《优化与机器学习》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:	SI151
课程名称:	优化与机器学习	英文名称:	Optimization and Machine Learning
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

本课程广泛介绍了用于机器学习和统计学习的实用算法和理论，尤其着重于学习模型，优化算法和统计分析。我们涵盖的主题包括监督学习（例如，回归，分类，lasso，逻辑回归，支持向量机，神经网络），无监督学习（例如，聚类，降维，非负矩阵分解，密度估计），统计学习理论（偏差/方差折衷，PAC 学习理论，样本复杂度），概率图模型，强化学习和主动学习。本课程还将介绍适用于机器学习应用中出现的大规模问题的优化方法（例如，梯度方法，近端方法，ADMM，拟牛顿方法以及大规模数值线性代数）。该课程将为本科生提供关于优化和机器学习在算法，理论和应用方面的基础入门知识。

三、教学内容、教学方式和学时安排

SI151 优化与机器学习课程表

(这是暂定的课程表，之后可能会有变化)

教学周	**课程**	**主题**	**阅读**
第 1 周	课程介绍 监督学习概述 I	• 机器学习实例 • 课程安排 • 最小二乘法和最近邻域法 • 统计决策理论	ESL: 第 1，2 章 PRML: 第 1 章
	监督学习概述 II	• 高维局部方法 • 统计模型 • 模型选择	ESL: 第 1，2 章 PRML: 第 1 章
	线性回归方法 I	• 线性回归模型 • 子集选择法 • 前向选择和后向选择	ESL: 第 3 章 PRML: 第 3 章

第 2 周	线性回归方法 II	• 收缩法 • 岭回归,lasso 和 LAR • 子集选择法和收缩法的比较	ESL: 第 3 章 PRML: 第 3 章
	作业: 习题集 1 第 4 周（周四）23:59 到期		
第 3 周	线性分类方法 I	• 从回归到分类 • 用于分类的线性回归 • 线性判别分析	ESL: 第 4 章
	线性分类方法 II	• 正则化的判别分析 • 费雪公式 • 逻辑回归	PRML: 第 4 章 ESL: 第 5 章
第 4 周	概率与估计	• 贝叶斯准则 • 最大似然估计 • 最大后验概率	ML（第二版）: 第 2 章
	朴素贝叶斯	• 条件独立 • 朴素贝叶斯 • 高斯朴素贝叶斯	ML（第二版）: 第 3 章
	作业: 习题集 2 第 6 周（周四）23:59 到期		
第 5 周	概率图模型 I	• 贝叶斯网络 • 条件独立下的联合分布表示	PRML: 第 8 章
	概率图模型 II	• 马尔可夫随机场 • 概率图模型中的推断	PRML: 第 8 章
第 6 周	混合模型和 EM I	• EM 算法 • 混合高斯聚类	ML（第一版）: 第 6 章 PRML: 第 9 章
	混合模型和 EM II	• 混合高斯聚类 • Chow-Liu 算法	ML（第一版）: 第 6 章 PRML: 第 9 章
第 7 周	学习理论 I	• PAC 和统计学习理论 • 样本复杂度 • VC 维	ML（第一版）: 第 7 章
	学习理论 II	• Rademacher 复杂度 • 过拟合和正则化	ML（第一版）: 第 7 章

	作业：习题集 3 第 9 周（周四）23:59 到期		
第 8 周	集成学习 I	• 弱与强 (PAC) 学习理论 • boosting 和 bagging • boosting 精度	ESL：第 10,16 章
	集成学习 II	• Adaboost • Adaboost-泛化保证 • 几何边距和感知器算法	•
第 9 周	核方法	• 几何边距 • 内核化学习算法 • 内核化感知器	PRML：第 6 章
	支持向量机	• SVM：原始和对偶形式 • 内核化 SVM • 半监督 SVM	PRML：第 7 章 ESL：第 12 章
	作业：习题集 4 第 11 周（周四）23:59 到期		
第 10 周	半监督学习	• 转导 SVM • 联合训练和多视角学习 • 基于图的方法	•
	主动学习	• 批量主动学习 • 选择性采样和主动学习 • 采样偏差	•
	课程 project 第 18 周（周日）23:59 到期		
第 11 周	无监督学习 I	• k-means, 密度聚类 • 层次聚类 • 谱聚类	PRML：第 9 章 ESL：第 14 章
	无监督学习 II	• 主成分分析 • 内核化主成分分析 • 稀疏主成分分析	PRML：第 12 章
第 12 周	神经网络和深度学习	• 误差反向传播 • 卷积神经网络 • 递归神经网络	PRML：第 5 章 ESL： 第 11 章
	强化学习	• 马尔科夫决策过程 • 价值迭代 • Q 学习	ML（第一版）：第 13 章

