

به نام خدا

تمرین سری چهارم

یادگیری ماشین

استاد:

دکتر کمندی

دستیار آموزشی این قسمت:

آزیتا طاهری

دیتاست به پیوست تمرین بر روی سایت قرار گرفته  
است. این دیتاست شامل اطلاعات افراد سالم و بیمار  
(Breast Cancer) است.

۱- الگوریتم طبقه‌بند بیزی ساده را روی دیتاست پیوست شده، پیاده سازی کنید. تابع توزیع ویژگی‌ها را گوسین در نظر بگیرید (تابع گوسی را با استفاده از میانگین و واریانس خودتان تشکیل دهید.) نتایج (تشخیص فرد سالم از بیمار) را در قالب confusion matrix و دیگر پارامترهای سنجش دقت بیان کنید.

۲- با استفاده از PCA، تعداد بعدهای دیتاست را کاهش دهید. با توجه به نکاتی که قبلاً گفته شده، تعداد ویژگی‌ها را انتخاب کنید. در گزارش مراحل انجام کاهش بعد و دلیل انتخاب تعداد ویژگی‌ها را بیان کنید.

۳- الگوریتم طبقه‌بند بیزی را بر روی دیتاست جدید (کاهش بعد یافته) اجرا کنید. نتایج را با نتایج سوال ۱ مقایسه کنید. آیا نتایج تغییری کرده است و چرا؟

- 
- فایل‌های کد و پاسخ‌ها رابه صورت یک فایل فشرده بر روی ایلرن آپلود کنید.
  - پاسخ‌ها به صورت تایپ شده ارسال شود و پاسخ‌ها به زبان فارسی باشد.
  - به پاسخ‌های مشابه نمره‌ای داده نمی‌شود و صفر برای آنها لحاظ خواهد شد.
  - فرمت فایل ارسالی به صورت HW04.FistName.LastName\_StudentID باشد.
  - این تمرین بعد از مهلت تعیین شده تحویل گرفته نمی‌شود.