

دانشجویان درس داده کاوی، به عنوان پروژه این درس، یکی از موارد زیر را می توانند انتخاب کنند:

۱. **ارائه یکی از عناوین نظری زیر در کلاس، بصورت کامل و جامع** (همراه با تهیه اسلایدهای مرتبط)

- روش رگرسیون ناپارامتری (NPR)
- روش رگرسیون لجستیک (LR)
- الگوریتم رگرسیون بردار پشتیبان (SVR)
- الگوریتم طبقه‌بندی ماشین بردار پشتیبان (SVM)
- الگوریتم طبقه‌بندی Naïve Bayes
- الگوریتم‌های درخت تصمیم (C4.5 و CART): فصل ۶ کتاب مرجع
- الگوریتم درخت تصمیم CHAID
- الگوریتم درخت تصمیم ID3
- الگوریتم جنگل تصادفی (RF)
- شبکه‌های عصبی مصنوعی: فصل ۷ کتاب مرجع
- شبکه‌های عصبی بازگشتی (RNN)
- شبکه‌های عصبی با تاخیر زمانی (TDNN) برای تحلیل سری‌های زمانی
- شبکه‌های کوهنن: فصل ۹ کتاب مرجع
- مدل یادگیر هیجان‌ات مغز (BEL)
- الگوریتم خوشه‌بندی DBSCAN
- الگوریتم خوشه‌بندی CURD
- الگوریتم خوشه‌بندی K-Medoid
- روش‌های خوشه‌بندی سلسله مراتبی: فصل ۸ کتاب مرجع، از صفحه ۱۴۷ تا ۱۵۳
- خوشه‌بندی k-means سلسله مراتبی: فصل ۸ کتاب مرجع، از صفحه ۱۵۳ تا ۱۶۱
- روش A priori برای استخراج قواعد انجمنی: فصل ۱۰ کتاب مرجع، از صفحه ۱۸۰ تا ۱۹۰
- روش GRI برای استخراج قواعد انجمنی: فصل ۱۰ کتاب مرجع، از صفحه ۱۹۰ تا ۲۰۰
- مدل خودرگرسیو میانگین متحرک (ARMA)، برای تحلیل سری‌های زمانی
- الگوریتم‌های Boosting (مانند AdaBoost)
- الگوریتم‌های Bagging
- الگوریتم کاهش ابعاد (مانند PCA، ICA و ...)

## ۲. آشنایی با یکی از ابزارهای داده کاوی، انجام یک پروژه دست گرمی و ارائه آن در کلاس

علاوه بر نرم افزار متلب، صدها ابزار برای داده کاوی وجود دارد. تعدادی از آنها در ادامه ذکر شده است.

- R (Rattle، یک محیط گرافیکی مناسب برای کار با R)
- Weka
- RapidMiner (قبلا به YALE معروف بود)
- Orange
- KNIME
- NLTK (natural language toolkit)
- JHepWork
- CLUTO (یکی از ابزارهای مربوط به خوشه بندی)
- Fastcluster (یکی از ابزارهای مربوط به خوشه بندی برای R و پایتون)
- Arules (یک بسته R برای کاوش قواعد انجمنی)
- ARMiner (یکی از ابزارهای مربوط به کاوش قواعد انجمنی)
- Clementic (مربوط به SPSS)
- Insightful Miner
- IBM SPSS Modeler
- ابزارهای مربوط به متن کاوی (Persian Sentence Segmenter، Persian Tokenizer، Twitter Stream API)
- ...