



## تمرین چهارم یادگیری ماشین

توجه:

۱. استفاده از کتابخانه‌های از پیش آماده مجاز است.
۲. لطفاً علاوه بر ضمیمه کردن کد، نتایج را تحلیل و در فایل گزارش خود ضمیمه کنید.

### سوال اول

داده‌ها در فضای دو بعدی به شما داده شده است. (آ) با استفاده از PCA مولفه‌های اصلی<sup>۱</sup> را بدست آورید. (ب) داده‌ها را بر روی مولفه اول نگاشت کنید. (ج) اکنون داده‌ها را از روی نگاشت مرحله پیشین بازسازی کنید. پروسه باید به صورت تشریحی انجام شود.

$$\left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\}$$

### سوال دوم

۱. در چه مواقعی استفاده از PCA بر LDA برتری دارد؟ و برعکس؟
۲. آیا استفاده از PCA برای مسائل supervised، مانند دسته‌بندی، فایده‌ای دارد؟ توضیح دهید.

### سوال سوم

در این تمرین دو مجموعه داده آموزش و تست در اختیار شما قرار گرفته شده است. این دادگان شامل تصاویر دو بعدی از صورت است، ولی تصاویر به صورت یک بردار یک بعدی با طول ۴۰۹۶ ذخیره شده است (بعد هر تصویر 64 × 64 است).

۱. از طریق Repeated ۵ Fold Cross-Validation با ۱۰ تکرار، خطای MSE بین تصویر بازسازی شده پس از کاهش بعد با تصویر اصلی بر روی داده ارزیابی حساب کنید و در boxplot نمایش دهید. کاهش ابعاد ۵۰۰، ۳۰۰، ۲۰۰، ۱۴۰، ۱۰۰، ۷۰، ۴۵، ۳۰، ۱۵، ۸، ۳ مورد بررسی قرار دهید.

بر اساس boxplot رسم شده، تا چه مقدار کاهش بعد، اطلاعات درون دادگان، حتی با وجود کاهش بعد، به خوبی حفظ شده است.

۲. پس از انتخاب بعد مناسب، داده تست را به آن کاهش دهید و سپس داده را بازسازی کنید. خطای MSE را گزارش دهید.

۳. ۱۰ نمونه اول داده تست به همراه بازسازی شان را نمایش دهید.

---

<sup>1</sup>Principle Component