



لطفاً به نکات زیر توجه بفرمایید

الف) استفاده از همه منابع، حتی مشورت با سایرین (به جز دستیاران آموزشی) در امتحان مجاز است.

ب) در صورتی که از مشورت دیگران استفاده می‌کنید؛ پاسخ سوال باید با بیان خود شما در برگه‌ی پاسخ نوشته شود.  
پاسخ‌های عیناً مشابه مورد قبول نیست.

ج) در ابتدای پاسخ به هر سوال نام افرادی که با آنها مشورت کرده‌اید را بنویسید (حداکثر ۵ نفر)

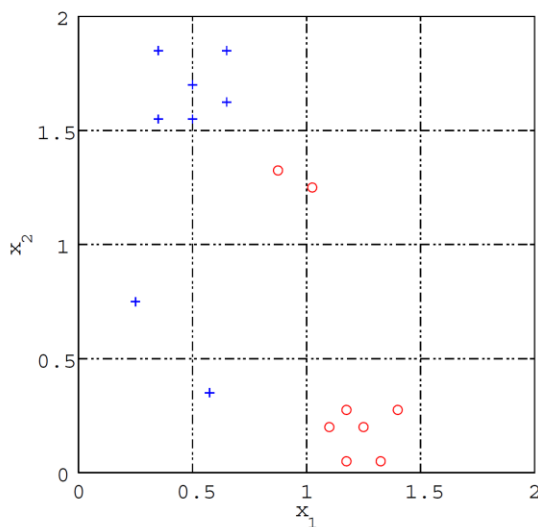
د) ۳۰ امتیاز از مجموع امتیازات امتحان به عنوان نمره تشویقی در نظر گرفته شده است.

۱. داده‌های یادگیری و تست یکسانی در اختیار دو تیم طراحی قرار گرفته است آنها دو طبقه‌بندی مختلف ساخته اند که عملکرد کاملاً یکسانی از نظر ماتریس کانفیوژن روی داده های یادگیری و تست دارند. حال شروع به استفاده از دو طبقه بند در شرایط کاری حقیقی می‌کنیم، اما مشاهده می‌شود که خطای طبقه‌بندی آن‌ها به نحو محسوسی با هم متفاوت است و یکی بهتر از دیگری عمل می‌کند. دلایل مختلفی که می‌تواند موجب این امر شده باشد را برشمارید و راجع به هر کدام از دلایلی که به ذهنتان می‌رسد مختصراً بحث کنید. (۲۵ امتیاز)

۲. یک تیم حرفه‌ای صنعتی طبقه بندی را ساخته اند و بعد از انجام تمام تست ها و ارزیابی های لازم طبقه بند را در شرایط عملیاتی زیر بار قرار داده اند. ابتدا همه چیز خوب است اما بعد از ۲ ماه خطای طبقه بندی شروع به بالا رفتن میکند. الف) به نظرتان چه دلایلی می‌تواند موجب این امر شده باشد؟ آنها را برشمارید و مختصراً درباره‌شان بحث کنید. ب) فرض کنید می‌خواهید بدون آنکه از شرایط عملیاتی خارج شوید طبقه بندی خود را مجدداً تنظیم کنید تا به عملکرد سابقش برگردد. از میان طبقه بندهایی که در طول درس با آنها آشنا شده اید به نظرتان کدام بیشتر به درد این کار می‌خورد؟ درباره روند تنظیم طبقه بند در حالی که زیر بار است صحبت کنید. (۳۰ امتیاز)

۳. قصد دارید یک سیستم طبقه بندی برای انواع سیب چیده شده به سه رده کیفی خوب، متوسط و بد بسازید. سیب‌ها به ترتیب برای فروش در مغازه های سطح شهر، آب گیری و خوراک دام استفاده خواهند شد. برای این کار یک خط نقاله دارید که سیب‌ها تک تک و به فاصله تقریبی یک متر از هم به روی آن حرکت می کنند. بعلاوه یک دوربین نور مرئی و یک کامپیوتر در اختیار دارید. الف) به کمک ابزارهای مذکور طبقه بند را طراحی نمایید. ب) یک بلوک دیاگرام از تمام مراحل طبقه بندی بکشید. درباره هر مرحله مختصراً توضیح دهید. خصوصاً سعی کنید درباره الزامات و محدودیت های طراحی در هر بلوک بحث کنید. ج) مسئله را در دو حالت بررسی کنید وقتی هیچ محدودیت بودجه در خرید تجهیزات سه گانه فوق ندارید و وقتی محدودیت بودجه دارید. (۳۵ امتیاز)

