

# 课程大纲

## COURSE SYLLABUS

1.	<b>课程代码/名称</b> Course Code/Title	CSE5002/智能数据分析 Intelligent Data Analysis										
2.	<b>课程性质</b> Compulsory/Elective	专业选修课 Major elective course										
3.	<b>课程学分/学时</b> Course Credit/Hours	3/48 (48 Lectures)										
4.	<b>授课语言</b> Teaching Language	英文 English										
5.	<b>授课教师</b> Instructor(s)	唐珂, 教授, 计算机科学与工程系, tangk3@sustech.edu.cn Ke Tang, Professor, Department of Computer Science and Engineering, tangk3@sustech.edu.cn										
6.	<b>是否面向本科生开放</b> Open to undergraduates or not	是 Yes										
7.	<b>先修要求</b> Pre-requisites	(如面向本科生开放, 请注明区分内容。 If the course is open to undergraduates, please indicate the difference.) CS208 算法设计与分析 (或讲授相关内容但课程名字不同的课) MA107A 线性代数 A Linear Algebra A MA212 概率论与数理统计 Probability and Statistics										
8.	<b>教学目标</b> Course Objectives	<p>本课程介绍智能数据分析的基本思想和算法, 并侧重于基本的分析任务如分类和聚类。通过一系列的理论和实验课, 本课程将涵盖多种分类和聚类算法、输出校正、模型选择、数据表征(特征分析)、排序学习、相关性分析等。本课程的目的是让学生了解和掌握针对智能数据分析的基本思想、算法和工具, 同时兼顾对前沿方向的分析与解读, 以便为学生将来的研究和工作打下基础。</p> <p>This course will provide an introduction to intelligent data analysis with special emphasis on fundamental data analysis tasks such as classification and clustering. Through lectures and labs, we will explore topics including classification and clustering algorithms, output calibration and model selection, data representation (feature analysis), learning-to-rank, correlation analysis, etc. Introduction and discussion on emerging topics will also be covered. The objective of this course is to enable each student to understand and master some basic ideas, algorithms and tools for intelligent data analysis tasks, so as to have a foundation for their further research and work.</p>										
9.	<b>教学方法</b> Teaching Methods	理论课 48 lectures										
10.	<b>教学内容</b> Course Contents	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center;"><b>Section 1</b></td> <td>课程介绍 Introduction</td> </tr> <tr> <td style="text-align: center;"><b>Section 2</b></td> <td>生成模型与 k 近邻 Generative Models and k-Nearest Neighbors</td> </tr> <tr> <td style="text-align: center;"><b>Section 3</b></td> <td>支持向量机 Support Vector Machines</td> </tr> <tr> <td style="text-align: center;"><b>Section 4</b></td> <td>决策树 Decision Trees</td> </tr> <tr> <td style="text-align: center;"><b>Section 5</b></td> <td>神经网络 Neural Networks</td> </tr> </table>	<b>Section 1</b>	课程介绍 Introduction	<b>Section 2</b>	生成模型与 k 近邻 Generative Models and k-Nearest Neighbors	<b>Section 3</b>	支持向量机 Support Vector Machines	<b>Section 4</b>	决策树 Decision Trees	<b>Section 5</b>	神经网络 Neural Networks
<b>Section 1</b>	课程介绍 Introduction											
<b>Section 2</b>	生成模型与 k 近邻 Generative Models and k-Nearest Neighbors											
<b>Section 3</b>	支持向量机 Support Vector Machines											
<b>Section 4</b>	决策树 Decision Trees											
<b>Section 5</b>	神经网络 Neural Networks											

	<b>Section 6</b>	输出校准、性能评估和模型选择 Output Calibration, Performance Evaluation and Model Selection
	<b>Section 7</b>	智能优化方法 Intelligent Optimization Methods
	<b>Section 8</b>	特征分析和降维 Feature Analysis and Dimensionality Reduction
	<b>Section 9</b>	机器学习理论 Theoretical Foundation of Machine Learning
	<b>Section 10</b>	多分类器系统 Multiple Classifier Systems
	<b>Section 11</b>	不平衡学习、代价敏感学习和半监督学习 Imbalanced Learning, Cost-sensitive Learning and Semi-Supervised Learning
	<b>Section 12</b>	图数据挖掘 Mining Graph Data
	<b>Section 13</b>	排序与聚类 Learning to Rank and Clustering
	<b>Section 14</b>	主动学习与强化学习 Active Learning and Reinforcement Learning
	<b>Section 15</b>	多任务学习、增量式学习与自动机器学习 Multi-task Learning, Incremental Learning and Automated Machine Learning
	<b>Section 16</b>	课程总结和回顾 Summary and Review
<b>11.</b>	<b>课程考核</b> <b>Course Assessment</b>	
	<p>( ① 考核形式 Form of examination; ②. 分数构成 grading policy; ③ 如面向本科生开放, 请注明区分内容。 If the course is open to undergraduates, please indicate the difference. )</p> <p><b>课程项目 Projects 50%</b> <b>期末考试 Final Exam 50%</b></p>	
<b>12.</b>	<b>教材及其它参考资料</b> <b>Textbook and Supplementary Readings</b>	
	<p>T. Hastie, R. Tibshirani, J. Friedman, 《The Elements of Statistical Learning: Data Mining, Inference, and Prediction》 Christopher M. Bishop, 《Pattern Recognition and Machine Learning》</p>	