سلام بر تمام دانشجویان عزیز، چند نکته مهم:

۱. حجم گزارش به هیچ عنوان معیار نمره‌دهی نیست، در حد نیاز توضیح دهید.
۲. در این تمرین شما حداکثر ۱۳۰ نمره میتوانید کسب کنید بنابراین از پاسخ دادن به سوالات با مجموع بارم بیشتر از ۱۳۰ نمره خودداری کنید.
۳. حداقل به سه سوال شبیه‌سازی باید پاسخ دهید.
۴. برای سوالات شبیه سازی، فقط از دیتاست داده شده استفاده از کنید. شکل ها، به طور واضح و در فرمت درست گزارش شود.
۵. در سوال ۱۰ از دیتاست FACES و در سوال ۹ تا ۱۳ از دیتاست TinyMNIST استفاده کنید. در صورت زمانبر بودن پردازش کد در مورد سوالات ۹، ۱۲ و ۱۳ می‌توانید از TinyMNIST_loader که پیوست شده نیز استفاده نمایید. اما دقت بفرمایید که در مورد سوالات ۱۰ و ۱۱ مجاز به کاهش تعداد ویژگی‌ها/ تعداد داده‌ها نیستید.
۶. هرگونه شباهت در گزارش و کد مربوط به شبیه سازی، به منزله تقلب می باشد و کل نمره تمرین صفر می‌شود.
۷. در صورت داشتن سوال، از طریق ایمیل afsaneh.h.ebrahimi@gmail.com، سوال خود را مطرح کنید.

---

۱. فرض کنید که $p_x(x|w_i)$ بیانگر چگالیهای احتمالی با میانگین μ_i و واریانس Σ_i برای $i = ۱, ۲$ باشد که لزوما هم نرمال نیست. $y = w^t x$ بیانگر projection باشد و چگالیهای تک بعدی $p(y|w_i)$ دارای میانگین μ_i و واریانس $\delta_i^۲$ هستند. نشان دهید که تابع هزینه‌ی J(w)

$$J(w) = \frac{(\mu_۱ - \mu_۲)^۲}{\delta_۱^۲ + \delta_۲^۲}$$

به وسیله‌ی $w = (\Sigma_۱ + \Sigma_۲)^{-۱}(\mu_۱ - \mu_۲)$ بیشینه میگردد.(۲۰ نمره)

---

۲. در مساله‌ی طبقه‌بندی c کلاسه، ماتریس پراکندگی درون کلاسی و بین کلاسی به ترتیب به صورت زیر تعریف میشود:

$$S_w = \sum_{k=۱}^{c} \sum_{x^q \in w_i} (x^q - \mu_k)(x^q - \mu_k)^T$$

$$S_B = \sum_{k=۱}^{c} N_k (\mu_k - \mu)(\mu_k - \mu)^T$$

أ. نشان دهید که $rank(S_B) \leq C - ۱$ و در چه شرایطی $rank(S_B) = C - ۱$؟(۵ نمره)

ب. درباره‌ی حداکثر تعداد مقادیر ویژه‌ی ناصفر ماتریس جداپذیری $S_w^{-۱} S_B$ بحث نمایید.(۵ نمره)

ت. نشان دهید که $S_T = S_W + S_B$. (۱۰ نمره)

---

۳. عبارت

$$J = \frac{۱}{n_۱ n_۲} \sum_{y_i \in Y_۱} \sum_{y_j \in Y_۲} (y_i - y_j)^۲$$

پراکندگی کل درون گروهی(within group scatter) را اندازه میگیرد نشان دهید که این عبارت را میتوان به صورت زیر هم نوشت: (۲۰ نمره)

$$J = (m_۱ - m_۲)^۲ + \frac{۱}{n_۱} s_۱^۲ + \frac{۱}{n_۲} s_۲^۲$$

---

۴. فرض کنید که پنجره‌ی پارزن به صورت زیر تعریف شده باشد:

