

به نام خدا



دانشگاه تهران  
پردیس دانشکده‌های فنی  
دانشکده برق و کامپیوتر



## شبکه های عصبی و یادگیری عمیق

تمرین شماره ۲

آبان ۹۸

## سوال ۱ – Multi-Layer Perceptron

### بخش اول:

به سوالات زیر پاسخ دهید.

**الف)** همان طوری که می دانید شبکه های عصبی می توانند تنها با ۲ لایه مخفی خاصیت general function approximator بودن خود را حفظ کنند. با این وجود، دلیل این که گاه بیش از ۲ لایه برای این شبکه ها انتخاب می کنیم چیست؟

ب) تابع هزینه cross entropy چگونه هزینه را محاسبه می کند؟

ج) تاثیر batch size و learning rate در آموزش شبکه به چه صورت است؟

د) چرا از validation set استفاده می کنیم؟

### بخش دوم:

در این بخش از تمرین توسط شبکه های multi-layer perceptron (MLP) به طبقه بندی داده های مجموعه داده MNIST می پردازید. می توانید با استفاده از [این لینک](#) با این مجموعه داده آشنا شوید.

با توجه به این که هدف آشنایی با شبکه های MLP است نه درگیری با داده، با استفاده از کد قرار داده شده می توانید داده ها را در سه دسته train، test و validation به برنامه وارد کرده و بخشی از این داده ها را رسم کنید.

**الف)** برای طراحی هر شبکه عصبی نیاز به تنظیم تعدادی هاپر پارامتر داریم: تعداد لایه های درونی، تعداد نوروها در هر لایه، تابع فعال ساز و... با در نظر گرفتن مقادیر داده شده زیر برای پارامترهای قابل تنظیم شبکه و با استفاده از داده های validation، جدول زیر را پر کنید.

تعداد epoch های یادگیری را ۳۰ و تعداد لایه های مخفی را ۲ و تعداد نوروهای لایه های مخفی را ۵۱۲ در نظر بگیرید و برای محاسبه هزینه از تابع cross entropy استفاده کنید. توجه داشته باشید که تمامی این پارامترها قابل تنظیم می باشند، اما ما تنها تاثیر تنظیم ۲ پارامتر ذکر شده را در این تمرین خواسته ایم.

<i>batch</i>	<i>learningrate</i>	<i>validation score</i>
4	0.1	
4	0.01	
32	0.1	
32	0.01	

ب) با انتخاب بهترین مدل، دقت شبکه را برای طبقه‌بندی داده‌های تست گزارش کنید.

## سوال ۲ – Autoencoder

**الف)** در دیتاست MNIST به کمک یک Autoencoder ابتدا فضای ویژگی را کاهش دهید و با استفاده از پارامترهای به دست آورده شده در سوال ۱، به آموزش یک شبکه MLP بپردازید. دقت نهایی روی داده‌های تست را گزارش کنید و سپس نتیجه را با نتیجه به دست آورده شده در سوال ۱ مقایسه کنید.

**ب)** نتایج قسمت الف را با الگوریتم PCA (principal component analysis) از نظر میزان و کیفیت فشرده‌سازی مقایسه کنید.

جهت راهنمایی برای حل این سوال، به این [لینک](#) مراجعه فرمایید.

### سوال ۳ - RBM

سوال ۲ را دوباره حل کنید؛ ولی بجای استفاده از autoencoder ها از RBM های به هم لینک شده برای کاهش بعد استفاده کنید و نتایج را با نتایج سوال ۲ مقایسه کنید.

## سوال ۴ - MLP

**الف)** با آموزش یک شبکه MLP با تک لایه مخفی تابعی را جهت پیش بینی قیمت خانه در شهر بوستون تقریب بزنید.

(این مجموعه داده شامل ۵۰۵ مشاهدات از ۱۳ ویژگی بعنوان ورودی و یک خروجی قیمت می باشد). بصورت تصادفی ۸۰ درصد داده ها را بعنوان داده test و باقی را به عنوان داده train استفاده کنید.

**ب)** یک لایه مخفی دوم را به شبکه اضافه کنید و با تنظیم تعداد نرون ها دوباره تابع را تقریب بزنید.

**ج)** ورودی ها را با روش PCA خطی کاهش ابعاد دهید و سعی کنید نتایج قسمت الف وب را بهبود دهید. در هر روش، شاخص خطای متوسط حداقل مربعات خطا را برای هر epoch رسم کنید. در یک جدول، شاخص متوسط مربعات خطا و حداکثر خطای مربوط به روش های آموزش داده شده سوال را با هم مقایسه کنید.

## نکات:

- مهلت تحویل این تمرین ۱۰ آبان می باشد.
- گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است. لطفا تمامی نکات و فرض‌هایی که برای پیاده سازی‌ها و محاسبات خود در نظر می‌گیرید را در گزارش ذکر کنید.
- در صورت مشاهده‌ی تقلب نمرات تمامی افراد شرکت کننده در آن صفر لحاظ می‌شود.
- استفاده از کدهای آماده برای تمرین‌ها مجاز نمی‌باشد. برای مینی پروژه‌ها فقط برای قسمت‌هایی از کد و به عنوان راهنمایی برای پیاده‌سازی می‌توانید از کدهای آماده استفاده کنید.
- نحوه‌ی محاسبه‌ی تاخیر به این شکل است: مهلت بدون کسر نمره تا تاریخ اعلام شده و تاخیر تا یک هفته بعد از مهلت، یعنی ۱۷ آبان ماه با ۳۰ درصد کسر نمره محاسبه خواهد شد.
- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه‌ی زیر با دستیار آموزشی مربوطه در تماس باشید (سوالات ۲ و ۳ خانم نیلوفر شه‌دوست و سوالات ۱ و ۴ خانم مرجان شاهی):

**[n.shahdoust@ut.ac.ir](mailto:n.shahdoust@ut.ac.ir)**

**[Marjan.shahi71@gmail.com](mailto:Marjan.shahi71@gmail.com)**