



پردیس دانشکده‌های فنی

بسمه تعالی
دانشکده مهندسی برق و کامپیوتر
تمرین سری ششم درس شناسایی الگو



دانشگاه تهران

سلام بر تمام دانشجویان عزیز، چند نکته مهم:

۱. حجم گزارش به هیچ عنوان معیار نمره‌دهی نیست، در حد نیاز توضیح دهید.
۲. نکته‌ی مهم در گزارش نویسی روشن بودن پاسخها می‌باشد، اگر فرضی برای حل سوال استفاده می‌کنید حتما آن را ذکر کنید، اگر جواب نهایی عددی است به صورت واضح آن را بیان کنید.
۳. برای سوال ۶ (شبیه‌سازی)، از دیتاست "TinyMNIST" و برای سوال ۹ از دیتاست مربوط به EEG استفاده شود. شکل ها، به طور واضح و در فرمت درست گزارش شود.
۴. هرگونه شباهت در گزارش و کد مربوط به شبیه سازی، به منزله **تقلب** می باشد و کل نمره تمرین **صفر** می‌شود.
۵. نمره کل این تمرین از ۲۰۰ می‌باشد، که حتما باید به ۱۰۰ نمره از سوالات پاسخ دهید و همچنین تا ۵۰ نمره امتیازی نیز می‌توانید تحویل دهید.
۶. در صورت داشتن سوال، از طریق ایمیل salar.nouri@ut.ac.ir سوال خود را مطرح کنید.

۱. با در نظر گرفتن رابطه $y = Wx$ ، مقدار مربوط به اطلاعات متقابل مربوط به y ، را به دست آورید. (۱۵ نمره)

۲. با استفاده از روش تخمین Maximum Likelihood، و با در نظر گرفتن رابطه $x = As = W^{-1}s$ ، تخمین مدل ICA، ماتریس W را تخمین بزنید. (۱۵ نمره)

۳. با در نظر گرفتن رابطه $x = As = W^{-1}s$ ، ابتدا آنتروپی مربوط به منبع‌ها (s) را به دست آورید. سپس، با استفاده از روش تندترین شیب، رابطه آپدیت مربوط به ماتریس W را به دست آورید. (۲۰ نمره)

۴. در مبحث مربوط به LDA، مقدار مربوط به Criterion Function را برحسب ماتریس‌های پراکندگی داخل کلاسی و بین کلاسی به دست آورید. با در نظر گرفتن $y = g(x) = W^t x + w_0$ ، نشان دهید که Criterion Function از w_0 مستقل می‌باشد. همچنین به ازای چه ماتریس W ، این Criterion Function حداکثر مقدار خود را دارد. (۲۵ نمره)

۵. یک راه جهت تعمیم‌سازی فرایند minimum mean-squared-error، تعریف کردن Criterion Function زیر می‌باشد:

$$J_T = \sum_{i=1}^c \sum_{x \in w_i} (x - m_i)^T S_T^{-1} (x - m_i)$$

که در آن m_i میانگین n_i نمونه در w_i و ماتریس S_T پراکندگی کلی می‌باشد. نشان دهید که انتقال نمونه \hat{x} از w_i به w_j موجب می‌شود که J_T به شکل زیر تغییر کند: (۲۵ نمره)

$$J_T^* = J_T + \left[\frac{n_j}{n_j + 1} (\hat{x} - m_j)^T S_T^{-1} (\hat{x} - m_j) - \frac{n_i}{n_i - 1} (\hat{x} - m_i)^T S_T^{-1} (\hat{x} - m_i) \right]$$

۶. بدون در نظر گرفتن لیبل‌های کلاس، در بین ویژگی‌های داده‌ی آموزش، وابستگی خطی پیدا کنید. همچنین یک تبدیل خطی جهت سفید کردن داده‌ها به دست آورید.

الف) حال لیبل‌های کلاس را در این داده‌های سفید شده در نظر بگیرید. ماتریس‌های پراکندگی S_T ، S_B را به دست آورید. با استفاده از LDA، به تعداد ۱۹۶ مورد از این ویژگی‌ها را به صورت نزولی مطابق با معیار جداپذیری مرتب کنید. بردارهای ویژه مربوطه را به صورت نزولی رسم کنید. (۱۵ نمره)

ب) یک سری زیرفضاهای دیگر در نظر بگیرید و داده‌ها را به این زیر فضاها تصویر کنید. نمودار مربوط به معیار جداپذیری را بر حسب تعداد Component هایی که برای هر زیرفضا استفاده شده است، رسم کنید. همچنین در مورد تاثیر بعد این زیرفضاها بر روی معیار جداپذیری بحث کنید. (۱۵ نمره)

۷. در این سوال قصد داریم، که 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. (۲۵ نمره)

۸. همانطور که در کلاس بیان شد، یکی از شرایط اینکه بتوان از الگوریتم ICA، جهت تشخیص کور منابع استفاده کنیم، این است که غیرگوسی باشند. در این سوال، سه تا منبع گوسی با میانگین و واریانس متفاوت، که به صورت همزمان کار می‌کنند، را در نظر بگیرید، حال با استفاده از ICA، به تشخیص این منابع بپردازید. مشاهده خود را شرح دهید. (۲۰ نمره)

۹. با استفاده از الگوریتم ICA، به پردازش سیگنال‌های EEG پرداخته و سورس‌های سیگنال‌های حاصل را تشخیص دهید. برای اینکار لازم است که در دیتاست مربوطه، نویز حاصل از سایر عوامل را حذف کنید. و همچنین در صورت لزوم، از الگوریتم‌های کاهش بعد، Feature Selection، قبل اعمال ICA به داده‌های مربوطه اعمال شود. (۲۵ نمره)

دیتاست: این دیتاست به صورت فایل mat. در اختیار شما قرار گرفته است، که داده‌های مربوطه در data قرار دارند. این دیتا به صورت ماتریس $64 \times 640 \times 99$ می‌باشد، که به ترتیب، ۶۴ معرف تعداد کانال‌ها، ۶۴۰ معرف زمان (۶۴۰ میلی‌ثانیه) و ۹۹ معرف تعداد epoch ها می‌باشد.

جهت حذف نویز از فایل زیر می‌توانید استفاده کنید:

<https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwiUiPyootnpAhVGQhUIHSE2BN8QFjAGegQIBxAB&url=https%3A%2F%2Fcore.ac.uk%2Fdownload%2Fpdf%2F11050484.pdf&usg=AOvVaw2NmfEtOzaXx5KO8DlsjAMx>