به نام خدا

دانشکده علوم و فنون نوین

مبانی محاسبات (رایانش) نرم (۰۱-۰۳۹-۰۵-۸۳)

نیم‌سال اول ۱۴۰۲-۱۴۰۱

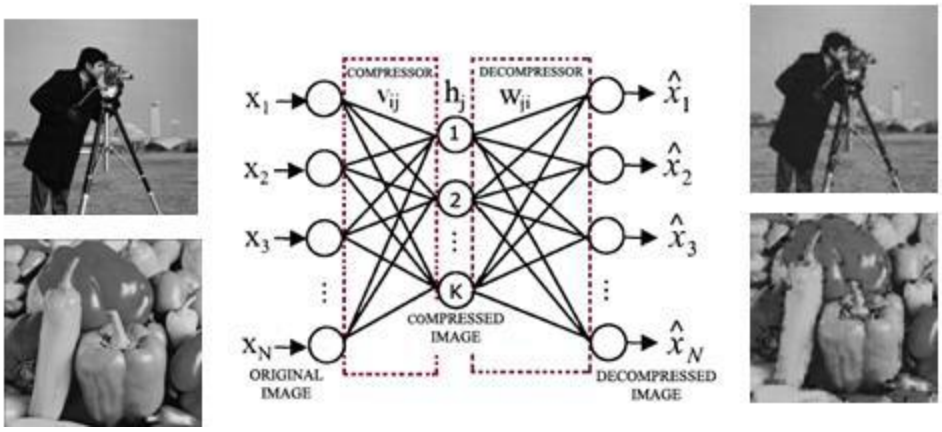
تمرین شماره ۲

تاریخ تحویل: ۱۴۰۱/۰۹/۰۷

---

۱. (۴۰٪) [**پیاده‌سازی: پرسپترون چندلایه برای فشرده‌سازی**] در این تمرین به پیاده‌سازی فشرده‌سازی تصاویر با استفاده از پرسپترون چندلایه می‌پردازیم. برای این کار، مشابه آنچه در مثال ۶-۵ (فصل ششم) و مطالب ارائه شده در کلاس بیان شده است، یک شبکه سه لایه ایجاد کنید که لایه ورودی معادل تصویر ورودی (با n نرون)، لایه میانی معادل تصویر فشرده شده (با m<n نرون) و لایه خروجی معادل تصویر بازسازی شده (با n نرون) است. شبکه را با داده‌های ارائه شده همراه این تمرین (داده‌های آموزش) آموزش دهید و با مجموعه آزمون ارائه شده تست کنید.

- **نکته:** تمامی مراحل مربوط به الگوریتم، بایستی توسط دانشجو پیاده‌سازی شود و استفاده از توابع آماده مجاز نیست مگر برای مقایسه با پیاده‌سازی خودتان.



- **نکته:** تمام تصاویر مورد استفاده در این تمرین خاکستری بوده و اندازه آنها ۲۵۶*۲۵۶ است که می‌توانید در فرآیند آموزش و آزمون، آن ها را به بلوک‌های ۸*۸ تقسیم کنید.
- **نکته:** برای راهنمایی بیشتر در مورد این تمرین و آشنایی با برخی ایده‌های کمکی جهت بهبود کارایی این روش می‌توانید به مقالات ارائه شده همراه با این تمرین مراجعه کنید. همچنین برای محاسبه کارایی شبکه، از معیار PSNR[1] بیان شده در این مقالات استفاده نمایید.

گزارش خود را از عملکرد شبکه برای موارد زیر بیان کنید.

[1]- Peak Signal to Noise Ratio (PSNR)

h.veisi@ut.ac.ir

به نام خدا

دانشکده علوم و فنون نوین

مبانی محاسبات (رایانش) نرم (۰۱-۰۳۹-۰۵-۸۳)

نیم‌سال اول ۱۴۰۲-۱۴۰۱

تاریخ تحویل: ۱۴۰۱/۰۹/۰۷

تمرین شماره ۲

۱-۱ [۳۰٪] [تعداد نرون‌های مخفی] شبکه را برای حداقل سه تعداد مختلف ۳۲، ۱۶، ۴ از نرون‌های لایه مخفی آموزش داده و در هر حالت، برای مجموعه آزمون، میانگین مقدار معیار PSNR را به همراه تصاویر اصلی و بازسازی شده ارائه دهید. نمودار خطای شبکه در حین آموزش را برای هر سه حالت رسم کنید.

