Networks</td>
</tr>
<tr>
<td>Lectures 24-26</td>
<td>Support Vector Machines &amp; Discriminant Analysis</td>
</tr>
<tr>
<td>Lecture 27</td>
<td>Nearest-Neighbor methods</td>
</tr>
<tr>
<td>Lectures 28-29</td>
<td>Unsupervised Learning</td>
</tr>
<tr>
<td>Lectures 30-31</td>
<td>Student presentations</td>
</tr>
</tbody>
</table>

**Required and Recommended Texts:**
The course will use the following textbook:

Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning,
The lectures will follow the textbook closely, but we will supplement the textbook with instructors' own lecture notes.

**Evaluation and Grading:** Students will be evaluated based on their performance on the homework assignments (50%), and the final project (50%). Letter grades for the course are assigned using the following scale:

A: 90-100; B: 80-89; C: 70-79; D: 60-69; F: less than 60. Within each letter grade, "+" and "-" will be assigned if the numeric score is in the top and bottom quintiles, respectively.

**Cheating and Plagiarism:**
Academic misconduct will not be tolerated and all cases will be reported. Examine the IU Code of Student Rights, Responsibilities, and Conduct at http://life.iupui.edu/help/code.asp and in particular examine the rules regarding academic misconduct at http://life.iupui.edu/help/docs/Part_3all.html. Violations of these rules will result in a grade of "F" (or 0%) for the assignment in question, and may result in an "F" for the course or even expulsion from the university (see http://life.iupui.edu/help/docs/Part_4all.html#sanction).

**Americans with Disabilities Act:**
If you need any special accommodations due to a disability, please contact Adaptive Educational Services at (317)-274-3241. The office is located in CA 001E.