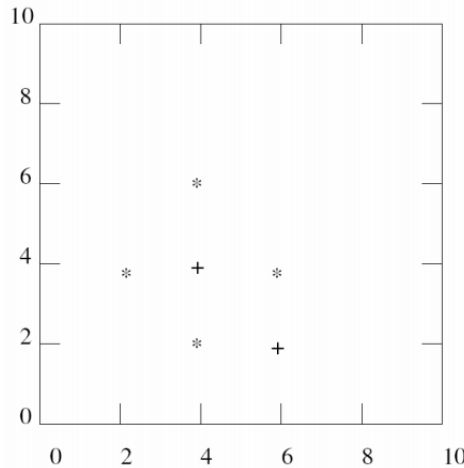


سوال ۱: k-NN (10 Points)

۱. (3 Points) مرز تصمیم دسته‌بند 1-NN را در مجموعه داده شکل زیر مشخص کنید.



۲. (7 Points) X, Y را دو متغیر تصادفی یکنواخت مستقل در بازه $[0, 1]$ در نظر بگیرید:

الف) میانگین و واریانس متغیر تصادفی $Z = (X - Y)^2$ را به دست آورید.

ب) با فرض d بعدی بودن X, Y میانگین و واریانس $R = \|X - Y\|^2$ را به دست آورید.

پ) پاسخ قسمت قبل را با حداکثر مربع فاصله اقلیدسی ممکن در d بعد مقایسه کنید. از این مقایسه چه نتیجه ای میتوان در مورد فاصله اقلیدسی نقاط در فضاهایی با ابعاد بالا گرفت؟

سوال ۲: Perceptron (10 Points)

۱. (3 Points) یک دسته بند پرسپترون با پارامترهای $(w_0, w_1, w_2)^T = (2, 1, 1)^T$ را در نظر بگیرید.

الف) مرز تصمیم این دسته بند را رسم کنید.

ب) کدام یک از پرسپترون های زیر مشابه قسمت الف فضا را دسته بندی میکنند؟

$$(w_0, w_1, w_2)^T = (1, 0.5, 0.5)^T$$

$$(w_0, w_1, w_2)^T = (2.0, 1.0, 1.0)^T$$

$$(w_0, w_1, w_2)^T = (\sqrt{2}, \sqrt{1}, \sqrt{1})^T$$

$$(w_0, w_1, w_2)^T = (-2, -1, -1)^T$$

۲. (7 Points) تابع هزینه پرسپترون به صورت زیر تعریف میشود که در آن M مجموعه‌ای از اندیس های نمونه های غلط دسته بندی شده است:

$$J(w) = - \sum_{i \in M} y^{(i)} w^T x^{(i)}$$

الف) نشان دهید تابع هزینه پرسپترون را میتوان به صورت زیر نوشت:

$$J(w) = \sum_{i=1}^n \max(0, -y^{(i)} w^T x^{(i)})$$

ب) تابع ضرر loss پرسپترون را برای $y = 0$ و $y = 1$ بر حسب $w^T x$ رسم کنید. تابع ضرر SVM چه مزیتی نسبت به تابع هزینه پرسپترون دارد؟
 پ) مسئله بهینه سازی تابع هزینه پرسپترون جواب به فرم بسته ندارد به همین دلیل از روش هایی مانند گرادیان نزولی Batch/Stochastic Gradient Descent برای به دست آوردن پارامتر های بهینه استفاده میشود. پس از به دست آوردن گرادیان تابع هزینه بر حسب w قوانین به روزرسانی وزن ها را در روش گرادیان نزولی batch و stochastic به دست آورید.

سوال ۳: SVM (30 + 5 Points)

۱. (12 Points) مسئله بهینه سازی Soft-Margin SVM به شکل زیر تعریف میشود:

$$\min_{w, b, \xi} \frac{1}{2} \|w\|^2 + C \sum_n \xi_n \quad (1)$$

$$s.t : y^{(n)}(w^T x^{(n)} + b) \geq 1 - \xi_n, \quad \forall n \quad (2)$$

$$\xi_n \geq 0, \quad \forall n \quad (3)$$

الف) در مسئله بهینه سازی Hard-Margin SVM نشان دهید با جایگذاری ۱ در محدودیت مسئله با هر عدد ثابت مثبت دیگر مسئله بهینه سازی تغییری نخواهد کرد.

