**ITM 6273-01: Big Data Technology and Applications**

|  |  |
| --- | --- |
| Instructor: Jiming Wu | E-mail address: jiming.wu@csueastbay.edu |
| Office: VBT 355 | Office Hours: Mo & Th 4:00-5:30PM |
| Class Time: Thu. 6:30-10:00PM | Classroom: Art & Education 142A |
| Course Units: 4 | Phone: (510)885-3099 |
| Class Website: http://bb.csueastbay.edu/webapps/portal/frameset.jsp | |

**TEXTBOOK:**

White, T. ***Hadoop: The Definitive Guide***, 3rd edition, O’Reilly.

**SOFTWARE:**

Hadoop, Amazon MapReduce, and a PC

**COURSE DESCRITION:**

This is an introduction to Big Data concepts, technologies, systems, and applications. Topics include MapReduce features, Hadoop distributed filesystem, input/output, application development, Pig, and Hive. Students will have opportunity to use Amazon MapReduce and Hadoop system installed on university server. The objective of this course is to impart working knowledge and skills associated with Big Data technologies and to let students better understand how companies leverage these technologies to analyze Big Data.

**OBJECTIVES AND LEARNING OUTCOMES:**

Upon successful completion of the course a student, at a level appropriate to graduate study, will be able to:

* Analyze technical components of big data technologies such as distributed computing.
* Apply advanced applications and technologies for effective data processing and analysis.
* Demonstrate the knowledge and skills for manipulating, storing, and analyzing Big Data in a project setting.

**PREREQUISITES:**

ITM 6015 & ITM 6271

**GRADING:**

20%-Project; 20%-Homework; 30%-Exam 1; 30%-Exam 2.

Grades: A = 93-100; A- = 90-92; B+ = 87-89; B = 83-86; B- = 80-82; C+ = 77-79; C = 73-76; C- = 70-72; D+ = 65-69; D = 60-64; F = 0-59.

**Project**

One project will be assigned during the quarter. The task is to modify existing Java programming codes to solve a problem related to the examples in the textbook. The project will be posted to Blackboard and you should submit your project to Blackboard before the due dates in the schedule. Late submission gets partial credit but late submission received two days after its original due day will **not earn** credit.

**Homework**

Three homework tasks will be assigned during the quarter. The tasks are mainly to work the examples in the textbook. The assignments will be posted to Blackboard and you should submit your completed work to Blackboard before the due dates in the schedule. Late homework gets partial credit but late homework received two days after its original due day will **not earn** credit.

**Exam**

Two exams will be given on the dates in the schedule. Make-ups will not be given unless absences are excused with formal documentation according to university policy.

**ATTENDANCE:**

Students are required to attend all the classes and be on time. If absence occurs, the student is responsible for making up missed assignments.

**POLICY ON CLASSROOM BEHAVIOR:**

Students must behave in a professional manner. Students cannot engage in constant chatting with others, must not use computer to play video games, should avoid using computer for purposes unrelated to the class, and need to turn off cell phones.

**POLICY ON ACADEMIC DISHONESTY:**

By enrolling in this class the student agrees to uphold the standards of academic integrity described in the catalog at:

<http://www.csueastbay.edu/ecat/current/i-120grading.html#section12>

The University has a published policy on cheating and academic dishonesty. Students are expected to be familiar with the policy and to abide by it. Cheating will result in: 1) a zero score on the test and the loss of all grading options; and/or 2) an "F" grade for the course; and/or 3) referral to the Academic Vice President for expulsion from the University.

**ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES:**

If you have a documented disability and wish to discuss academic accommodations, or if you would need assistance in the event of an emergency evacuation, please contact me as soon as possible. Students with disabilities needing accommodation should speak with the Accessibility Services.

**EMERGENCY INFORMATION:**

California State University, East Bay is committed to being a safe and caring community. Your appropriate response in the event of an emergency can help save lives. Information on what to do in an emergency situation (earthquake, electrical outage, fire, extreme heat, severe storm, hazardous materials, and terrorist attack) may be found at: h[ttp://www.aba.csueastbay.edu/EHS/emergency\_mgnt.htm](http://www.aba.csueastbay.edu/EHS/emergency_mgnt.htm)

**UNANTICIPATED CIRCUMSTANCES:**

Unanticipated circumstances including discovery of the need to spend more time mastering particular content may require changes to the syllabus. In such situations, I should discuss the need for making such changes with students prior to making them.

**SCHEDULE:**

|  |  |
| --- | --- |
| **Please read each chapter prior to the date on which it is covered!** | |
| **Date** | **Topics** |
| The 1st Week Thu, Mar 29 | Course Overview; Meet Hadoop (ch 1); MapReduce (ch 2) |
| The 2nd Week  Thu, Apr 05 | The Hadoop Distributed Filesystem (ch 3) |
| The 3rd Week  Thu, Apr 12 | Hadoop I/O (ch 4); |
| The 4th Week  Thu, Apr 19 | Developing a MapReduce Application (ch 5); **Homework 1 due Saturday, 4/21** |
| The 5th Week  Thu, Apr 26 | How MapReduce Works (ch 6) |
| The 6th Week  Thu, May 03 | MapReduce Types and Formats (ch 7); **Homework 2 due Saturday, 5/5** |
| The 7th Week  Thu, May 10 | MapReduce Features (ch 8); **Exam 1 (ch 1-6)** |
| The 8th Week  Thu, May 17 | Pig (ch 11); |
| The 9th Week  Thu, May 24 | Hive (ch 12); **Homework 3 due Saturday, 5/26** |
| The 10th Week  Thu, May 31 | HBase (ch 13); ZooKeeper (ch 14); Sqoop (ch 15); **Project due Thursday, 5/31** |
| Thu, Jun 07 | **6:30PM-8:00PM, Exam 2** |