$$\varphi(x) = \begin{cases} e^{-x} & x > ۰ \\ ۰ & o.w \end{cases}$$

و همچنین $p(x) \sim U(۰, a)$ باشد. نشان دهید که میتوان میانگین پنجره پارزن را با n نمونه به صورت زیر تخمین زد: (۲۰ نمره)

$$\overline{p_n}(x) = \begin{cases} ۰ & x < ۰ \\ \frac{۱}{a}\left(۱ - e^{-\frac{x}{h_n}}\right) & ۰ \le x \le a \\ \frac{۱}{a}\left(e^{\frac{a}{h_n}} - ۱\right) e^{-\frac{x}{h_n}} & a \le x \end{cases}$$

---

۵. یک رستوران بر این است که بررسی نماید با توجه به عوامل موثر، افرادی که به رستوران مراجعه میکنند در صورتی که تمام میزها پر باشد آیا برای خالی شدن میز صبر میکند یا نه؟

داده‌های ثبت شده از ۱۲ مراجعه کننده، جنبه‌های مختلف و اینکه صبر میکنند/نمیکنند را در جدول ۱ مشاهده میفرمایید.

Example	Input Attributes										Goal
	Alt	*Bar*	*Fri*	*Hun*	*Pat*	*Price*	*Rain*	*Res*	*Type*	*Est*	*WillWait*
$\mathbf{x}_1$	*Yes*	*No*	*No*	*Yes*	*Some*	*$$$*	*No*	*Yes*	*French*	*0–10*	$y_1 = Yes$
$\mathbf{x}_2$	*Yes*	*No*	*No*	*Yes*	*Full*	*$*	*No*	*No*	*Thai*	*30–60*	$y_2 = No$
$\mathbf{x}_3$	*No*	*Yes*	*No*	*No*	*Some*	*$*	*No*	*No*	*Burger*	*0–10*	$y_3 = Yes$
$\mathbf{x}_4$	*Yes*	*No*	*Yes*	*Yes*	*Full*	*$*	*Yes*	*No*	*Thai*	*10–30*	$y_4 = Yes$
$\mathbf{x}_5$	*Yes*	*No*	*Yes*	*No*	*Full*	*$$$*	*No*	*Yes*	*French*	*>60*	$y_5 = No$
$\mathbf{x}_6$	*No*	*Yes*	*No*	*Yes*	*Some*	*$$*	*Yes*	*Yes*	*Italian*	*0–10*	$y_6 = Yes$
$\mathbf{x}_7$	*No*	*Yes*	*No*	*No*	*None*	*$*	*Yes*	*No*	*Burger*	*0–10*	$y_7 = No$
$\mathbf{x}_8$	*No*	*No*	*No*	*Yes*	*Some*	*$$*	*Yes*	*Yes*	*Thai*	*0–10*	$y_8 = Yes$
$\mathbf{x}_9$	*No*	*Yes*	*Yes*	*No*	*Full*	*$*	*Yes*	*No*	*Burger*	*>60*	$y_9 = No$
$\mathbf{x}_{10}$	*Yes*	*Yes*	*Yes*	*Yes*	*Full*	*$$$*	*No*	*Yes*	*Italian*	*10–30*	$y_{10} = No$
$\mathbf{x}_{11}$	*No*	*No*	*No*	*No*	*None*	*$*	*No*	*No*	*Thai*	*0–10*	$y_{11} = No$
$\mathbf{x}_{12}$	*Yes*	*Yes*	*Yes*	*Yes*	*Full*	*$*	*No*	*No*	*Burger*	*30–60*	$y_{12} = Yes$

جدول ۱-داده‌های ثبت شده از ۱۲ مراجعه کننده

توضیح جنبه‌های مختلف که در Input attributes آمده است به شرح زیر است:

1.	Alternate: whether there is a suitable alternative restaurant nearby.
2.	Bar: whether the restaurant has a comfortable bar area to wait in.
3.	Fri/Sat: true on Fridays and Saturdays.
4.	Hungry: whether we are hungry.
5.	Patrons: how many people are in the restaurant (values are None, Some, and Full).
6.	Price: the restaurant's price range ($, $$, $$$).
7.	Raining: whether it is raining outside.
8.	Reservation: whether we made a reservation.
9.	Type: the kind of restaurant (French, Italian, Thai or Burger).
10.	WaitEstimate: the wait estimated by the host (0-10 minutes, 10-30, 30-60, >60).

