

## 课程大纲 COURSE SYLLABUS

1.	<b>课程代码/名称 Course Code/Title</b>	<b>MAT7100 统计深度学习 MAT7100 Statistical Deep Learning</b>
2.	<b>课程性质 Compulsory/Elective</b>	专业选修课 Major Elective Courses
3.	<b>课程学分/学时 Course Credit/Hours</b>	3/48
4.	<b>授课语言 Teaching Language</b>	英文 English
5.	<b>授课教师 Instructor(s)</b>	陈欣 CHEN Xin
6.	<b>是否面向本科生开放 Open to undergraduates or not</b>	是 Open to undergraduates
7.	<b>先修要求 Pre-requisites</b>	(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.) 统计线性模型 (MA329) Statistical Linear Models (MA329) 广义线性模型 (MA403) Generalized Linear Models (MA403)
8.	<b>教学目标 Course Objectives</b>	<p>本课程首先介绍统计学习的基本概念, 它们是现代统计学的基石。本课程也为进一步学习其他领域比如大数据, 深度学习, 人工智能 打下良好的基础。本课程还重点介绍了现代统计很多方法。统计深度学习可应用于许多学科领域。基本教学目标是掌握比如主成分分析, 正则化回归分析, LASSO 型变量选择, 充分降维, 决策树, 分类方法等。基本目标是教会学生掌握统计学习和现代统计方法, 培养学生的统计学思维和分析数据的能力, 并为后续课程打下良好的基础。</p> <p>This course begins with an introduction to the basic concepts of statistical learning, which is the cornerstone of modern statistics. This course also lays a good foundation for further study in other areas such as big data, deep learning, and artificial intelligence. This course also highlights many of the methods of modern statistics. Statistical deep learning can be applied to many subject areas. The basic teaching objectives are to master such as principal component analysis, regularized regression analysis, LASSO-type variable selection, sufficient dimension reduction, decision tree, and classification methods. The basic goal is to teach students to master statistical learning and modern statistical methods, to develop students' statistical thinking and ability to analyze data, and to lay a good foundation for follow-up courses.</p>
9.	<b>教学方法 Teaching Methods</b>	讲授 Lectures
10.	<b>教学内容 Course Contents</b>	<p><b>Section 1</b> 基本回归线性方法的介绍 (8 学时) 最小二乘法, 模型选择, 岭回归, 主成分分析, 偏最小二乘法。 Introduction of Basic Linear Methods for Regression (12 hours): Least squares, Model selection, ridge regression, Principal Components Regression, Partial Least Squares.</p> <p><b>Section 2</b> 正则化回归分析 (8 小时): Lasso, 自适应 Lasso, 群 Lasso, 偏差 - 方差权衡, 对逻辑回归模型的扩展。</p>

	Regularized regression analysis (8 hours): Lasso, adaptive lasso, group lasso, Bias–Variance Tradeoff, Extension to the logistic regression model.
<b>Section 3</b>	充分降维 (12 学时) 充分降维的基本定义, 切片逆回归, 稀疏型充分降维, 充分性检验。 Sufficient dimension reduction (12 hours) Definition of Sufficient of dimension reduction, sliced inverse regression, sparse dimension reduction, test for sufficiency .
<b>Section 4</b>	局部平均方法(12 学时) CART, K-最近邻, 样条和核平滑方法。 Local averaging methods (12 hours) CART, K-nearest neighbor, spline and kernel smoothing .
<b>Section 5</b>	Boosting 和神经网络简介 (8 学时) Boosting 和神经网络简介 Introduction to Boosting and Neural networks (8hours) Introduction to Boosting and Neural networks.
<b>Section 6</b>	
<b>Section 7</b>	
<b>Section 8</b>	
<b>Section 9</b>	
<b>Section 10</b>	
.....	
<b>11. 课程考核</b> <b>Course Assessment</b>	
	(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.) 小测验 Quiz 10% 平时作业 Assignments 20% 期中考试 Mid-Term Test 30% 期末考试 Final Exam 40%
<b>12. 教材及其它参考资料</b> <b>Textbook and Supplementary Readings</b>	
	The Elements of Statistical Learning 2nd edition February 2009 By Jerome H. Friedman, Robert Tibshirani, and Trevor Hastie. Publisher: Springer