۱-۲ [۱۰٪] [پس‌انتشار با گشتاور] روش یادگیری شبکه را به پس‌انتشار با گشتاور(ممان) تغییر دهید و نتیجه آن را در آموزش شبکه با روش استاندارد مقایسه کنید. برای این کار، علاوه بر تعداد تکرارها، زمان همگرایی شبکه را نیز گزارش کنید. کارایی شبکه در این حالت برای تعداد ۱۶ نرون در لایه مخفی را با حالت بدون گشتاور (با همین تعداد نرون مخفی) برای داده‌های آزمون به صورت میانگین مقدار معیار PSNR گزارش کنید.

---

۲. (۵۰٪) **[پیاده‌سازی: پرسپترون چندلایه برای تشخیص یک دسته]** در این تمرین با استفاده از یک شبکه‌ی پرسپترون چندلایه قرار است مدلی را آموزش دهیم که تصویر اعداد دست‌نویس مربوط به اعداد ۰ تا ۹ از مجموعه داده MNIST را تشخیص دهد. این مجموعه داده از تصاویر مربعی با ابعاد ۲۸*۲۸ پیکسل تشکیل شده که هرکدام حاوی یکی از ارقام صفر تا ۹ به صورت دست‌نویس است. در این مجموعه داده از هر رقم تعداد ۶ هزار تصویر در بخش آموزش (train) و هزار تصویر در بخش آزمایش (test) موجود است. نمونه‌هایی از این تصاویر در شکل زیر آورده شده است. مجموعه‌ی کامل این داده‌ها و توضیحات بیشتر در این مورد در لینک زیر موجود است.

http://yann.lecun.com/exdb/mnist

h.veisi@ut.ac.ir

به نام خدا

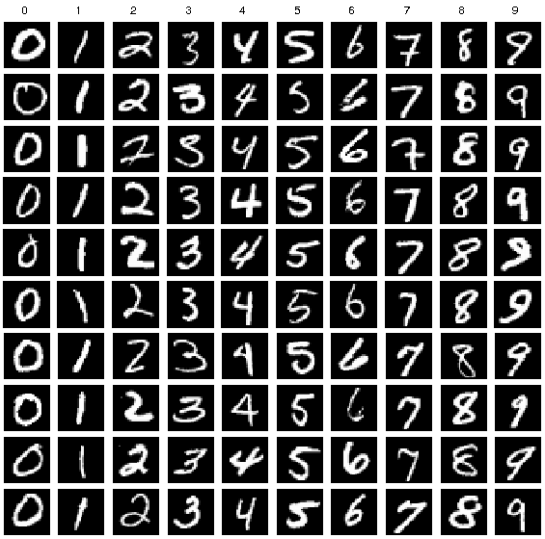
دانشکده علوم و فنون نوین

مبانی محاسبات (رایانش) نرم (۰۱-۰۳۹-۰۵-۸۳)

نیم‌سال اول ۱۴۰۲-۱۴۰۱

تمرین شماره ۲

تاریخ تحویل: ۱۴۰۱/۰۹/۰۷



شکل ۱ نمونه تصاویر مجموعه دادگان MNIST

۲-۱ [۳۵٪] [آموزش auto-encoder] در این بخش به کمک شبکه‌ی پرسپترون چندلایه یک شبکه auto-encoder (مشابه با مثال فشرده‌سازی تصویر) برای همه اعداد در مجموعه داده‌ی MNIST آموزش می‌دهیم. برای این منظور لازم است الگوریتم پرسپترون چندلایه را به صورت کلی (با امکان تغییر تعداد لایه‌های مخفی، تعداد نرون‌های هر لایه، تنظیم تابع فعال‌ساز و...) پیاده‌سازی کنید. **توجه کنید که تمامی مراحل مربوط به این الگوریتم باید توسط دانشجو پیاده‌سازی شود و استفاده از توابع آماده مجاز نیست.**

