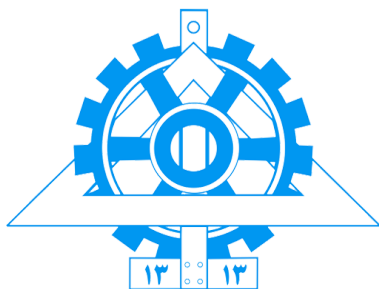


باسمه تعالی



یادگیری ماشین

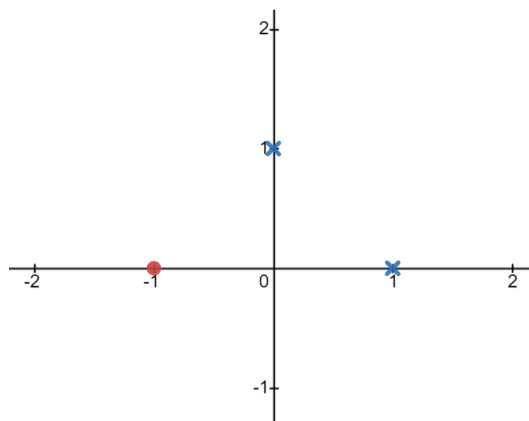
تمرین سری پنجم

آذر ۹۹

- لطفا در پاسخ‌گویی به سوالات این تمرین، نکات زیر را مدنظر قرار دهید.
- از اطناب پرهیز کنید و پاسخ‌ها را تا حد ممکن خلاصه و شفاف ذکر کنید. نکات نگارشی را رعایت و پاسخ هر سوال را در صفحه‌ی جداگانه تایپ کنید.
 - سوالات تحلیلی را تا حد امکان به صورت تایپ‌شده و فقط در قالب فایل pdf تحویل دهید. برای نگارش روابط ریاضی می‌توانید از روابط عکس گرفته و در فایل pdf قرار دهید.
 - کدهای مربوط به سوالات کامپیوتری را حتما ضمیمه کنید. گزارش و تحلیل این سوالات بدون انضمام کد نمره‌ای نخواهد داشت.
 - کدها فقط می‌توانند در زبان‌های پایتون باشند. تنها فرمت مورد قبول فرمت py می‌باشد. همچنین برای هر سوال (یا بخش)، فایل مربوط به آن سوال را جداگانه و با نام خود سوال ضمیمه کنید.
 - برای هر کد که در فایل نهایی ضمیمه می‌کنید، گزارش بنویسید. کدهای ضمیمه‌شده بدون گزارش مربوطه نمره‌ای نخواهند داشت. (این گزارش‌ها تنها معیار تفکیک کد شما و کدهای موجود در منابع مختلف مانند اینترنت خواهند بود).
 - عکس‌ها را به صورت واضح و همراه با زیرنویس در گزارش خود بیاورید.
 - فایل نهایی خود را در یک فایل زیپ شامل یک فایل pdf گزارشات و فایل کدهای خود آپلود کنید. نام فایل زیپ حتما الگوی ml-hw5-SID داشته باشد.
 - مجموع نمرات این تمرین ۲۵۰ نمره است که با انتخاب سوالات مناسب، شما حداکثر ۱۲۰ نمره‌ی آن را می‌توانید بگیرید. همچنین برای نمره‌ی کامل این تمرین نیز، ۱۰۰ نمره کافیت.
 - در صورت داشتن هرگونه سوال راجع به این تمرین، با ایمیل rahimiazghan@gmail.com در ارتباط باشید.

۱ سوال اول (۲۵ نمره)

داده‌های شکل ۱ را که به دو کلاس تقسیم شده‌اند را در نظر بگیرید. با نوشتن روابط موجود، معادله‌ی جداساز خطی را بر اساس SVM برای آن‌ها پیدا کنید. SV ها را نیز مشخص کنید.



شکل ۱

۲ سوال دوم (۲۵ نمره)

در مورد SVM به سوالات زیر پاسخ دهید.

