به نام خدا

دانشگاه تهران

دانشکده مهندسی برق و کامپیوتر

# درس شبکه‌های عصبی و یادگیری عمیق

## تمرین Extra

| | | |
|---|---|---|
| پرسش ۱ | نام دستیار طراح | سیاوش رزمی |
| | رایانامه | siavashrazmi74@gmail.com |
| پرسش 2 | نام دستیار طراح | سجاد علیخانی |
| | رایانامه | ichbinsajadalikhani@gmail.com |
| پرسش 3 | نام دستیار طراح | مصطفی حاجی قاسملو |
| | رایانامه | mos.haji@ut.ac.ir |
| | مهلت ارسال پاسخ | ۱۴۰۱.۰۹.۲۱ |

# فهرست

قوانین ..... ۱

پرسش ۱. تشخیص تقلب (fraud detection) با استفاده از شبکه عمیق ..... ۳

پرسش ۲ – Liveness Detection ..... ۴

پرسش ۳ – تشخیص کاراکتر نوری ..... ۶

# قوانین

قبل از پاسخ دادن به پرسش‌ها، موارد زیر را با دقت مطالعه نمایید:

- از پاسخ‌های خود یک گزارش در قالبی که در صفحه‌ی درس در سامانه‌ی Elearn با نام ***REPORTS_TEMPLATE.docx*** قرار داده شده تهیه نمایید.
- پیشنهاد می‌شود تمرین‌ها را در قالب گروه‌های دو نفره انجام دهید. (بیش از دو نفر مجاز نیست و تحویل تک نفره نیز نمره‌ی اضافی ندارد) توجه نمایید الزامی در یکسان ماندن اعضای گروه تا انتهای ترم وجود ندارد. (یعنی، می‌توانید تمرین اول را با شخص A و تمرین دوم را با شخص B و ... انجام دهید)
- کیفیت گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است؛ بنابراین، لطفا تمامی نکات و فرض‌هایی را که در پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید در گزارش ذکر کنید.
- در گزارش خود مطابق با آنچه در قالب نمونه قرار داده شده، برای شکل‌ها زیرنویس و برای جدول‌ها بالانویس در نظر بگیرید.
- الزامی به ارائه توضیح جزئیات کد در گزارش نیست، اما باید نتایج بدست آمده از آن را گزارش و تحلیل کنید.
- تحلیل نتایج الزامی می‌باشد، حتی اگر در صورت پرسش اشاره‌ای به آن نشده باشد.
- دستیاران آموزشی ملزم به اجرا کردن کدهای شما نیستند؛ بنابراین، هرگونه نتیجه و یا تحلیلی که در صورت پرسش از شما خواسته شده را به طور واضح و کامل در گزارش بیاورید. در صورت عدم رعایت این مورد، بدیهی است که از نمره تمرین کسر می‌شود.
- در صورت مشاهدهٔ تقلب امتیاز تمامی افراد شرکت‌کننده در آن، ۱۰۰- لحاظ می‌شود.
- تنها زبان برنامه نویسی مجاز **Python** است.
- استفاده از کدهای آماده برای تمرین‌ها به هیچ وجه مجاز نیست.
- نحوه محاسبه تاخیر به این شکل است: پس از پایان رسیدن مهلت ارسال گزارش، حداکثر تا یک هفته امکان ارسال با تاخیر (به ازای هر روز ۵ درصد کسر نمره) وجود دارد، پس از این یک هفته نمره آن تکلیف برای شما صفر خواهد شد. (در مورد کسر ۵ درصد نمره در یک هفته‌ی ارسال با تاخیر، دقت بفرمایید که در انتهای ترم در مجموع ۱۸روز بخشش جریمه، برای کمک به شما عزیزان در نظر گرفته شده‌است).
- لطفا گزارش، کدها و سایر ضمایم را به در یک پوشه با نام زیر قرار داده و آن را فشرده سازید، سپس در سامانه‌ی Elearn بارگذاری نمایید:

HW[Number] _[Lastname]_[StudentNumber]_[Lastname]_[StudentNumber].zip

(مثال: HW1_Ahmadi_810199101_Bagheri_810199102.zip)

- برای گروه‌های دو نفره، بارگذاری تمرین از جانب یکی از اعضا کافی است ولی پیشنهاد می‌شود هر دو نفر بارگذاری نمایند.

## پرسش ۱. تشخیص تقلب (fraud detection) با استفاده از شبکه عمیق

**۱-۱** مقاله ایی که برای پیاده‌سازی در نظر گرفته شده Credit Card Fraud Detection Using Autoencoder Neural Network است، پس از مطالعه ی مقاله به سؤالات پاسخ دهید:

۱- بزرگترین چالش ها در توسعه مدل های تشخیص تقلب چیست؟ این مقاله برای حل این چالش ها از چه متد هایی استفاده کرده است؟

۲- در مورد معماری شبکه ارائه شده در مقاله به شکل مختصر توضیح دهید.

۳- انواع روش‌های Resampling موجود برای balance کردن دیتاست را نامبرده و مزایا و معایب هر کدام را توضیح دهید؟

۴- مدل ارائه شده را پیاده‌سازی کرده و با استفاده از این داده آموزش دهید.

( برای جلوگیری از overfitting آموزش مدل را طوری تنظیم کنید که در انتهای آموزش بهترین وزن های مدل بر اساس خطای validation set برگردانده شوند)

۵- نمودار Heatmap را برای Confusion matrix پیش‌بینی مدل بر روی داده‌های تست رسم کنید و مقادیر accuracy، precision، Recall و f1score را در گزارش ذکر کنید، فکر میکنید در مسائلی که توزیع Label ها نامتوازن است استفاده از معیار Accuracy به تنهایی عمل‌کرد مدل را به درستی نمایش میدهد؟ چرا؟ اگر نه کدام معیار میتواند به عنوان مکمل استفاده شود؟

۶ – با threshold های مختلف برای oversampling عمل‌کرد مدل را بررسی و نمودار recall& accuracy را مانند شکل ۷ مقاله رسم کنید.

