



Microsoft®  
SQL Server®



۷ ساعت  
آموزش مفید



مدرس دوره : بهزاد رضایی

# آموزش کامل دوره پیشرفته داده کاوی



برای استعلام اسکن کنید

پشتیبانی آنلاین مدرس

پیش نمایش رایگان جهت آشنایی با دوره

شروع یادگیری بلافاصله پس از پرداخت هزینه

## جلسه اول

- الگوریتم ماشین بردار پشتیبان (*Support Vector Machine*)
- الگوریتم *Kernel SVM*
- الگوریتم بیز ساده (*Naïve Bayes*)

## جلسه دوم

- شبکه‌های عصبی مصنوعی (*Artificial Neural Networks*)
- کلاسه بندی چند کلاسه (*Multiclass Classification*)
- کلاس بندی نیمه نظارتی (*Semi-Supervised Classification*)
- بیش‌برازش و کم‌برازش (*Overfitting & Underfitting*)
- *k-fold Cross Validation*
- ماتریس درهم‌ریختگی (*Confusion Matrix*)

## جلسه سوم

- تعریف خوشه‌بندی (*Clustering*)
- کاربردهای خوشه‌بندی
- خوشه‌بندی خوب چه ویژگی‌هایی دارد؟
- ملاحظات خوشه‌بندی
- چالش‌ها و نیازمندی‌های خوشه‌بندی
- انواع روش‌های خوشه‌بندی
- *K-means*
- نکاتی در مورد الگوریتم *K-means*
- خوشه‌بندی سلسله‌مراتبی (*Hierarchical Clustering*)
- *AGNES vs. DIANA* - خوشه‌بندی سلسله‌مراتبی
- روش‌های تعیین فاصله بین خوشه‌ها
- مثال از خوشه‌بندی سلسله‌مراتبی

## جلسه چهارم

- ارزیابی خوشه‌بندی در داده کاوی
- گرایش به خوشه‌بندی
- تعداد خوشه ها
- *Elbow Method*
- *Cross Validation*
- کیفیت خوشه‌بندی
- کیفیت خوشه‌بندی – روش *Extrinsic*
- کیفیت خوشه بندی *Bcubed*
- کیفیت خوشه‌بندی – روش *intrinsic*
- کیفیت خوشه‌بندی – مثال روش سیلوئت
- خوشه‌بندی گراف‌ها بر اساس معیار *betweenness*

## جلسه پنجم

- کاوش الگوهای متناوب (*Frequent Pattern Mining*) چیست؟
- مفاهیم اولیه
- الگوهای نزدیک (*closed*) و الگوهای بیشینه (*max*)
- الگوهای نزدیک و الگوهای بیشینه – مثال
- ویژگی *Downward Closure*
- الگوریتم *Apriori*
- تولید قوانین انجمنی از مجموعه آیتم‌های پرتکرار
- معیار *Lift*

## جلسه ششم

- داده‌های پرت (*Outlier*) و تحلیل آنها
- انواع داده‌های پرت
- روش‌های تشخیص داده‌های پرت
- روش‌های تشخیص داده‌های پرت - نوع اول (بدون نظارت)
- روش‌های تشخیص داده‌های پرت - نوع اول (نیمه‌بانظارت)
- روش‌های تشخیص داده‌های پرت - نوع دوم (آماري)
- روش‌های تشخیص داده‌های پرت - نوع دوم (مبتنی بر مجاورت)
- روش‌های تشخیص داده‌های پرت - نوع دوم (مبتنی بر خوشه)
- روش‌های آماری -
- تشخیص داده‌های پرت تک‌متغیره بر اساس توزیع نرمال (*parametric*)
- روش *non-parametric* با استفاده از هیستوگرام

## جلسه ششم

- روش‌های مبتنی بر مجاورت (*proximity-based*)
- روش مبتنی بر فاصله
- روش‌های مبتنی بر خوشه‌بندی

## جلسه هفتم

- متن کاوی (*Text Mining*) چیست؟
- چرا متن کاوی دشوار است؟
- کاربردهای متن کاوی
- پیش پردازش در متن کاوی
- *Tokenization*
- نرمال سازی
- حذف *noise*
- استخراج کلمات کلیدی
- الگوریتم *RAKE*
- تشخیص موجودیت های اسمی
- (*Named Entity Recognition*)
- نمونه ابزار
- ساخت مدل در *Stanford Named Entity Recognizer*