۱. چرا به مسالهی ثانویه (dual problem) احتیاج داریم و از همان ابتدا سعی نمی‌کنیم به‌طور مستقیم مسالهی اولیه (primal problem) را حل کنیم؟

۲. تفاوت hard margin و soft margin در چیست؟ مثالی ساده را در فضای دوبعدی نشان دهید که استفاده از soft margin به‌جای hard margin، در ازای وجود تعدادی خطا، طبقه‌بند را «بهتر» می‌کند.

۳. کاربرد و مفهوم ضریب c در معادله‌ی مربوط با soft margin چیست و افزایش و کاهش آن چه تاثیری در طبقه‌بند خواهد داشت؟

$$\max \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n \xi_i$$

۳ سوال سوم (۲۵ نمره)

مفهوم کلی کرنل و دلایل استفاده از روش‌های مبتنی بر کرنل را بیان کنید و برای یک کرنل معتبر، اثبات کنید که داریم

$$K(x, y)^2 \leq K(x, x)K(y, y)$$

هم‌چنین فرض کنید مجموعه داده‌ی $S = \{x_q\}_{q=1}^Q$ داده شده است. تحت تبدیل ϕ این نقاط از فضای d_1 بعدی به فضای d_2 بعدی انتقال پیدا کرده‌اند. اگر میانگین نقاط در این فضای جدید و کرنل مربوط با تبدیل ϕ را به ترتیب با μ_ϕ و K_ϕ نشان دهیم، اثبات کنید داریم

$$\|\mu_\phi\| = \frac{1}{Q} \sqrt{\sum_{m=1}^Q \sum_{n=1}^Q K_\phi(x_m, x_n)}$$

۴ سوال چهارم (۲۵ نمره)

اگر کرنل‌های معتبر $K_1(x, y)$ و $K_2(x, y)$ را داشته باشیم، اثبات کنید کرنل‌های زیر همگی معتبر هستند.

۱. $K(x, y) = f(x)K_1(x, y)f(y)$ (که در آن f یک تابع دلخواه است).

۲. $K(x, y) = \exp(K_1(x, y))$

۳. $K(x, y) = K_1(x, y) + K_2(x, y)$

۴. $K(x, y) = K_1(x, y)K_2(x, y)$

۵. $K(x, y) = x^T A y$ که در آن ماتریس A یک ماتریس مثبت معین است.

۵ سوال پنجم (۲۵ نمره)

یک کرنل گاوسی را با رابطه‌ی $K(x, y) = \exp(-\frac{\|x - y\|^2}{2})$ را در نظر بگیرید. اگر با استفاده از تابع نگاشت مربوط با این کرنل، نقاط $x_1 = [1, 1]^T$ و $x_2 = [3, 4]^T$ را به فضای جدید انتقال دهیم، فاصله‌ی دو نقطه در این فضای جدید چقدر خواهد بود؟

حال کرنل $K(x, y) = (x^T y + 1)^2$ را در نظر بگیرید. اثبات کنید که فضای جدید ایجاد شده توسط تابع نگاشت این کرنل، یک فضای $\frac{(d+1)(d+2)}{2}$ بعدی است. d تعداد ابعاد فضای اولیه است.

۶ سوال ششم (۲۵ نمره)

به سوالات زیر پاسخ کوتاه دهید.

۱. تفاوت بین رویکردهای generative و discriminative را برای مسائل طبقه‌بندی توضیح دهید.

۲. مزایا و معایب رویکردهای one-vs-another، one-vs-rest، linear machine را نسبت به هم‌دیگر بیان کنید.

۳. در مسائل طبقه‌بندی، در چه شرایطی (یا به چه قصدی) ابعاد فضا را کاهش و در چه شرایطی افزایش می‌دهند. مثالی ساده را ذکر کنید که افزایش ابعاد فضای اصلی، منجر به نتایج «بهتری» برای مسئله‌ی طبقه‌بندی شده است.

