



## تمرین سری سوم

مهلت تحویل: پایان روز ۲۵ آبان ۱۴۰۳  
طراح تمرین: فاطمه فولادی - علی رستمی

### نکات ضروری

۱. لازم است کد تمرین به همراه گزارش تحویل داده شود که شامل پاسخ سوالات نظری، توضیحات و تحلیل کد و تصاویر و نمودارهای لازم می باشد.
۲. کد تمرین باید در زبان پایتون، به صورت notebook و شامل نتایج اجرا باشد.
۳. فایل گزارش باید به فرمت pdf باشد. حجم گزارش به هیچ عنوان معیار نمره دهی نیست، لذا لطفا در حد نیاز توضیح دهید.
۴. استفاده از کتابخانه های از پیش آماده مجاز است، مگر اینکه نقیض آن گفته شود. در گزارش کتابخانه ها ذکر شوند.
۵. نمره هر تمرین از ۱۰۰ است (مگر آنکه نمره امتیازی در تمرین مشخص شده باشد). تا سه روز بعد از مهلت اعلام شده برای تمرین امکان بارگذاری وجود دارد اما به ازای هر روز ۱۰ درصد از نمره تمرین کسر خواهد شد.
۶. مشاهده شباهت نامتعارف در گزارش و کد به منزله تقلب می باشد و طرفین مشمول کسر نمره خواهند شد.
۷. برای این تمرین، می توانید گزارش و پاسخ هایتان را به زبان های فارسی و یا انگلیسی تحویل دهید.

### مسئله ۱.

داده ها در فضای دوبعدی به صورت زیر داده شده است. قسمت الف و ب را به صورت تشریحی پاسخ دهید.

(آ) با استفاده از روش LDA بهترین برداری را پیدا کنید که داده های دو کلاس بعد از نگاشت بر روی آن، بیشترین فاصله را مطابق معیار LDA داشته باشند. معیار، پارامترهای بهینه مدل و نتایج نهایی بدست آمده را گزارش دهید.

$$C_0 = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

$$C_1 = \left\{ \begin{pmatrix} -1 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\}$$

(ب) با استفاده از روش PCA بهترین برداری را پیدا کنید که واریانس داده ها بعد از نگاشت بر روی آن، بیشینه شود. مولفه های اصلی PCA را به صورت تشریحی محاسبه کنید، سپس نگاشت داده ها بر روی مولفه اول را به دست آورید و در آخر نگاشت داده ها را به فضای دوبعدی برگردانید.

$$\left\{ \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \end{pmatrix} \right\}$$

[۱۰ نمره]

### مسئله ۲.

الگوریتم LDA و PCA را از ابتدا بدون استفاده از کتابخانه های پیشرفته پیاده سازی کنید. سپس داده های سوال قبل را به عنوان ورودی به الگوریتم داده و نتیجه را گزارش کنید.

[۴۰ نمره]

### مسئله ۳.

مجموعه داده Fashion MNIST را از طریق قطعه کد زیر دانلود کنید. (ترجیحا از google colab استفاده کنید.)

```
import tensorflow as tf
import numpy as np
(train_images, train_labels), (test_images, test_labels) = tf.keras.
    datasets.fashion_mnist.load_data()
```

(آ) این دادگان شامل تصاویر دوبعدی ۲۸\*۲۸ است. ده تصویر اول را نمایش دهید. سپس داده های دوبعدی را به بردارهای یک بعدی تبدیل کنید و نرمال سازی انجام دهید.

(ب) به کمک LDA دادگان را به ابعاد ۲، ۳، ۵، ۷ و ۹ کاهش دهید. سپس دسته بند های Perceptron متفاوتی را بر روی هر یک از خروجی های مرحله قبل آموزش دهید. همچنین بایستی دقت های آموزشی و ارزیابی را با استفاده از روش ۵ Fold Validation-Cross Repeated با تکرار ۱۰ گزارش دهید. سپس نتایج را در Boxplot رسم کنید (از مجموعه داده تست استفاده نکنید).

(ج) قسمت ب را به کمک PCA انجام دهید و دادگان را به ابعاد ۲، ۳، ۵، ۷، ۹، ۲۵، ۵۰، ۱۰۰، ۲۰۰ و ۳۰۰ کاهش دهید.

(د) با توجه به Boxplot مدل ها را از دیدگاه Overfit و Underfit بررسی کنید و بهترین مدل LDA و PCA را گزارش کنید. آیا میتوان تنظیمات بهتری ارائه داد؟ توضیح دهید.

(ه) مدل با تنظیم بهینه را بر روی داده تست اجرا کنید، سپس دقت<sup>۱</sup> مدل را محاسبه کنید و همچنین ماتریس درهم ریختگی<sup>۲</sup> را رسم کنید.

[۵۰ نمره]

---

[بارم کل: ۱۰۰ نمره]

با آرزوی موفقیت