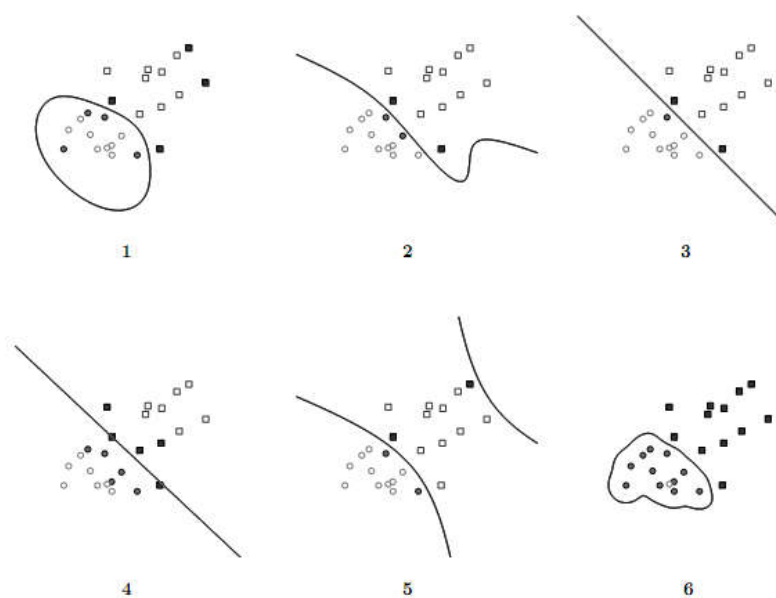


- طراح هریک از سوالات در زیر مشخص شده است. در صورتی که سوالی دارید می‌توانید سوالاتان را در گروه مطرح کنید یا از طراح سوال بپرسید.
-طراح سوال ۱ تا ۲: آقای حق دوست
-طراح سوال ۳ تا ۶: خانم محمودی‌نژاد

- ۱- استفاده از توابع Dirichlet ، Gamma و Rayleigh را به جای تبدیل گوسی در RBF بررسی کنید.
- ۲- فرض کنید در بحث soft-margin SVM تابع جریمه را از $\sum_{i=1}^m \epsilon_i$ به $\max_{i=1}^m \epsilon_i$ تغییر دهیم. رابطه بهینه‌سازی دوگان را برای این حالت نوشته و نشان دهید که این حالت تنها در ضرایب معادله با حالت اول تفاوت دارد.
- ۳- در تصاویر زیر به کمک الگوریتم SVM، دو کلاس که با مربع و دایره نشان داده شده، از هم جدا شده‌اند. اشکال توپر بردار پشتیبان را مشخص می‌کنند. در زیر ۵ عبارت بیان شده است. با دلیل بیان کنید که هر عبارت مربوط به کدامیک از تصاویر است.
الف) SVM خطی با حاشیه نرم و $C=0.1$
ب) SVM خطی با حاشیه نرم و $C=10$
ج) SVM با حاشیه سخت و تابع کرنل $K(u,v) = u \cdot v + (u \cdot v)^2$
د) SVM با حاشیه سخت و تابع کرنل $K(u,v) = \exp(-1/4 \|u-v\|^2)$
ه) SVM با حاشیه سخت و تابع کرنل $K(u,v) = \exp(-4 \|u-v\|^2)$



شکل مربوط به سوال سوم



۴- داده‌های ps3-data.mat به دودسته‌ی آموزشی و تست تقسیم شده است. در هر کدام، ستون آخر نشان دهنده‌ی کلاس مربوطه است. به کمک این داده‌ها یک شبکه‌ی RBF بسازید. از دستور newrbf متلب استفاده کنید و مقدار spread را {1,10,100} قرار دهید و نتیجه را گزارش کنید. توجه: چون خروجی باید باینری باشد، می‌توانید مقادیر خروجی کمتر از میانگین را صفر و بزرگ‌تر از میانگین را برابر یک قرار دهید.

۵- دو مجموعه‌ی داده‌ی data1.csv و data2.csv داده شده است. به کمک الگوریتم k-means هر کدام از این دو مجموعه داده را به دو خوشه تقسیم کنید. نتایج هر کدام را نمایش دهید (خروجی را به گونه‌ای نشان دهید که هر خوشه مشخص باشد). آیا این الگوریتم برای هر دو مجموعه داده خوب کار می‌کند؟ توضیح دهید.

۶- داده‌های مربوط به نوع ایمیل در فایل spambase.data آمده است که ستون آخر مشخص می‌کند هر ایمیل اسپم (۱) هست یا خیر (۰). به کمک الگوریتم SVM قصد داریم که نوع کلاس هر ایمیل مشخص کنیم. از ۷۰٪ داده‌ها برای آموزش و از مابقی برای تست استفاده کنید. الف) مقادیر مختلف $c = \{0.1, 0.4, 0.7, 1\}$ را تست کنید و دقت را برای هر کدام گزارش نمایید. ب) توابع کرنل مختلف خطی، چندجمله‌ای، درجه دو (quadratic)، mlp و rbf را امتحان کنید و نتیجه را گزارش کنید. توجه: در هر مرحله الگوریتم را ۵ بار اجرا کرده و در نهایت میانگین‌گیری کنید.