۴. رابطه‌ی بین convex hull ساخته‌شده توسط داده‌های دو کلاس و ابرصفحه‌ی با بیش‌ترین حاشیه‌ی ایجاد شده را بیان کرده و با رسم مثالی ساده، این رابطه را نشان دهید.

۵. با نوشتن روابط موجود، از مساله‌ی اولیه در الگوریتم SVM، به مساله‌ی ثانویه در آن برسید.

۶. آیا با اضافه کردن مفهوم soft margin، در تعریف SVM تفاوتی (نسبت به hard margin) ایجاد خواهد شد؟

۷ سوال هفتم (۲۵ نمره)

برای این سوال محدود به استفاده از پکیج خاصی نیستید.

یک الگوریتم پرسپترون دولایه را تشکیل داده و تغییر پارامترهای زیر را در آن امتحان کنید. ابتدا توضیحی دقیق از هرکدام از پارامترها ارائه دهید. (مثلا برای پارامتر optimizer، تفاوت الگوریتم sgd و adam را توضیح دهید) سپس نتایج (اعم از دقت ترین و تست و ماتریس آشفتگی) را بر روی مجموعه داده‌ی Tiny-MNIST را گزارش کنید. (زمانی که یک پارامتر را تغییر می‌دهید، بقیه را پیش فرض خود پکیج استفاده‌شده بگذارید)

```
1 hidden_layer_size = "[100, 50]"
2 activation_function = "[relu, sigmoid]"
3 optimizer = "[adam, sgd]"
4 max_iteration = "[20, 30]"
5 early_stopping = "[True, False]"
```

۸ سوال هشتم (۲۵ نمره)

برای این سوال، مجاز به استفاده از هیچ پکیج مربوط با یادگیری ماشین نیستید.

برای فضای یک بعدی شکل ۲، با استفاده از کرنل $K(x, y) = (xy + 1)^2$ SV ها و هم چنین ضرایب α_i را برای دو کلاس بدست آورید.

دقت کنید که گزارش شما در این سوال، باید ربط جز به جز کد و روابط ریاضی موجود را نمایش دهد.



شکل ۲

۹ سوال نهم (۲۵ نمره)

در این سوال بر روی مجموعه داده‌ی iris کار خواهید کرد و محدودیتی نیز بر استفاده از هیچ پکیجی ندارید.

با استفاده از الگوریتم SVM و روش‌های زیر، برای هرکدام، ناحیه‌ی کلاس‌های مختلف را رسم کنید. از ویژگی‌های Petal Length و Petal Width استفاده کنید.

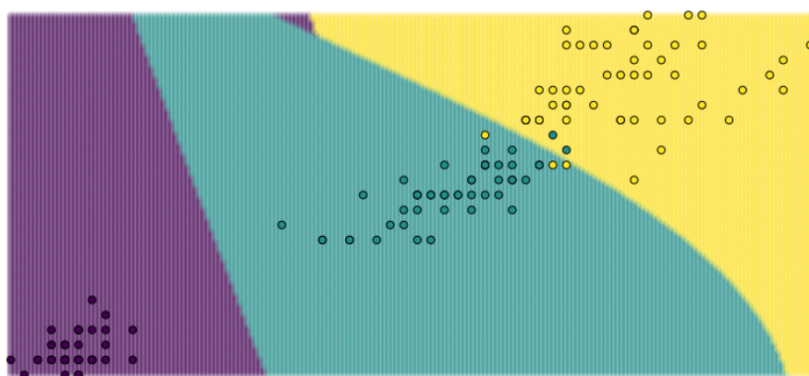
● SVM with linear kernel, one-vs-rest

● SVM with linear kernel, one-vs-one

● SVM with rbf kernel, one-vs-rest

● SVM with polynomial kernel ($d=3$), one-vs-rest

علاوه بر رسم ناحیه‌های مختلف هرکلاس، برای هر کدام مقادیر دقت ترین و ماتریس آشفستگی را نیز گزارش کنید. شکل‌های ضمیمه‌شده‌ی شما، باید چیزی شبیه به شکل ۳ باشد.



