《机器学习引论》教学大纲

一、课程基本信息

课程名称/英文名称:	机器学习引论/Introduction to Machine Learning	课程代码:	CS182
课程层次:	本科生课程	学 分/学 时:	4/64
主要面向专业:	电子信息工程 , 计算机科学与技术	授课语言:	中英文
先修课程:	无	建议先修课程说明:	Mathematical Analysis II, Linear Algebra,Probability and Statistic
开课单位:	信息科学与技术学院	课程负责人:	师玉娇

二、课程简介

本课程广泛介绍了用于机器学习的实用理论和算法，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，逻辑回归，神经网络)，无监督学习(例如，聚类，
维)，集成学习等内容。

三、课程教学目标

通过本课程的学习，学生将掌握机器学习的基本问题定义、基本模型，对机器学习学科有概览性的认识；掌握目前主流的机器学习算法和模型，并能够根据实际问题的需要选择并实现相应的算法；编
成机器学习典型应用实例，对机器学习工程编程有初步的训练；团队合作完成1项机器学习项目，并做展示与交流。通过课程的学习，让学生能理解工程职业道德和规范，具备科学精神和工程师的基本
养，具备科技报国的家国情怀和使命担当；能进行团队协作，具备合作精神和人际沟通能力。

四、课程教学方法

课堂讲授与动手实践相结合。突出重点，深入浅出，在重视基础理论的同时，注意培养学生独立思考的能力，同时引导学生用学到的理论来解决一些实际问题。采用启发式教学方法，引导学生对问题
思考和讨论。

•课后辅导：通过适量的习题课设置与固定答疑时间的安排，巩固学生的基础，。

•课程设计：结合学生从事各研究方向的当前热点领域，提出机器学习的方法来解决实际问题，深刻理解与掌握机器学习在各个学科中的应用。

五、课程教学内容与安排

CS182机器学习引论课程表

（这是暂定的课程表，之后可能会有变化）

教学周	课时	课程主题	阅读	
分类与回归				
第一周	2	Lecture 1 : 课程介绍	Command Line and File I/O Tutorial. 10601 Course Staff (2020).10601 Learning Objectives. Matt Gormley (2023). Math Resources. 10601 Course Staff (2023).	HW1 Out
	2	Lecture 2 : Machine Learning as Function Approximation	10601 Notation Crib Sheet. Matt Gormley (2023).	
第二周	2	HW1习题课与Python实践		
	2	Lecture 3 : 决策树	Visual Information Theory. Christopher Olah (2015). blog.Decision Trees. Hal Daumé III (2017). CIML, Chapter 1.	HW1 Due HW2 Out
第三周	2	HW2习题课与Python实践		
	2	Lecture 4 : k-近邻	Geometry and Nearest Neighbors. Hal Daumé III (2017). CIML, Chapter 3.	
第四周	2	Lecture 5 : 模型选择与实验设计	Limits of Learning. Hal Daumé III (2017). CIML, Chapter 2.	
线性模型				
第四周	2	Lecture 6 : 感知机	The Perceptron. Hal Daumé III (2017). CIML, Chapter 4.	
第五周	2	Lecture 7 : 线性回归	Linear Regression. Kevin P. Murphy (2014). Machine Learning: A Probabilistic Perspective. Chapter 7.1-7.3.	HW2 Due HW3 Out
	2	Lecture 8 : 机器学习中的优化	An overview of gradient descent optimization algorithms. Sebastian Ruder (2017).	
第六周	2	HW3习题课与Python实践		
	2	Lecture 9 : 随机梯度下降/逻辑回归	Estimating Probabilities: MLE and MAP. Tom Mitchell (2016).Probability and Information Theory. Ian Goodfellow, Yoshua Bengio, & Aaron Courville (2016). Deep Learning, Chapter 3.Logistic Regression. Kevin P. Murphy (2014). Machine Learning: A Probabilistic Perspective. Chapter 1.4.6, 8.1-8.3, 8.6.	HW3 Due (only two grace/late days permitte
第七周	2	Lecture 10 : 特征工程/正则化	Regularization for Deep Learning. Ian Goodfellow, Yoshua Bengio, & Aaron Courville (2016). Deep Learning, Chapter 7.1 and 7.8.	
神经网络				
第七周	2	Lecture 11 : 神经网络	Deep Feedforward Networks. Ian Goodfellow and Yoshua Bengio and Aaron Courville (2016). Deep Learning, Chapter 6.1-6.4.	HW4 Out
第八周	2	Lecture 12 : 反向传播 I		
	2	HW4习题课与Python实践		
第九周	2	Lecture 13 : 反向传播 II		
深度学习				
第九周	2	Lecture 14 : 卷积神经网络（CNNs）和循环神经网络（RNNs）		HW4 Due