حال برای این مساله decision tree را رسم نمایید و با توجه به داده‌های موجود قضاوت نمایید که آیا مراجعه کننده در صورت پر بودن میزها صبر میکند یا خیر.(۲۰ نمره)

۶.

در مورد روش Adaboost مطالعه بفرمایید.

به عنوان مثال میتوانید به لینک زیر مراجعه نمایید:

http://www.cs.man.ac.uk/~nikolaon/~nikolaon_files/Introduction_to_AdaBoost.pdf

أ. به طور کلی در مورد این روش و نقاط قوت/ضعف آن نسبت به decision tree توضیح مختصری بدهید.(۵ نمره)

ب. فرض کنید $h_t(x)$ یک طبقه بند ضعیف باشد که در مرحله t مشاهده شده است و α_t بیانگر وزنها باشد. طبقه بند نهایی به شکل زیر است:

$$H(x) = sign\big(f(x)\big),\ \ f(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$$

نشان دهید که برای خطای داده‌های اموزشی طبقه بند نهایی رابطه زیر برقرار است: (y_i لیبل حقیقی برای x_i است.) (۱۵ نمره)

$$\frac{1}{m}\sum_{i=1}^{m}(H(x_i) \neq y_i) \leq \frac{1}{m}\sum_{i=1}^{m}\exp(-f(x_i)y_i)$$

۷.

أ. برای نمونه‌های $X = \{-7, -5, -4, -3, -2, 0, 2, 3, 4, 5, 7\}$ دریک مساله یک بعدی، تخمین پنجره‌ی پارزن $P_j(x)$ را برای یک پنجره‌ی مستطیلی بیابید. ضمنا $h_j = \frac{1}{\sqrt{j}}$ است. پنجره را برای j=1, 4, 11 در یک شکل رسم نمایید. (۱۰ نمره)

ب. در رابطه $h_j = \frac{h}{\sqrt{j}}$ ، h را ثابت بگیرید و درباره شکل $P_j(x)$ به ازای hهای مختلف بحث کنید. (۱۰ نمره)

---

۸. یکی از تکنیکهای متداول در فشرده‌سازی تصاویر PCA میباشد که تعداد Principle Components(PCs) در کیفیت تصویر و نرخ فشرده سازی تصویر تاثیرگذار است. حال ما برآنیم که با استفاده از PCA تصاویر را فشرده کنیم و به فضایی با تعداد ویژگیهای کمتر منتقل شویم تا عملیات تشخیص تصویر بهتر انجام شود. در این سوال شما از شامل ۲۱۳ تصویر که دارای ۶ حالت (Happy, Fear, Angry, Disgust, Surprise and sad) استفاده میکنید. برای بارگذاری تصاویر از image_loader.py که با تصاویر پیوست شده است استفاده نمایید.

أ. مقادیر ویژه از PCA را به ترتیب کاهشی رسم نمایید و بیان نمایید که چگونه میتوان تعداد کامپوننت مناسب را در فرآیند فشرده سازی تشخیص داد.(۱۰ نمره)

ب. ۴ مقدار ویژه اول و ۴ مقدار ویژه نهایی(eigenfaces) را نشان دهید و تحلیل کنید.(۵ نمره)

أ. حال با یک طبقه بند k-NN با مقادیر k = 1, 2 طبقه بندی را انجام دهید و CCR را گزارش کنید. (۱۰ نمره)

---

۹. الگوریتم forward selection را پیاده‌سازی نمایید. شما میتوانید از Naïve bayes optimal classifier در الگوریتم خود به عنوان طبقه‌بند استفاده کنید.