	作业：习题集 5 第 14 周（周四）23:59 到期		
第 13 周	优化入门 凸集 I	• 最小二乘法 • 线性规划 • 仿射集，凸集和凸锥集	CVX：第 1，2 章
	凸集 II 凸函数 I	• 保留集合凸性的运算 • 广义不等式 • 分离和支持超平面 • 凸函数的基本特性	CVX：第 2，3 章
第 14 周	凸函数 II	• 共轭函数 • 拟凸函数 • 关于广义不等式的凸性	CVX：第 3 章
	凸优化问题	• 线性优化 • 二次优化 • 向量优化	CVX：第 4 章
	作业：习题集 6 第 16 周（周四）23:59 到期		
第 15 周	对偶	• 拉格朗日对偶问题 • 拉格朗日对偶的解释 • 最优条件	CVX：第 5 章
	无约束优化	• 梯度下降法 • 最陡下降法 • 牛顿方法	CVX：第 9 章
第 16 周	等式约束优化 不等式约束优化 I	• 等式约束的牛顿方法 • 广义牛顿方法 • 内点法	CVX：第 10，11 章
	不等式约束优化 II 课程复习	• 基于广义不等式的优化 • 原始-对偶内点法 • 课程回顾	CVX：第 11 章
第 18 周	**期末考试** **时间：（周二）08:15-09:55，地点：教学中心 xxx，形式：闭卷+2 张 A4 大小 cheat sheet**		

四、考核方式和成绩评定

作业：30%

课程 project：30%
期末考试：40%

五、推荐教材

书名	作者	译者	出版社	出版时间	ISBN
The Elements of Statistical Learning: Data Mining, Inference, and Prediction	Trevor Hastie, Robert Tibshirani, and Jerome Friedman		Springer		
Pattern Recognition and Machine Learning	Christopher Bishop		Springer		
Machine Learning	Tom Mitchell		McGraw Hill		
Convex Optimization	Stephen Boyd and Lieven Vandenberghe		Cambridge University Press		

六、参考书目

书名	作者	译者	出版社	出版时间	ISBN
Numerical Optimization	Jorge Nocedal and Stephen J. Wright		Springer-Verlag		
Machine Learning: A Probabilistic Perspective	Kevin P. Murphy		The MIT Press		
Deep Learning	Aaron Courville, Ian Goodfellow, and Yoshua Bengio		The MIT Press		

机器学习	周志华	清华大学出版社

七、其他说明

1. 请及时查阅 Blackboard 上的课程通知，相关资料和作业；
2. 请使用英文完成作业，project 以及期末考试；
3. 对于晚交的作业，其分数将根据迟交天数成倍降低；
4. 任何抄袭和作弊行为一经确认，相关作业或考试将进入 0 分;
5. 不努力学习的人可能会挂掉本课程。

八、教师信息和开课单位审核意见

授课教师	（签名） 年 月 日	邮 箱	shiym@shanghaitech.edu.cn
		电 话	
	（签名） 年 月 日	邮 箱	sunlu1@shanghaitech.edu.cn
		电 话	
课程负责人 （大纲负责人）	（签名） 年 月 日		
开课单位审核意见	（签名） 年 月 日		

《Optimization and Machine Learning》Syllabus

1.Basic course information

unit:	School of Information Science and Technology	course code:	SI151
course name:	优化与机器学习	course name:	Optimization and Machine Learning
credits:	4	period:	64
teaching object:		teaching language:	Chinese and English
previous course:			

2.Course introduction and teaching purpose

This course provides a broad introduction to practical algorithms and theory for machine learning and statistical learning, with particular emphasis on learning models, optimization algorithms and statistical analysis. We cover topics such as supervised learning (e.g., regression, classification,lasso, logistic regression, support vector machines, neural networks), unsupervised learning (e.g.,clustering, dimensionality reduction, non-negative matrix factorization, density estimation),
statistical learning theory (bias/variance tradeoffs; PAC learning theory, sample complexity), graphical models, reinforcement learning and active learning. This course will also introduce optimization methods (e.g., gradient methods, proximal methods, ADMM, quasi-Newton methods, as well as large-scale numerical linear algebra) that are suitable for large-scale problems arising in machine learning applications. The course is expected to provide an undergraduate student with a fundamental understanding on machine learning and optimization in terms of algorithm, theoryand application.

3.Teaching content, teaching method and teaching time arrangement

Course Schedule of SI151 Optimization and Machine Learning

(This is a tentative schedule and is subject to change)

Week	Lecture	Topics	Readings
Week 1	Introduction Overview of Supervised Learning I	• Machine learning examples • Course logistics • Least squares and nearest neighbors • Statistical decision theory	ESL: Ch 1, 2 PRML: Ch 1
	Overview of Supervised Learning II	• Local methods in high dimensions • Statistical models • Model selection	ESL: Ch 1, 2 PRML: Ch 1
	Linear Methods for Regression I	• Linear regression models • Subset selection • Forward/Backward Selection	ESL: Ch 3 PRML: Ch 3