۴. شکل زیر نمونه‌های مربوط به دو کلاس + و O را در یک فضای دو بعدی نشان می‌دهد. به روش GMM و به کمک دو تابع گوسی که ماتریس کواریانس آن‌ها identity است ( $\Sigma=I$ )، داده‌ها را مدل کرده‌ایم. فراوانی نسبی کلاس‌ها و میانگین توابع گوسی به عنوان پارامترهای مدل در نظر گرفته شده‌اند. (۲۵ امتیاز)



الف) تخمین درست‌نمایی بیشینه برای میانگین هر کدام از گوسی‌ها را به صورت جداگانه در شکل نمایش دهید. ب) چنانچه از این مدل برای طراحی طبقه‌بند استفاده کنیم، مرز تصمیم را در همان شکل قسمت الف رسم کنید. ج) اگر از این مدل برای تولید داده جدید استفاده کنیم، احتمال تولید داده متعلق به کلاس O چیست؟

د) چه تعداد نمونه اشتباه دسته بندی می‌شوند؟

ه) فرض کنید بخواهیم این دو دسته از داده‌ها را به کمک الگوریتم  $k\_means$  خوشه بندی کنیم ( $k=2$ ). نقاط اولیه را به صورت دلخواه فرض کرده و گام‌های مختلف الگوریتم را با رسم شکل‌های مختلف نشان دهید. جواب خود با روش GMM بیان شده در قسمت‌های قبل مقایسه کنید.

۵. الف) در برخی از موارد می‌توان VC dimension را متناسب با تعداد پارامترهایی که اعضای یک کلاس فرضیه ( $H$ : hypothesis class) را توصیف می‌کند، در نظر گرفت. در این راستا VC dimension مربوط به مجموعه فرضیه زیر (توابع چند جمله‌ای) را بیابید. (۱۰ امتیاز)

$$h_{\alpha}(x) = \begin{cases} 1 & \text{if } \alpha_d x^d + \alpha_{d-1} x^{d-1} + \dots + \alpha_0 > 0 \\ 0 & \text{otherwise} \end{cases}$$

ب) مشخصاً نتیجه قسمت الف در همه شرایط برقرار نیست. کلاس فرضیه‌ای را نشان دهید که به ازای هر عدد صحیح  $m$ ، این کلاس فرضیه  $m$  پارامتر داشته باشد اما VC dimension آن ۱ باشد (راهنمایی: به این فکر کنید که چند پارامتر به کمک یک پارامتر بازنمایی می‌شود) (۱۰ امتیاز)

ج) کلاس فرضیه زیر را در نظر بگیرید

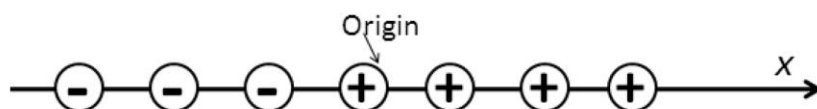
$$h_{\alpha}(x) = \begin{cases} 1 & \text{if } \sin(\alpha x) > 0 \\ 0 & \text{otherwise} \end{cases}$$

نشان دهید این کلاس فرضیه با یک پارامتر، VC dimension نامتناهی دارد. (۱۵ امتیاز)

۶. داده‌های شکل زیر را در فضای یک بعدی، در نظر بگیرید. نقاط مثبت  $\{0,1,2,3\}$  و منفی  $\{-3,-2,-1\}$  دو کلاس مختلف را نشان می‌دهند. طبقه‌بند خطی SVM با حاشیه نرم (soft-margin linear SVM) را بر روی داده‌های ارائه شده آموزش می‌دهیم. این طبقه‌بند به صورت زیر ارائه می‌شود که در آن پارامتر  $C$  یک پارامتر تنظیمی، جهت ایجاد تعادل در اندازه حاشیه ( $w^T w$  کمتر) و میزان تعدی از حاشیه ( $\sum_{i=1}^m \epsilon_i$ ) است.