۷- مدل را با استفاده از داده‌های unbalanced و بدون حذف نویز، آموزش و موارد سؤال ۶ را گزارش دهید، نتایج دو مدل را با هم مقایسه کنید.

(نکته: در صورتی که مقادیر هرکدام از هایپر پارامتر های مدل در مقاله ذکر نشده باشد، در انتخاب آن‌ها آزادی عمل کامل دارید.)

## پرسش ۲ – تشخیص زنده بودن

الف) با توجه به مقاله مرجع، Liveness Detection به چه منظور انجام می‌شود؟ انواع راهکارهای بر پایه مشخصات بیومتریکی برای مقابله با حملات کلاه‌برداری را نام ببرید. با استفاده از چه ویژگی‌های اثر انگشت می‌توان به زنده بودن آن در سیستم‌های Liveness Detection پی برد؟

ب) در Dataset_1، که از لینک مشخص شده قابل دسترسی می باشد، دو دسته تصویر واقعی و جعلی وجود دارند. با استفاده از شبکه‌های عصبی کانولوشنی مدلی برای تفکیک کردن این دو دسته آموزش دهید. ضمن ارائه بهترین معماری شبکه عصبی و توابع فعال‌ساز و بهینه‌سازی که به آن دست پیدا کرده اید، نمودارهای خطا و دقت را برای آموزش و تست مدل رسم کنید.

پ) با استفاده از یک شبه‌کد نشان دهید احراز هویت با استفاده از تشخیص پلک زدن چشم طی چه مراحلی می تواند در سیستم های Liveness Detection پیاده‌سازی شود. برای هر مرحله توضیحی مختصر ارائه کنید.

ت) مدل‌های LeNet-5 و AlexNet مبتنی بر شبکه‌های عصبی کانولوشنی را رسم کنید و کاربردهای رایج آن‌ها را نام ببرید. به نظر شما کدام یک از این دو مدل برای تشخیص باز یا بسته بودن چشم مناسب‌تر است؟ چرا؟

ث) در Dataset_2 دو دسته تصویر مربوط به چشم های باز و بسته وجود دارند. مراحلی که در قسمت (پ) برای تشخیص زنده بودن چهره به وسیله تشخیص پلک زدن چشم ارائه دادید را اکنون پیاده‌سازی کنید. بدین منظور می‌توانید یکی از دو راه زیر را انتخاب کنید.

راه اول) استفاده از چهره خودتان در ویدیوی وب‌کم و نشان دادن هم‌زمان تصویر شخصی دیگر: در این حالت پس از تشخیص چهره و چشم‌ها، به محض پلک زدن باید به عنوان یک چهره زنده شناسایی شوید، درحالی‌که تصویر شخص دیگر با وجود شناسایی چشم‌ها به دلیل باز و بسته نشدن چشم‌ها زنده شناسایی نشود.

راه دوم) استفاده از قسمتی کوتاه از یک فیلم که در آن حداقل دو نفر برای تشخیص پلک زدن وجود دارند: در این حالت پس از شناسایی چشم‌ها، به محض پلک زدن هر یک از اشخاص باید به نشانه صحت زنده بودن، نام شخص موردنظر در ویدیو نشان داده شود.

برای انجام هر یک از دو راه ذکر شده می‌توانید از طبقه‌بند از پیش آموزش دیده شده Haar-cascade ، که فایل‌های آن به تمرین ضمیمه شده‌اند، برای تشخیص چهره و چشم‌ها و از شبکه‌های عصبی

کانولوشنی برای تشخیص باز یا بسته بودن چشم‌ها استفاده کنید. برای کسب اطلاعات بیشتر درباره نحوه کار Haar Cascades می‌توانید به این لینک مراجعه کنید.

نمودارهای خطا و دقت طبقه‌بندی تصاویر چشم‌های باز و بسته را به همراه ویدیوی خروجی از تشخیص زنده بودن چهره‌ها گزارش کنید. تمامی بخش‌های کد باید دارای کامنت باشند.

## پرسش ۳ – تشخیص کاراکتر نوری(Optical character recognition)

در این  تمرین به شبیه سازی مقاله:

A recognition model for handwritten Persian/Arabic numbers based on optimized deep convolutional neural network

با استفاده از دیتا ست HODA می باشد. این مقاله به بررسی تشخیص اعداد فارسی با استفاده از معماری DCNN می باشد.( مقاله و دیتاست پیوست شده است).

الف) تفاوت بین شبکه های CNN و DCNN را توضیح دهید.

ب) سه روش بهینه سازی  Adam و Adadelta و Momentum را توضیح دهید.

ج) معماری DCNN استفاده شده در مقاله راپیاده سازی کنید. پیش پردازش ونرمالایز سازی های مورد استفاده را بیان کنید. تعداد لایه های بکار رفته و نوع لایه و علت استفاده ازآن ها را توضیح دهید. به منظور جلوگیری از overfitting چه تکنیکی به کار رفته است.

د) نمودارهای  accuracy  و loss وهم چنین Confusion matrix و مقادیرprecision، Recall و f1score را برای هر کدام از سه روش  بهینه سازی بیان کرده و مقایسه کنید.

ه ) معماری و پارامترهای بهترین شبکه را بیان کنید.