

时间	主题	简介	地点
13 日晚 , (周五) 19:00 -22:00	Introduction to Data Mining and Data Processing Techniques	This lecture will give a short introduction to data mining: what is data mining, why data mining is important for data analytics, and how to achieve high efficiency. This lecture also presents some data preprocessing techniques which are often used to get the data cleaned and get ready the data for subsequent management.	信 息 工 程 学 院 副 203
14 日下午 (周六) 14: 00- 17: 00	Frequent pattern mining and association study (5-6 hours)	This lecture will present definitions for frequent itemsets, closed patterns, frequent generators, and maximal patterns, and describe the algorithms for mining these patterns and the related association rules. This is foundation lecture for the whole subject	信 息 工 程 学 院 副 203
15 日下午 (周日) 14: 00- 17: 00	Emerging patterns and statistically important pattern	This lecture will describe definition and theories related to emerging patterns. Emerging patterns are a type of contrast patterns useful for descriptive data mining and classification. In this lecture, I also give introduction to statistically important patterns such as odd ratio patterns, relative risk patterns, propensity patterns, and how they are related to emerging patterns.	信 息 工 程 学 院 副 203
18 日晚上 (周三) 19:00 -22:00	Graph mining	I will describe algorithms for mining maximal bicliques, cliques, k-cores, quasi-bicliques and dense subgraphs. I also give a brief introduction to biomedical graph mining.	信 息 工 程 学 院 副 203
19 日晚上 (周四) 19:00 -22:00	Classification algorithms	In this lecture, I will talk about the classical classification methods such as decision trees, ensemble learning algorithms, neural networks, support vector machines, instance-based learning and eager leaning methods.	信 息 工 程 学 院 副 203
21 日下午 (周六) 14:00- 17:00	Clustering and biclustering algorithms	I will give an introduction to some widely used clustering algorithms. This include k-means, hierarchical and divisive clustering methods. Biclustering is another important topic. I will use example to show the importance of biclustering for real-life applications.	信 息 工 程 学 院 副 203
22 日下午 (周日) 14:00- 17:00	Gene expression data analysis.	I will give an introduction to gene expression data analysis: what is gene expression, how the expression levels of the genes in a genome are measured and acquired, where to get high-quality data sets. In this lecture, I will also talk about rule discovery from paired miRNA-mRNA data sets for functional module identification.	信 息 工 程 学 院 副 203

25 日晚上 (周三) 19:00 -22:00	Graph mining for protein binding hotspot detection	I will describe what is a protein binding hotspot, how to translate this biological problem into a matched graph mining problem, and where is the algorithm complexity. I will show real examples of binding hotspots and the importance to detect them.	信 息 工 程 学 院 副 203
26 日晚上 (周四) 19:00 -22:00	Conformational B-cell epitope prediction methods	In this lecture, I will give definitions for conformational B-cell epitopes. I will introduce positive-unlabeled learning algorithms for the discovery of unknown epitopes related to antigen-antibody binding.	信 息 工 程 学 院 副 203
Note: Lecture 2 may take 5-6 hours to complete. All the other lectures may take 3 hours each.			