Week 2	Linear Methods for Regression II	• Shrinkage methods • Ridge regression, lasso and LAR • Comparison of subset selection and shrinkage	ESL: Ch 3 PRML: Ch 3
	Assignment: Problem Set 1 Due Thursday, Week 4 at 11:59pm		
Week 3	Linear Methods for Classification I	• From regression to classification • Linear regression for classification • Linear discriminant analysis	ESL: Ch 4
	Linear Methods for Classification II	• Regularized discriminant analysis • Fisher's formulation • Logistic regression	PRML: Ch 4 ESL: Ch 5
Week 4	Probability and Estimation	• Bayes rule • MLE • MAP	ML (2nd edition): Ch 2
	Naive Bayes	• Conditional Independence • Naive Bayes: why and how • Gaussian Naïve Bayes	ML (2nd edition): Ch 3
	Assignment: Problem Set 2 Due Thursday, Week 6 at 11:59pm		
Week 5	Graphical Models I	• Bayesian networks • Representation joint distribution with conditional independence assumption	PRML: Ch 8
	Graphical Models II	• Markov random fields • Inference in graphical models	PRML: Ch 8
Week 6	Mixture Models and EM I	• EM algorithm • Mixture of Gaussian clustering	ML (1nd edition): Ch 6 PRML: Ch 9
	Mixture Models and EM II	• Mixture of Gaussian clustering • Chow-Liu algorithm	ML (1nd edition): Ch 6 PRML: Ch 9

Week 7	Learning Theory I	• PAC and learning theory • Sample complexity • VC dimension	ML (1nd edition): Ch 7
	Learning Theory II	• Rademacher complexity • Overfitting and regularization	ML (1nd edition): Ch 7
	Assignment: Problem Set 3 Due Thursday, Week 9 at 11:59pm		
Week 8	Ensemble Learning I	• Weak vs strong (PAC) learning • Boosting and bagging methods • Boosting accuracy	ESL: Ch 10, 11
	Ensemble Learning II	• Adaboost • Adaboost: generalization guarantees, geometric margins and perceptron	
Week 9	Kernels	• Geometric margins • Kernels: kernelizing a learning algorithm • Kernelized Perceptron	PRML: Ch 6
	Support Vector Machines	• SVM: primal and dual forms • Kernelizing SVM • Semi-supervised SVM	PRML: Ch 7 ESL: Ch 12
	Assignment: Problem Set 4 Due Thursday, Week 11 at 11:59pm		
Week 10	Semi-supervised Learning	• Transductive SVM • Co-training and multi-view learning • Graph-based methods	•
	Active Learning	• Batch active learning • Selective sampling and active learning • Sampling bias	•
	Course Project Due Sunday, week 18 at 11:59pm		
	Unsupervised Learning I	• k-means, density clustering • Hierarchical clustering • Spectral clustering	PRML: Ch 9 ESL: Ch 14

Week 11	Unsupervised Learning II	• Principal components • Kernel principal components • Sparse principal components	PRML: Ch 12
Week 12	Neural Networks and Deep Learning	• Error backpropagation • Convolutional neural networks • Recurrent neural networks	PRML: Ch 5 ESL: Ch 11
	Reinforcement Learning	• Markov decision processes • Value Iteration • Q-learning	ML (1nd edition): Ch 13
	Assignment: Problem Set 5 Due Thursday, Week 14 at 11:59pm		
Week 13	Introduction to Optimization Convex Set I	• Least squares • Linear programming • Affine set, convex set and convex cone	CVX: Ch 1, 2
	Convex Set II Convex Function I	• Operations that preserve convexity • Generalized inequality • Separating and supporting hyperplane. • Basic properties of convex functions	CVX: Ch 2, 3
Week 14	Convex Function II	• Conjugate function • Quasiconvex function • Convexity w.r.t. generalized inequality	CVX: Ch 3
	Convex Optimization Problems	• Linear optimization • Quadratic optimization • Vector optimization	CVX: Ch 4
	Assignment: Problem Set 6 Due Thursday, Week 16 at 11:59pm		
	Duality	• The Lagrange duality problem • Interpretations of the Lagrange duality • Optimality conditions	CVX: Ch 5

Week 15	Unconstraint Optimization	• Gradient descent method • Steepest descent method • Newton's method	CVX: Ch 9
Week 16	Equality Constraint Optimization Inequality Constraint Optimization I	• Newton's method with equality constraints • Generalized Newton's method • Interior point method	CVX: Ch 10, 11
	Inequality Constraint Optimization II Course Review	• Optimization with generalized inequality constraints • Primal-dual interior point method • Course review	CVX: Ch 11
Week 18	**Final Exam** 08:15—09:55 (Tuesday) at xxx, closed book (two A4 cheat sheets are allowed)		

4.Assessment methods and performance evaluation

Homework: 30%
Course project: 30%
Final exam: 40%

5.Other instructions

1. Please check the course notice, related materials and assignments on Blackboard in time;
2. Please write your homework, project and exam in English;
3. For late homework or project, the score will be exponentially decreased;
4. Once any plagiarism or cheating is confirmed, relevant assignments or exams will receive 0 points;
5. Some may fail if they don't work hard.

6.Teachers' information and audit institute

teacher	(signature) / / /	email	shiym@shanghaitech.edu.cn
		telephone	
	(signature) / / /	email	sunlu1@shanghaitech.edu.cn
		telephone	

Course head (Syllabus head)	(signature) / / /
Institute of audit opinion	(signature) / / /