第十周	2	Lecture 15 : 基于RNN的语言模型（RNN-LMs）和基于Transformers的语言模型（Transformers-LMs）		
	2	Lecture 16 : 自动微分（AutoDiff）、预训练、微调（Fine-Tuning）、上下文学习（In-context Learning）		HW6 Due (only two grace/late days permit
Project Out				
Societal Impacts				
第十一周	2	Lecture 17 : Societal Impacts of ML		HW5 Out
	2	HW5习题课与Python实践		
学习理论				
第十二周	2	Lecture 18 : PAC Learning		
	2	Lecture 19 : PAC Learning / MLE & MAP		HW6 Out
第十三周	2	HW6习题课与Python实践		HW5 Due
Learning Paradigms				
第十三周	2	Lecture 20 : 推荐系统 / 集成学习: Boosting		
第十四周	2	Lecture 21 : 集成学习: Bagging / K-Means		
	2	Lecture 22 : 降维: PCA		HW6 Due (only two grace/late days permit HW7 Out
第十五周	2	HW7习题课与Python实践		
	2	Lecture 23 : Special Topics: Generative Models for Vision / Significance Testing for ML		
第十六周	4	Advanced Topics， 内容回顾，考试重点		

六、考核方式和成绩评定方法

```
作业：30%
课程project：30%
期末考试：40%
```

七、教材和参考书目

(一)、推荐教材

(二)、参考书目

书名	作者	译者	出版社	出版年月	ISBN	版次
机器学习	周志华		清华大学出版社			
李航	统计学习方法		清华大学出版社			
Introduction to Machine Learning	Ethem Alpaydin		MIT Press Ltd	2020-03	978-0262043793	Fourth Editic

Probabilistic Machine Learning: An Introduction	Kevin P. Murphy	The MIT Press
Pattern Recognition and Machine Learning	Christopher Bishop	Springer
Convex Optimization	Stephen Boyd and Lieven Vandenberghe	Cambridge University Press
The Elements of Statistical Learning: Data Mining, Inference, and Prediction	Trevor Hastie, Robert Tibshirani, and Jerome Friedman	Springer

八、学术诚信教育

本课程高度重视学术诚信，严禁抄袭、作弊等行为。
“在学习、科研、实习实践等活动中，学生应恪守学术道德，坚守学术诚信，保护知识产权，坚持勇于创新、求真务实的科学精神，努力培养自己严谨求实、诚实自律、真诚协作的科学态度，成为良好学术风气的维护者、严谨治学的力行者、优良学术道德的传承者。”

九、其他说明(可选)

无

《Introduction to Machine Learning》 Syllabus

1.Basic course information

course name	Introduction to Machine Learning	course code	CS182
Course Level	Undergraduate	Credit/Contact Hour:	4/64
Major:	Electronic Information Engineering , Computer Science and Technology	Teaching Language	Chinese and English
Prerequisite	NULL	Prerequisite suggestion	
School/Institute	School of Information Science and Technology	Instructor	shiyujiao

2.Course Introduction

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, logistic regression, neural networks), unsupervised learning (e.g., clustering, dimensionality reduction), and ensemble learning. The course is expected to provide an undergraduate student with a fundamental understanding of machine learning and optimization in terms of algorithm, theory, and application.

3.Learning Goal

4.Instructional Pedagogy

• Teaching and discussion: The knowledge points of this courses are basically introduced via on-site teaching. We pay attention to the teaching of key and difficult points of the course, and use heuristic teaching methods to guide students to think and discuss the problems, and encourage students to analyze and solve related problems in the field of database in terms of mathematics, physics and engineering.
• Offline/online discussion and Q&A: Course Q&A will be conducted both offline and online (piazza) to provide supplements for course teaching and discussion, and to answer students' personalized questions in a timely manner.
• Curriculum practice: The practical teaching of this course is conducted based on the course project. Through analyzing and solving the practical project on database and data mining, students can understand the course knowledge from the viewpoint of engineering. It stimulates the research interest and innovative thinking of students.

5. Course Content and Schedule

6.Grading Policy

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

7. Textbook & Recommended Reading

(1) Textbook

(2) Recommended Reading

book name	author	translator	press	publication time	ISBN	edition
机器学习	周志华		清华大学出版社			
李航	统计学习方法		清华大学出版社			
Introduction to Machine Learning	Ethem Alpaydin		MIT Press Ltd	2020-03	978-0262043793	Fourth Edition
Probabilistic Machine Learning: An Introduction	Kevin P. Murphy		The MIT Press			
Pattern Recognition and Machine Learning	Christopher Bishop		Springer			
Convex Optimization	Stephen Boyd and Lieven Vandenberghe		Cambridge University Press			
The Elements of Statistical Learning: Data Mining, Inference, and Prediction	Trevor Hastie, Robert Tibshirani, and Jerome Friedman		Springer			

8.Academic Integrity

```
This course highly values academic integrity. Behaviors such as plagiarism and cheating are strictly prohibited. Please list more if you have more specific requirements.
```

9.Other Information (Optional)

```
NA
```