《数据库与数据挖掘》教学大纲

一、课程基本信息

开课单位:	信息科学与技术学院	课程代码:	CS150
课程名称:	数据库与数据挖掘	英文名称:	Database and Data Mining
学 分:	4	学 时:	64
授课对象:		授课语言:	中英文
先修课程:			

二、课程简介和教学目的

This course covers database design and the use of databases in applications, with an introduction to the internals of relational database engines as well as basic data mining methods.

The main content of this course includes the following topics:
- Relational data model, relational algebra, and SQL.
- Database design and relational design principles based on dependencies and normal forms.
- Database services including protection, integrity control, and alternative views of data.
- Query optimization and introduction to transaction processing.
- Modern database systems, including Parallel Databases, NoSQL, Hadoop and Spark for large-scale data processing.
- The integration between data mining, machine learning and database technology.

三、教学内容、教学方式和学时安排

- Discussion sessions: Once every week with quiz, which is mandatory

Content	Schedule	Details
Introduction to DMS	Week 1	
SQL I		

SQL II	Week 2	
SQL III		HW1
Disk, Buffers, Files I	Week 3	
Disk, Buffers, Files II		
File Organization	Week 4	
Indexes and B+ Trees		HW1 due
Buffer Management	Week 5	
National Day		
Relational Algebra	Week 6	
Sorting and Hashing		
Iterations and Joins I	Week 7	
Iterations and Joins II		HW2
Query Optimization I	Week 8	
Query Optimization II		
Midterm	Week 9	
Transactions and Concurrency I		
Transactions and Concurrency II	Week 10	Project
Recovery I		HW2 due
Recovery II	Week 11	
ER and Relational Modeling		
Data Mining and ML I	Week12	
Data Mining and ML II		HW3
Data Mining and ML III	Week13	
Data Mining and ML IV		
Data Mining and ML V	Week14	
Data Mining and ML VI		HW3 due
NoSQL I	Week15	
NoSQL II		

NoSQL III	Week16	
Course Review		
	Week18 (January 10th)	Project Due

四、考核方式和成绩评定

- Problem Sets (20%) + Quizs (10%)
- Course project (25%)
- Midterm (20%)
- Final exam (25%)
- Notes on grading (please read):
  + The percentages are subject to change as circumstances dictate.
  + Active participation in piazza and/or section can positively affect your final grade.
  + Since exams are the main indicators we have of individual grasp of the material, we reserve the right to adjust final letter grades based on exam performance. In particular, students who do not achieve a passing average on the exams will not receive a passing grade in the class.
  + Work that you submit must be your own (or for two-person projects, the team's). We will run the standard software duplication checkers on submitted assignments.
  + We will be following the SIST policy on Academic Honesty, so be sure you are familiar with it.

五、推荐教材

Reference:
- Database Management Systems (3rd edition) by Ramakrishnan and Gehrke (R/G)
- Database Systems: The Complete Book (2nd edition) by Garcia-Molina, Ullman, and Widom (G/U/W)
- Fundamentals of Database Systems (7th edition) by Elmasri and Navathe (E/N)
- Database System Concepts (6th edition) by Silberschatz, Korth, and Sudarshan (S/K/S)