أ. CCR را بر حسب تعداد ویژگی‌های انتخاب شده در یک نمودار رسم نمایید. (۱۰ نمره)

ب. تعداد بهینه‌ی ویژگی‌ها را برای بهترین عملکرد طبقه‌بند بیان نمایید. (۱۰ نمره)

---

۱۰. حال لیبل کلاس‌ها را در نظر گرفته و ماتریس پراکندگی درون کلاسی و بین کلاسی را محاسبه نمایید تا روش LDA را پیاده‌سازی کنیم.

أ. از LDA کمک بگیرید و مقادیر ویژه را مرتب نمایید و مقادیر ویژه ماتریس جداپذیری را در قالب نزولی رسم نمایید. (۱۰ نمره)

ب. در یک نمودار مقدار $trace(S_W^{-1}S_B)$ (*separability measure*) نسبت به تعداد ویژگیها رسم نمایید و در مورد تاثیر تعداد ویژگی‌ها بر آن بحث کنید. (۱۰ نمره)

ت. تعداد بهینه‌ی ویژگی‌ها را بر اساس دو بخش پیشین بیان نمایید و حال با انتقال به زیرفضای جدید که فقط شامل ویژگی‌های بهینه است، طبقه‌بند Naïve Bayes optimal classifier با تخمین پارامتری گوسی را پیاده‌سازی کنید و مقدار CCR را گزارش کنید.(۵ نمره)

---

۱۱. حال لیبل کلاس‌ها را در نظر گرفته و مقادیر و بردار ویژه‌های ماتریس کواریانس را حساب نمایید تا در این مساله روش PCA را اعمال کنیم.

أ. مقادیر ویژه‌ی ماتریس کواریانس را بر حسب شماره ویژگی رسم نمایید. (۱۰ نمره)

ب. تعداد بهینه ی ویژگی‌ها را بر اساس بخش پیشین بیان نمایید و حال با انتقال به زیرفضای جدید که فقط شامل ویژگی‌های بهینه است، طبقه‌بند Naïve Bayes optimal classifier با تخمین پارامتری گوسی را پیاده‌سازی کنید و مقدار CCR را گزارش کنید.(۱۰ نمره)

---

۱۲. برای داده‌ی مذکور، ابتدا یک طبقه بند بهینه‌ی بیز با روش پنجره پارزن برای تخمین pdf با حالتهای زیر طراحی کنید و مقدار CCR را گزارش نمایید.:

أ. دو پنجره‌ی مستطیلی و گوسی را بررسی نمایید. (۱۰ نمره)

ب. تاثیر اندازه‌ی پنجره را بررسی کنید.(برای ۳ مقدار مختلف) (۱۰ نمره)

---

۱۳. سوال قبل را با استفاده از روش K نزدیکترین همسایه برای تخمین pdf دوباره تکرار کنید.

أ. برای سه مقدار مختلف k الگوریتم را تکرار کنید و نتایج را گزارش و تحلیل نمایید. (۱۰ نمره)

ب. در چه صورتی نتایج این سوال و سوال پیش میتوانند به یکدیگر همگرا شوند؟ توضیح دهید.. (۱۰ نمره)

---

۱۴. در این سوال قصد داریم، با استفاده از ICA، ۳ منبع تولیدکننده موسیقی که به طور همزمان کار میکنند را تشخیص دهیم. جهت اینکار، ابتدا سه سیگنال با طول‌های برابر و به صورت سینوسی، پالسی و دندان اره‌ای تولید کنید. سپس، این سه سیگنال را با یکدیگر و با یک ضریب از نویز گوسی جمع کنید. در نهایت، پس از تولید mixing matrix، با استفاده از ICA، این سه منبع را تشخیص دهید. (۲۰ نمره)