$$\operatorname{argmin}_{\{w,b\}} \frac{1}{2} w^t w + C \sum_{i=1}^m \epsilon_i$$

Subject to :  $y_i(w^t x_i + b) \geq 1 - \epsilon_i$   
 $\epsilon_i \geq 0 \quad \forall i$



الف) تعداد بردارهای پشتیبان (support vector) را در دو حالت زیر با ذکر دلیل بیان کنید: (۲۰ امتیاز)

الف-۱) اگر  $C = 0$  باشد، یعنی تنها اندازه حاشیه برای ما مهم است

الف-۲) اگر  $C \rightarrow \infty$ ، یعنی تنها تعدی از حاشیه برای ما مهم باشد.

ب) اگر از روش leave-one-out (LOO) جهت اندازه گیری میزان خطا استفاده کرده باشیم. توضیح دهید چرا رابطه زیر با فرض  $C$  مشخص می‌تواند کران بالای خطای طبقه بند را توصیف کند. (۱۵ امتیاز)

$$\text{LOO error} \leq \frac{\#SVs}{l},$$

در رابطه فوق  $l$  سایز داده‌های آموزش و  $\#SV$  تعداد بردارهای پشتیبان است.

۷. مجموعه داده‌های زیر برای دو کلاس + و 0 در نظر بگیرد. طبقه‌بندهای زیر را از نظر میزان خطا و زمان اجرای الگوریتم با یکدیگر مقایسه کنید و بر اساس این دو معیار مرتب کنید. (۳۰ امتیاز)

الف) logistic regression

ب) SVM (quadratic kernel)

ج) 3-KNN

$$\begin{array}{cc} \bigcirc & + \\ + & \bigcirc \end{array}$$

۸. یک طبقه‌بندی دو کلاسه را در نظر بگیرید.  $\mathcal{Y}$  بردار ویژگی یک بعدی است و رابطه زیر، ماتریس پراکندگی کلی بین گروهی را محاسبه می‌کند.

$$J_1 = \frac{1}{n_1 n_2} \sum_{y_i \in Y_1} \sum_{y_j \in Y_2} (y_i - y_j)^2$$

الف) نشان دهید که رابطه زیر، یک بیان دیگر از رابطه ماتریس پراکندگی بالا می‌باشد:

$$J_1 = (m_1 - m_2)^2 + \frac{1}{n_1} s_1^2 + \frac{1}{n_2} s_2^2$$

که در آن  $m_1$  و  $m_2$  میانگین و مقادیر  $s_1^2$  و  $s_2^2$  نیز به ترتیب بیانگر پراکندگی  $Y_1$  و  $Y_2$  می‌باشند. (۱۰ امتیاز)

ب) از طرفی مجموع واریانس‌ها از رابطه زیر محاسبه می‌شوند:

$$J_2 = \frac{1}{n_1} s_1^2 + \frac{1}{n_2} s_2^2$$

حال، فرض کنید  $\mathcal{Y} = \mathbf{w}^t \mathbf{x}$  باشد؛ در صورتی که  $\mathbf{w}$  مقدار  $J_1$  را تحت شرط  $J_2 = 1$  بهینه کند، نشان دهید رابطه زیر برای  $\mathbf{w}$  برقرار است. (۲۵ امتیاز)

$$\mathbf{w} = \lambda \left( \frac{1}{n_1} \mathbf{S}_1 + \frac{1}{n_1} \mathbf{S}_2 \right)^{-1} (\mathbf{m}_1 - \mathbf{m}_2)$$

که در رابطه فوق داریم:

$$\lambda = \left[ (\mathbf{m}_1 - \mathbf{m}_2)^t \left( \frac{1}{n_1} \mathbf{S}_1 + \frac{1}{n_1} \mathbf{S}_2 \right)^{-1} (\mathbf{m}_1 - \mathbf{m}_2) \right]^{\frac{1}{2}}$$

$$\mathbf{m}_i = \frac{1}{n_i} \sum_{\mathbf{x} \in X_i} \mathbf{x} \qquad \text{and} \qquad \mathbf{S}_i = \sum_{\mathbf{x} \in X_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^t$$