



درس یادگیری ماشین

مدرس: دکتر سامان هراتی زاده

الگوریتم K_2

نیم سال اول ۹۶-۱۳۹۵

تمرین شماره‌ی شش (جبرانی)

تاریخ تحویل حضوری: ۱۳۹۵/۱۱/۲۴

مهلت ارسال تمرین: ۱۳۹۵/۱۱/۲۳

- انجام این تمرین اختیاری است و تنها به درخواست مکرر دانشجویان ارائه شده است.
 - هدف از این تمرین نوشتن برنامه‌ای برنامه‌ای برای ساختن یک مدل بر مبنای الگوریتم K_2 (Bayesian Network) است.
 - مجموعه داده‌ها: داده‌های پیوست شده در مورد انتخاب یا عدم انتخاب خودرو برای خرید است. تمام ویژگی‌های این داده‌ها اسمی و ستون آخر داده‌ها نشان‌دهنده‌ی کلاس مربوط به خودرو است.
۱. الگوریتم K_2 را روی کل داده‌ها پیاده‌سازی کنید. معیارهای $recall$ ، $precision$ ، $accuracy$ و $F\text{-measure}$ را گزارش کنید.
 ۲. با استفاده از $5\text{-fold-crossvalidation}$ الگوریتم K_2 را پیاده‌سازی کنید. میانگین کل اجراها را برای تمام معیارهای فوق در این حالت گزارش کنید.
 ۳. تحلیل: معیارهای $recall$ ، $precision$ ، $accuracy$ و $F\text{-measure}$ را برای تابع‌های مختلف امتیازدهی (LL ، AIC و MDL) گزارش کنید و نتایج خود را با ذکر دلیل تحلیل کنید و در گزارش خود بنویسید.

توجه: برای اطلاعات بیشتر در مورد این الگوریتم می‌توانید به فصل ۶ کتاب داده‌کاوی و یتن مراجعه نمایید.

نکات مهم:

۱. نمره سوالات پیاده‌سازی و تحلیلی با کد به صورت زیر در نظر گرفته میشود:

۱. پیاده‌سازی ۶۰٪

۲. گزارش ۲۰٪

۳. تحویل حضوری ۲۰٪

II.نمره سوالات **تحلیلی بدون کد** به صورت زیر در نظر گرفته میشود:

۱.گزارش ۶۰٪

۲.تحویل حضوری ۴۰٪

III.نمرات کدها، گزارش و تحویل حضوری منوط به ارسال به موقع **کدها و گزارش** است