

# تمرین سری اول درس بازشناسایی آماری الگو

موعد تحویل: اول آبان ۱۳۸۶

تمرین:

فصل ۲: ۲، ۱۲، ۱۳، ۱۴، ۱۹، ۲۴.

شبیه‌سازی کامپیوتری:

- ۱- تمرین ۳۱ فصل ۲.
- ۲- تعداد ۱۰۰۰ نمونه از هر یک از سه کلاس گوسی به ترتیب با میانگین  $[0,0]^T$ ،  $[2,2]^T$ ،  $[-2,-2]^T$  و ماتریس کوواریانس  $\begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix}$ ،  $\begin{bmatrix} 0.5 & 0 \\ 0 & 1 \end{bmatrix}$ ،  $\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$  ایجاد کنید. توزیع نمونه‌ها را رسم کنید. پس از تخمین پارامترهای pdf هر کلاس (به روش ML)، طبقه‌بندی کننده‌ی بیز را با فرض احتمال یکسان کلاس‌ها پیاده‌سازی کنید. مقدار خطای طبقه‌بندی کننده را به ازای نمونه‌های شبیه‌سازی بدست آورید و آنرا با خطای تئوری مقایسه کنید.
- ۳- طبقه‌بندی کننده‌ی  $k$  نزدیکترین همسایه را پیاده‌سازی کنید و خطای آنرا به ازای نمونه‌های مساله قبل بدست آورید. خطای طبقه‌بندی به ازای مقادیر مختلف  $k$  (با شروع از مقدار ۱) را در جدولی بیاورید. اثر افزایش  $k$  بر مقدار خطا را توجیه کنید.
- ۴- تعداد ۱۰۰۰ نمونه از هر یک از دو کلاس گوسی  $N(0,1)$  و  $N(1,1)$  ایجاد کنید و کلاسیفایر بیز را بدون اطلاع از احتمال کلاس‌ها پیاده‌سازی کنید (minMax Classifier). مقدار خطای طبقه‌بندی را بر حسب مقادیر مختلف احتمال کلاس اول (از ۰ تا ۱) رسم کنید. حداکثر خطا به ازای کدام مقدار  $p(\omega_1)$  رخ می‌دهد؟
- ۵- ۱۴ نوع نویز صوتی (فایل ضمیمه) در اختیار شما قرار گرفته است.
  - a. با استفاده از روش Kernel تابع چگالی احتمال هر یک را تخمین زده و با نتیجه هیستوگرام عادی مقایسه کنید (رسم در یک نمودار)
  - b. برای هر یک از چهارده فایل، پارامترهای گوسی (متوسط و واریانس) را به روش ML تخمین زده و با یافتن معیاری عددی برای گوسی بودن، مشخص کنید، کدام نویز واقعاً گوسی است.