با داشتن الگوریتم پرسپترون چندلایه می‌توان از شبکه‌ای با ۷۸۴ نرون ورودی (پیکسل‌های یک تصویر) و لایه‌ی میانی با تعداد نرون کمتر (مثلاً ۶۴) و لایه‌ای خروجی با ۷۸۴ نرون (تصویر بازسازی شده) به عنوان auto-encoder استفاده کرد. با آموزش چنین شبکه‌ای می‌توانیم تصویر ورودی را در لایه‌ی وسط شبکه encode کنیم و سپس مجدداً با بخش دوم شبکه، آن را decode کرده و تصویر اصلی را بازسازی کنیم. از Mean Squared Error به عنوان خطای بازشناسی استفاده کنید.

یک شبکه‌ی پرسپترون سه لایه با تعداد ۷۸۴، ۶۴ و ۷۸۴ نرون در هر لایه و با استفاده از تابع فعال‌ساز sigmoid ایجاد کنید و یادگیری آن را با نرخ یک هزارم انجام دهید (بدون پس‌انتشار با گشتاور و بدون

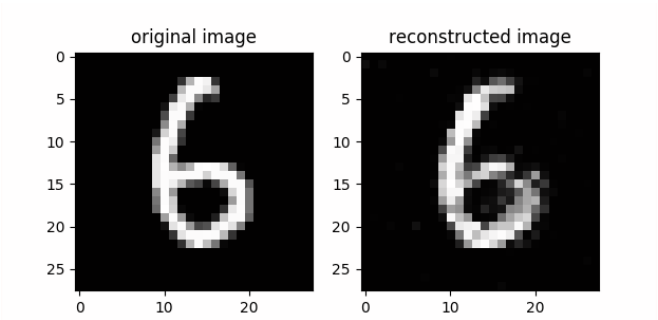
h.veisi@ut.ac.ir

به‌روز کردن دسته‌ای). در هریک از بخش‌های **الف** و **ب** تغییرات گفته شده را بر روی شبکه با همین پارامترها اعمال کنید.

نکته ۱: در این بخش از همه داده‌ها استفاده کنید و قسمتی از این داده‌های آموزش را به عنوان داده‌های validation جدا کرده و میزان خطای آن را معیاری برای زمان توقف الگوریتم و ارزیابی مدل قرار دهید. از هیچ‌یک از داده‌های آزمایش در آموزش استفاده‌ای نمی‌کنیم و آن‌ها را برای ارزیابی نهایی نگه می‌داریم.

نکته ۲: برای هریک از اجراهای خواسته شده نموداری از روند تغییرات خطای MSE شبکه بر روی داده‌های train و validation رسم کنید (این دو نمودار را روی یکدیگر قرار دهید). در گزارش خود با تحلیل این نمودارها برای هر بخش، تأثیر موارد خواسته شده بر روی «سرعت همگرایی بر اساس تعداد epoch» و «توان شبکه در یادگیری الگو» را بررسی کرده و علت آن را بیان کنید.

نکته ۳: همچنین یک تصویر نمونه از مجموعه دادگان (بخش validation) را در نظر بگیرید و برای هر بخش، خروجی بهترین شبکه (تصویر بازسازی شده) را با تصویر اصلی مقایسه کنید (مانند شکل ۲).



شکل ۲ مقایسه‌ی تصویر اصلی و تصویر بازسازی شده توسط auto-encoder

الف) تعداد نرون‌های لایه‌ی مخفی را هریک از مقادیر ۶۴ و ۱۲۸ قرار دهید و برای هر حالت، بعد از آموزش، کارایی شبکه را بر روی داده آزمایش گزارش کنید. برای محاسبه کارایی، مقدار MSE بین تصویر اصلی آزمون و تصویر بازسازی شده در Decoder را حساب کرده و میانگین آن را گزارش کنید

h.veisi@ut.ac.ir

ب) تعداد لایه‌های مخفی را از یک به ۳ و ۵ به صورت زیر تغییر دهید و مقدار میانگین MSE را روی داده آزمون در هر حالت گزارش کنید.

- تعداد نرون‌ها برای سه لایه‌ی مخفی به ترتیب ۲۵۶ – ۶۴ – ۲۵۶
- تعداد نرون‌ها برای پنج لایه‌ی مخفی به ترتیب ۲۵۶ – ۱۲۸ – ۶۴ – ۱۲۸ – ۲۵۶

پ) برای بهترین شبکه در بخش‌های قبل، تابع فعالسازی را به ReLU تغییر داده و کارایی آن را گزارش کنید.

