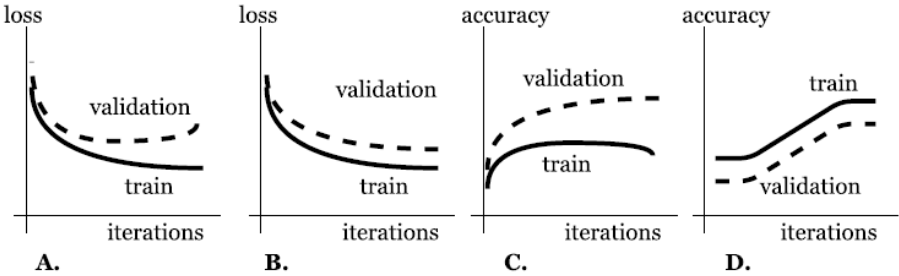
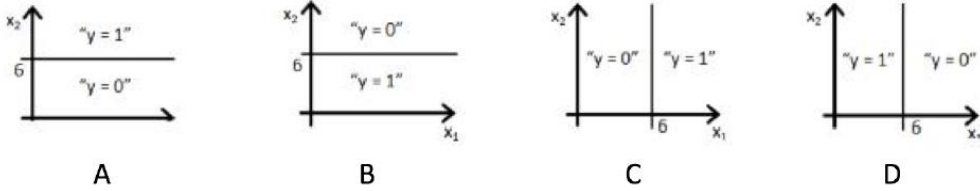




لطفا به نکات زیر توجه کنید.

- (۱) مدت زمان امتحان ۲ ساعت و ۳۰ دقیقه است.
- (۲) هیچ گونه مشارکتی در امتحان مجاز نیست.
- (۳) برای دریافت نمره کامل باید به تمام سوالات به غیر از سوال ۴ پاسخ کامل دهید. در صورت پاسخ به سوال ۴، نمره کسب شده از این سوال به عنوان نمره امتیازی به شما تعلق خواهد گرفت.

ردیف	سوال
۱ (۱۰)	<p>کدام نمودار نشان دهنده $overfitting$ در مسائل واقعی است. (ممکن است چند جواب وجود داشته باشد).</p> <div></div> <p>A. B. C. D.</p>
۲ (۱۰)	<p>فرض کنید یک طبقه‌بند $logistic\ regression$ را آموزش می‌دهید که تابع فرض آموزش دیده (learned hypothesis function) در آن بصورت $h_{\theta}(x) = \sigma(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$ است که داریم $\theta_0 = 6, \theta_1 = 0, \theta_2 = -1$ است. کدام یک از شکل‌های زیر نشان دهنده مرز تصمیم برای $h_{\theta}(x)$ است.</p> <div></div> <p>A B C D</p>
۳ (۱۵)	<p>به سوالات زیر پاسخ کوتاه دهید.</p> <p>(الف) یک مسئله طبقه‌بندی دو کلاسه را در نظر بگیرید. برای محاسبه ریسک به کدام یک از موارد زیر نیاز است.</p> <ol style="list-style-type: none">(۱) نمونه‌های مشاهده شده(۲) احتمالات پیشین کلاس‌ها(۳) تابع هزینه (loss function)(۴) محاسبه گرادیان کاهشی <p>(ب) جملات زیر درباره $logistic\ regression$ درست است یا خیر.</p> <ol style="list-style-type: none">(۱) تابع هزینه در آن $convex$ است.(۲) فرض می‌کنیم مشاهدات هر کلاس از توزیع گوسی پیروی می‌کنند.

	پ) تفاوت دو روش Stochastic Gradient Descent و Batch Gradient Descent را شرح دهید.
۴ (۲۰)	روابط موجود در مدل های GMM را در چهارچوب ساختار بیز بیان کنید و تعابیر معادل احتمال پیشین و پسین را بیان کنید.
۵ (۳۰)	<p>یک مسئله طبقه‌بندی دو کلاسه را در نظر بگیرید. چگالی شرطی مشاهده x برای هر کلاس به صورت زیر است.</p> $p(x w_1) = k_1 e^{-\frac{x^2}{20}}$ $p(x w_2) = k_2 e^{-\frac{(x-6)^2}{12}}$ <p>الف) مقادیر k_1 و k_2 را بدست آورید.</p> <p>ب) فرض کنید احتمالات پیشین دو کلاس برابر هستند. هزینه انتخاب درست صفر است. اما هزینه انتخاب نادرست بصورت $\lambda_{12} = \sqrt{3}$ و $\lambda_{21} = \sqrt{5}$ (هزینه λ_{ij} به معنی انتخاب کلاس i برای نمونه است در صورتی که واقعا متعلق به کلاس j است) است. عبارتی برای ریسک مشروط (conditional risk) بدست آورید.</p> <p>پ) مرزهای تصمیم را با توجه به ریسک تعریف شده در بخش (ب) بدست آورید.</p>
۶ (۳۰)	<p>دو توزیع احتمال دخواه f_1 و f_2 را در نظر بگیرید. در این سوال می‌خواهیم فرمولی برای مخلوط این دو توزیع (mixture) بدست آوریم. به بیان دقیق‌تر فرض کنید</p> $f_\lambda(x) = \lambda f_1(x) + (1 - \lambda) f_2(x)$ <p>که در آن f_1 و f_2 توابع چگالی احتمال دلخواه هستند و λ پارامتری نامعلوم (mixture parameter) است.</p> <p>الف) با فرض داشتن یک مشاهده x و معلوم بودن پارامتر λ احتمال اینکه مشاهده x از توزیع f_1 آمده باشد را حساب کنید.</p> <p>ب) فرض کنید اکنون مجموعه‌ای از مشاهدات به صورت $\{x_1, \dots, x_n\}$ و به ازای هر مشاهده یک مقدار c_i داریم. مشاهدات بصورت <i>i.i.d</i> از توزیع مخلوط بدست آمده اند. اگر مقدار c_i برابر ۱ باشد یعنی مشاهده x_i از توزیع f_1 آمده است و اگر مقدار c_i برابر ۰ باشد یعنی مشاهده x_i از توزیع f_2 آمده است. با فرض مشخص بودن λ عبارتی برای مقدار log-likelihood مشاهدات یعنی $\log P(x_1, c_1, \dots, x_n, c_n \lambda)$ بدست آورید.</p> <p>پ) حال فرض کنید که دیگر مشاهدات c_i را نداریم. به عبارت دیگر نمی‌دانیم هر مشاهده x_i از کدام توزیع آمده است. با استفاده روابطی که در قسمت های قبل بدست آورده اید، E-Step و M-Step در الگوریتم EM را برای تخمین پارامتر λ بدست آورید.</p>
۷ (۱۵)	<p>فرض کنید از شما به عنوان یک متخصص یادگیری ماشین در حل مسئله طبقه‌بندی سه مدل سیب کمک می‌خواهند. ابزار در اختیار شما یک دوربین و سخت افزاری جهت اجرای الگوریتم نهایی است. در گام توسعه محدودیتی از جهت هزینه ندارید اما در فاز اجرا محدودیت بسیار زیاد است و باید هزینه هر دستگاه (مجموع دوربین و سخت افزار اجرای الگوریتم) کمینه شود. شما چه روشی را برای حل این مسئله پیش خواهید گرفت. لطفا صرفا به مسئله محدودیت هزینه اشاره شود و از بیان موارد کلی پرهیز کنید.</p>
اگر دانشی مردراند سخن / تو بشو که دانش نکر دو کهن	