

داده کاوی پیشرفته

نام درس			داده کاوی پیشرفته
نام درس به انگلیسی			Advanced Data Mining
نوع درس، مقطع، و واحد			گروه ۲
رشته و گرایش			مهندسی کامپیوتر
درس‌های پیش‌نیاز			هوش مصنوعی و رباتیکز
اهداف درس			برای تحلیل داده‌هایی مانند متون، گراف‌ها (شبکه‌های اجتماعی و وب)، داده‌های مکانی (spatial)، زمانی (temporal) و سری‌های زمانی روش‌های به مراتب پیچیده‌تری در مقایسه با روش‌های معمول داده کاوی مورد نیاز است. در این درس روش‌های کاوش در داده‌های پیچیده مطرح و بحث می‌گردند. همچنین، در این درس مروری کامل بر روش‌های خوشه‌بندی صورت خواهد گرفت. علاوه بر این، مفروض مباحث مطرح شده کار روی داده‌های بزرگ است.
سرفصل‌های درس			- مقدمه - مروری بر تئوری‌های احتمالات، تصمیم‌گیری و اطلاعات - یادگیری بی‌نظارت (خوشه‌بندی) ○ خوشه‌بندی مبتنی بر مرکز نقل (K-means, K-medoids, ...) ○ خوشه‌بندی مبتنی بر توزیع (EM) ○ خوشه‌بندی مبتنی بر اتصال (سلسله مراتبی) ○ خوشه‌بندی مبتنی بر چگالی ○ خوشه‌بندی مبتنی بر گراف (Chameleon, ...) ○ خوشه‌بندی داده‌ها با ابعاد بزرگ (Subspace clustering, ...) ○ روش‌های ارزیابی کیفیت خوشه‌بندی - کاهش ابعاد (Dimensionality Reduction)، مروری بر تکنیک‌هایی نظیر: - Filter-type methods, F-test, mutual information - max-relevance min-redundancy algorithm, feature stability algorithms - Wrapper methods, search methods, floating search methods - کاوش دنباله‌ها و سری‌های زمانی (مدل‌هایی برای داده‌های سری زمانی و دنباله‌ها) - روش‌های کاوش در شبکه‌های اجتماعی - روش‌های کاوش در گراف‌ها و درخت‌ها - کاربردهایی در وب (مانند تبلیغات در وب، بازاریابی ویروسی، سیستم‌های توصیه‌گر...)
نرم‌افزارهای مورد نیاز			R, Matlab, and RapidMiner
تکالیف پیشنهادی			پنج تکلیف تئوری و دو پروژه درسی
کتاب(های) مرجع			[1] Anand Rajaraman, Jure Leskovek, Jeffery D. Ullman, Mining of Massive Datasets, Cambridge University Press, 2012. [2] Jiawei Han, Micheline Kamber, Jian Pei, Data Mining: Concepts and Techniques, Third Edition, The Morgan Kaufmann Series in Data Management Systems, 2011. [3] Lei Tang, Huan Liu, Community Detection and Mining in Social Media, Morgan and Claypool Publishers, 2012. [4] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. Foundations of Machine Learning. MIT Press, 2012. [5] Kevin Murphy, Machine Learning: a Probabilistic Perspective, 2012. [6] Christopher M. Bishop, Pattern Recognition and Machine Learning. Springer



Verlag, 2006. [7] Thodoridis, S. and Koutroumbas, K. Pattern Recognition. Edition 4. Academic Press, 2008.	
[1] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition, 2009. [2] Larry Wasserman, All of Statistics: A Concise Course in Statistical Inference. Springer, 2003.	سایر مراجع