ب) در رابطه (۱) ξ_n را بر حسب داده و پارامتر ها بازنویسی کرده و نمودار loss را رسم کنید.

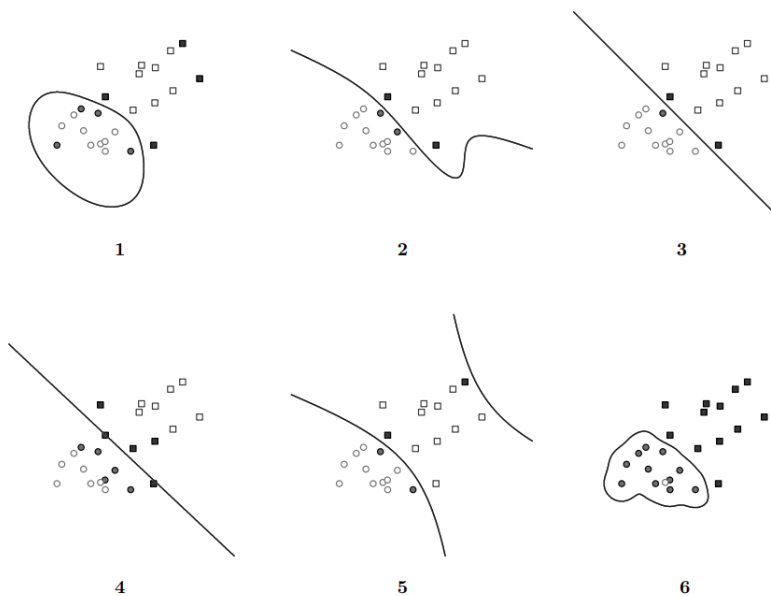
پ) با استفاده از ضرایب لاگرانژ مسئله بالا را بازنویسی کنید سپس بررسی کنید در صورت نقض شدن هر یک از محدودیت ها بهینه سازی چه جوابی خواهد داشت.

ت) از تابع لاگرانژ نسبت به ξ , b , w مشتق گرفته و با استفاده از آنها مسئله را به فرم دوگان بازنویسی کنید.

ث) فرم دوگان به دست آمده از قسمت قبل را به فرم ماتریسی [Quadratic Programming](#) بازنویسی کنید.

ج) فرم اولیه و دوگان مسئله را مقایسه کنید. وجود چه شرایطی باعث میشود که فرم دوگان ترجیح داده شود؟

۲. (10 Points) شکل زیر ۶ پلات مختلف دسته بندی SVM را نشان میدهد که در آن $y^{(n)} \in \{-1, +1\}$ به ترتیب با دایره و مربع و بردار های پشتیبان بصورت پر رنگ نشان داده شده اند:



هر یک از سناریو های زیر متعلق به کدام پلات است؟ دلیل خود را برای هر انتخاب بیان کنید:

الف) SVM حاشیه نرم با $C = 0.02$

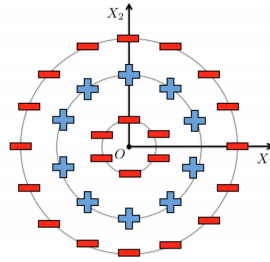
ب) SVM حاشیه نرم با $C = 20$

پ) SVM حاشیه سخت با کرنل $K(x, x') = x^T x' + (x^T x')^2$

ت) SVM حاشیه سخت با کرنل $K(x, x') = \exp(-5\|x - x'\|^2)$

ث) SVM حاشیه سخت با کرنل $K(x, x') = \exp(-\frac{1}{5}\|x - x'\|^2)$

۳. (8 Points) الف) برای داده های ۲ بعدی شکل زیر یک تبدیل ۱ بعدی معرفی کنید که در آن داده ها به صورت خطی جدا پذیر باشند.



