

الگوشناسی آماری (CE-725)

دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف

تمرینات سری پنجم - بهار ۱۳۸۹

به نکات زیر توجه فرمائید:

1. زمان تحویل تمرینات در سایت درس مشخص شده است. دقت نمائید که زمانبندی‌های تعیین شده قابل تغییر نیستند.
  2. تمرینات را با عنوان SPR-HWx-8xxxxxxx (مثلا SPR-HW5-88300785) و در یک فایل فشرده با همین نام به آدرس rahban@ce.sharif.edu ایمیل زده و در اولین جلسه بعد از زمان تحویل، بصورت پرینت شده تحویل استاد درس دهید.
  3. گزارش شما باید مختصر و مفید باشد. برای تمرینات پیاده‌سازی که با لوگوی شخص شده‌اند باید کد مطلب نوشته شده ضمیمه گزارش شده و تمامی خروجی‌های برنامه‌ها در گزارش شما ذکر شوند.
- سوال ۱) یک مجموعه از تصاویر ارقام دستنویس انگلیسی را در نظر بگیرید (مجموعه استاندارد USPS). هر تصویر به صورت یک ماتریس  $28 \times 28$  نشان داده می‌شود. می‌توان با سطری کردن، هر تصویر را به صورت یک بردار  $1 \times 784$  نمایش داد. هدف تشخیص رقم متناظر با یک تصویر است.
- با توجه به اینکه در این مسئله ۱۰ کلاس مختلف وجود دارد، از روش One-Against-All برای جداسازی استفاده می‌کنیم. همچنین قبل از اعمال داده‌ها به روش‌های یادگیری، از سفید کردن (Whitening) داده‌ها با ۵۰ بعد اول استفاده می‌کنیم. برای دسته بندی داده از دو روش استفاده و آنها را با هم مقایسه کنید :

الف) روش شبکه عصبی چندلایه

ب) روش SVM با هسته های خطی و RBF

در این پایگاه داده ها، ۶۰۰۰۰ تصویر وجود دارد.

برای هر روش آزمایش های زیر را ترتیب دهید :

(۱) تعداد داده آموزشی ۲۰۰۰ باشد

(۲) تعداد داده آموزشی ۱۰۰۰ باشد

(۳) تعداد داده آموزشی ۲۰۰ باشد

برای هر آزمایش از Cross-Validation (با مجموعه Validation 100 تایی) برای بدست آوردن C بهینه SVM و پهنای

باند بهینه SVM در حالت هسته RBF استفاده کنید.

هر آزمایش را ۲۰ بار تکرار کنید (با داده های آموزشی تصادفی).

الف) میانگین دقت هر روش را بدست آورید.

ب) Support Vector ها در SVM را نمایش دهید. آیا این داده با داده های دیگر یادگیری تفاوت خاصی دارند یا خیر.

ج) متوسط مقدار Margin را در هر حالت بدست آورید.

د) آیا استفاده از هسته RBF بهبودی نسبت به حالت خطی SVM ایجاد می کند؟

ضمیمه : فایل USPS.mat

متغیر Dt شامل داده ها که هر سطر آن یک داده را مشخص می کند

متغیر labels شامل برچسب متناظر با هر داده