شکل ۳

۱۰ سوال دهم (۲۵ نمره)

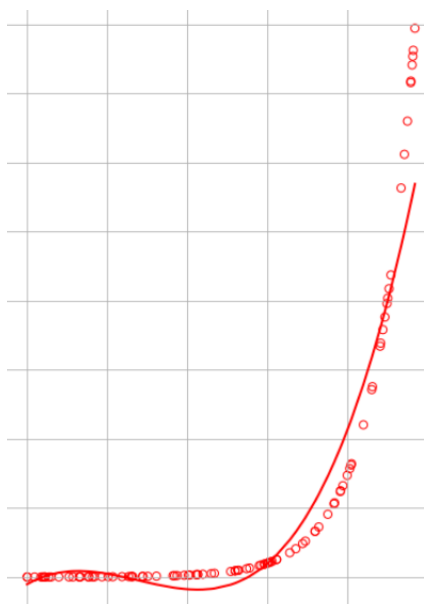
در این سوال، محدودیتی بر استفاده از پکیج ندارید.

۱.۱۰ الف

ابتدا تعدادی داده‌ی تصادفی با اندازه‌ی ۱۰۰ مابین ۰ تا ۱۰ تولید کنید. سپس با استفاده از Support Vector Regression و کرنل‌های خطی، درجه‌ی ۳ و rbf، تابع نمایی زیر را تخمین زده و نمودار تخمین زده شده توسط هر کرنل را همراه با نقاط اصلی نشان دهید. همچنین خطای تخمین زن خود را نیز گزارش کنید.

$$y = 5e^x + 3, x \in [0, 10]$$

نمودارهای به دست آمده‌ی شما چیزی شبیه شکل ۴ خواهد شد.

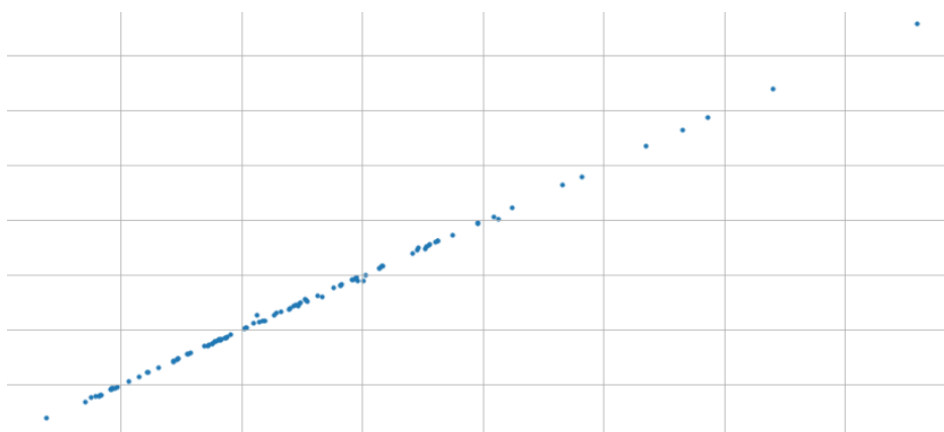


شکل ۴

۲.۱۰ ب

در این قسمت بر روی مجموعه داده‌ی Boston Housing کار خواهید کرد. این مجموعه داده را می‌توانید از این‌جا دانلود کنید. برای کسب اطلاعات راجع به این مجموعه داده نیز به این لینک می‌توانید مراجعه کنید.

کاری که در این قسمت می‌کنید، مشابه قسمت قبل است. با استفاده از این مجموعه داده و ویژگی‌های آن و الگوریتم Support Vector Regression و استفاده از کرنل‌های خطی، درجه‌ی ۳ و rbf قیمت خانه را تخمین بزنید. در انتها برای هر کرنل، نمودار قیمت واقعی و قیمت تخمین زده شده توسط الگوریتم را نمایش دهید. نمودارهای ضمیمه شده‌ی شما بایستی چیزی شبیه به شکل ۵ باشد. هم‌چنین خطای هر کرنل را نیز گزارش کنید.



شکل ۵