ت) برای بهترین شبکه حاصل در بخش‌های قبلی (الف تا پ)، ضریب یادگیری را مقادیر یک میلیونم و یک صدم قرار دهید و نتیجه را گزارش دهید.

ث) برای بهترین شبکه حاصل در بخش‌های قبلی (الف تا پ)، از به‌روز کردن دسته‌ای، با اندازه‌های دسته ۸ و ۱۶ استفاده کنید و مقایسه کارایی آن‌ها با حالتی که از به‌روز کردن دسته‌ای استفاده نمی‌کنید، بیان کنید.

ج) برای بهترین شبکه حاصل در بخش‌های قبلی، تعداد ۵ تصویر (از داده‌های آزمایشی) با بیشترین میزان خطای بازسازی را رسم کنید (مانند شکل ۲ آن‌ها را با تصویر اصلیشان مقایسه کنید). همین کار را برای ۵ تصویر با کمترین خطای بازسازی انجام دهید.

در نهایت با جمع‌بندی موارد بررسی شده بهترین پارامترهای پیشنهادی خود را بیان کنید و نتیجه‌گیری کدام شبکه و پارامترها بهترین هستند (از این شبکه برای قسمت بعدی سوال استفاده کنید).

۲-۲ [۱۵٪] [دسته‌بندی اعداد] با استفاده بهترین auto-encoder آموزش داده شده در سوال ۱-۱، می‌خواهیم یک دسته‌بند مبتنی بر MLP برای دسته‌بندی اعداد آموزش دهیم. برای این کار، داده‌های اعداد را از Encoder عبور داده و خروجی آن را به عنوان بردار ویژگی در نظر می‌گیریم. حال یک شبکه MLP سه لایه می‌سازیم با تعداد نرون‌های ورودی برابر با خروجی Encoder و تعداد نرون‌های خروجی برابر با ۱۰ (هر عدد یک نرون) و تعداد نرون دلخواه. این شبکه را با داده آموزش در بخش دادگان، آموزش داده و با داده آزمون ارزیابی کنید.

به نام خدا

مبانی محاسبات (رایانش) نرم (۰۱-۰۳۹-۰۵-۸۳)
نیم‌سال اول ۱۴۰۲-۱۴۰۱

تاریخ تحویل: ۱۴۰۱/۰۹/۰۷

تمرین شماره ۲

الف) شبکه MLP را با حداقل دو تعداد مختلف نرون در لایه مخفی آموزش دهید و مقدار Accuracy را روی داده آزمون محاسبه کنید. مقادیر پارامترهای مورد استفاده دیگر مانند نرخ یادگیری را نیز گزارش کنید.

ب) بهترین نتایج خود را با بهترین نتایج گزارش شده در صفحه این دادگان مقایسه کنید، به ویژه با نتایج گزارش شده برای شبکه MLP. به نظر شما چطور می‌توان کارایی را بهتر کرد تا به بهترین نتایج گزارش شده در این وب سایت دست یافت؟

---

۳. (۱۰٪) [**بدست آوردن قانون یادگیری**] فرض کنید یک شبکه عصبی تک لایه برای طبقه‌بندی دودویی(Binary classification) دارید. ورودی این شبکه $X \in R^{n*m}$ و خروجی $\hat{y} \in R^{1*m}$ و برچسب واقعی $y \in R^{1*m}$ است. فرمول‌های پیش‌خور(Forward propagation) به صورت زیر هستند:

$$z^{[1]} = W^{[1]}X + b^{[1]}$$
$$a^{[1]} = \sigma(z^{[1]})$$
$$\hat{y} = a^{[1]}$$
$$J = -\sum_{i=1}^{m} y^{(i)} log(\hat{y}^{[i]}) + (1 - y^{(i)}) log(1 - \hat{y}^{[i]})$$

عبارت $\frac{\partial J}{\partial W^{[1]}}$ را با استفاده از قانون زنجیر (Chain rule) به دست آوید.

h.veisi@ut.ac.ir