Course Syllabus: Machine Learning - CS 229

<table>
<thead>
<tr>
<th>Division</th>
<th>Computer, Electrical and Mathematical Sciences &amp; Engineering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course Number</td>
<td>CS 229</td>
</tr>
<tr>
<td>Course Title</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>Academic Semester</td>
<td>Spring</td>
</tr>
<tr>
<td>Academic Year</td>
<td>2019/2020</td>
</tr>
<tr>
<td>Semester Start Date</td>
<td>01/26/2020</td>
</tr>
<tr>
<td>Semester End Date</td>
<td>05/13/2020</td>
</tr>
<tr>
<td>Class Schedule</td>
<td>10:30 AM - 12:00 PM</td>
</tr>
</tbody>
</table>

**Instructor(s)**

<table>
<thead>
<tr>
<th>Name</th>
<th>Email</th>
<th>Phone</th>
<th>Office Location</th>
<th>Office Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xiangliang Zhang</td>
<td><a href="mailto:Xiangliang.Zhang@kaust.edu.sa">Xiangliang.Zhang@kaust.edu.sa</a></td>
<td>+966128080313</td>
<td>4413, 1, Al-Khawarizmi (bldg. 1)</td>
<td></td>
</tr>
</tbody>
</table>

**Teaching Assistant(s)**

<table>
<thead>
<tr>
<th>Name</th>
<th>Email</th>
</tr>
</thead>
</table>

**Course Information**

**Comprehensive Course Description**

Machine Learning is a science of getting machines to learn, more specifically, designing algorithms that allow computers to learn from empirical data. In the past decade, Machine Learning has successfully made computers to recognize speeches and hand-written characters, to convert spoken words to text, to effectively search our needed information, and to recommend products/books/movies that we potently like. In this class, you will learn the most important machine learning techniques, not only the theoretical foundations of these techniques, but also the practice implementation of them. The main topics will include linear and non-linear regression, nonparametric methods, Bayesian methods, support vector machines, kernel methods, Artificial Neural Networks, deep network, model selection, learning theory, VC dimension, clustering, EM, dimensionality reduction, PCA, SVD, and reinforcement learning.

**Course Description from Program Guide**

Topics: linear and non-linear regression, nonparametric methods, Bayesian methods, support vector machines, kernel methods, Artificial Neural Networks, model selection, learning theory, VC dimension, clustering, EM, dimensionality reduction, PCA, SVD, and reinforcement learning.

**Goals and Objectives**

Students will understand the most important machine learning techniques, and be able to implement and apply these techniques on solving real problems.

**Required Knowledge**

Students should know linear algebra and basic probability and statistics. Familiarity with artificial intelligence recommended. Students should also know programming, at least be good at one programming skill.

**Reference Texts**

Pattern Recognition and Machine Learning, Chris Bishop, 2006. The lectures may not be compatible with the textbook. Reading literatures coupling with the course content will be post on-line, as well as the slides.

**Method of evaluation**

15.00% - Midterm exam
50.00% - Homework /Assignments
15.00% - Final exam
20.00% - Course Project(s)
### Nature of the Assignments

The homework will be assigned after the conclusion of each topic (see the homework assignment and due date at the course webpage). If submitting the homework after the due date, the homework will be graded for correctness, but not credited. The homework will consist of both conceptual questions and practical exercises, which require programming works in C/C++, or Matlab, or Java, etc. Each homework has 100 pts. The final homework score will be the average taking on all assigned homework.

The project of CS 229 should be an implementation of what you learned from the course. It can be a work:

1. designing a novel algorithm for supervised or unsupervised learning;
2. extending, improving, or speeding-up some existing algorithms;
3. comparing and discussing a bunch of machine learning algorithms; etc.

The project will be evaluated by:
Technical quality (30) + significance (30) + novelty/impact (20) + report/presentation (20)?

### Course Policies

Only one homework deadline extension is allowed for each student. Students should notify the instructor in advance of missing any class or as soon as possible thereafter. Any absence without notification will be reported to Graduate Affairs.

### Additional Information

The instructor reserves the right to make changes to this syllabus as necessary.

### Tentative Course Schedule

*(Time, topic emphasis & resources)*

<table>
<thead>
<tr>
<th>Week</th>
<th>Lectures</th>
<th>Topic</th>
</tr>
</thead>
</table>
| 1    | Mon 01/27/2020  
      Wed 01/29/2020 | Regression        |
| 2    | Mon 02/03/2020  
      Wed 02/05/2020 | Regression        |
| 3    | Mon 02/10/2020  
      Wed 02/12/2020 | Regression        |
| 4    | Mon 02/17/2020  
      Wed 02/19/2020 | Discriminative Algorithms |
| 5    | Mon 02/24/2020  
      Wed 02/26/2020 | Bayesian classifier |
| 6    | Mon 03/02/2020  
      Wed 03/04/2020 | Decision Tree Learning |
| 7    | Mon 03/09/2020  
      Wed 03/11/2020 | Neural Networks |
| 8    | Mon 03/16/2020  
      Wed 03/18/2020 | Deep Learning |
| 9    | Mon 03/23/2020  
      Wed 03/25/2020 | Support Vector Machine |
| 10   | Mon 03/30/2020  
      Wed 04/01/2020 | SVM and VC-Dimension |
| 11   | Mon 04/06/2020  
      Wed 04/08/2020 | K-means and EM |
| 12   | Mon 04/13/2020  
      Wed 04/15/2020 | Clustering  |
| 13   | Mon 04/20/2020  
      Wed 04/22/2020 | PCA Learning |
| 14   | Mon 04/27/2020  
      Wed 04/29/2020 | SVD |
| 15   | Mon 05/04/2020  
      Wed 05/06/2020 | Reinforcement Learning |
| 16   | Mon 05/11/2020  
      Wed 05/13/2020 | Project presentation |