ب) هر تابع کرنل را میتوان به صورت ضرب داخلی دو بردار نمایش داد. برای توابع کرنل زیر فرم برداری آنها را به دست آورید:

$$K_1(x, x') = (x^T x' + r)^2 \quad x, x' \in \mathbb{R}^2, \quad r \in \mathbb{R}$$

$$K_2(x, x') = \exp(-\|x - x'\|^2), \quad x, x' \in \mathbb{R}^d$$

۴. (5 Points) امتیازی) فرم مبتنی بر کرنل پرسپترون را به دست آورید. (راهنمایی: فرض کنید تعداد کل دفعات اشتباه دسته بندی شدن یک نمونه در طول آموزش به شما داده شده است)

سوال ۴: Practical Classification (50 + 10 Points)

در این سوال هدف دسته بندی داده های Ischemia که از حدود ۱۰۰۰ بیمار برای تعیین وضعیت Ischemia به دست آمده است می باشد. توجه: برای load کردن داده ها میتوانید از کتابخانه pandas استفاده کنید. برای پیاده سازی سایر قسمت ها فقط مجاز به استفاده از کتابخانه numpy هستید و نمیتوانید از کتابخانه های آماده یادگیری ماشین استفاده کنید. در بعضی سوالات نیاز به استفاده از کتابخانه های دیگری دارید که در متن همان سوال به آنها اشاره شده است.

برای ارزیابی و مقایسه روش های دسته بندی در سوالات زیر، در تمام موارد دقت دسته بندی، sensitivity, specificity, F-score را گزارش کنید.

۱. (10 points) برای $k = 1, 3, \dots, 9$ با استفاده از روش k-NN داده ها را دسته بندی کرده و بهترین k را بدست آورید (برای این کار از داده های Validation استفاده کنید). با استفاده از k ی بدست آمده، موارد گفته شده در بالا را روی داده های تست گزارش کنید.

۲. (10 points) دسته بند Perceptron را پیاده سازی کرده و سپس این دسته بند را روی داده های داده شده بکار بگیرید. در حین آموزش دقت دسته بند را روی داده های Validation نمایش دهید و در انتها نمودار این تغییرات را رسم کنید (برای این کار میتوانید از کتابخانه matplotlib استفاده نمایید). موارد گفته شده در بالا و confusion matrix را روی داده های تست گزارش کنید.

۳. (5 points) مراحل سوال قبل را با استفاده از Voted Perceptron تکرار کرده و نتایج را با قسمت قبل مقایسه کنید.

۴. در فرم دوگان مسئله SVM داریم:

$$\min_{\alpha} -L(\alpha) = \frac{1}{2} \sum_n \sum_m \alpha_n \alpha_m y_n y_m K(x_n, x_m) - \sum_n \alpha_n$$

$$\text{subject to : } 0 \leq \alpha_n \leq C \quad \forall n$$

۱.۴. (20 points) مقدار C را برابر ۱ در نظر بگیرید و مسئله بهینه سازی بالا را با استفاده از کتابخانه های حل مسائل quadratic programming حل کنید. (به عنوان مثال میتوانید از کتابخانه cvxopt استفاده کنید). سپس با استفاده از α های بدست آمده، وزن های دسته بند SVM را بدست آورید و داده ها را دسته بندی کرده و معیارهای خواسته شده در بالا و precision را گزارش کنید.

۲.۴. (5 points) قسمت قبل را با استفاده از کرنل RBF انجام داده و نتایج را مقایسه کنید.

۳.۴. (10 points) امتیازی) برای حل مسئله بهینه سازی فرم دوگان SVM میتوان از روش gradient descent استفاده کرد. (دقت کنید که مسئله دارای شرط روی متغیرها است بنابراین نیاز است که کمی الگوریتم را تغییر دهید. برای ارضای شروط میتوانید از الگوریتم projected gradient استفاده کنید که در بخش ۶.۱۱ کتاب به آن اشاره شده است.)