وَنام خدا

دانشکده سامانه‌های هوشمند

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

نیم‌سال اول ۱۴۰۵-۱۴۰۴

تمرین شماره ۲

تاریخ تحویل: ۱۴۰۴/۰۹/۰۵

---

۱. (۱۰٪) [تحقیق و پژوهش درباره توابع خطا] در این تمرین، هدف بررسی و مقایسه چند تابع خطا و معیار ارزیابی پرکاربرد در یادگیری ماشین است. موارد زیر را در نظر بگیرید:

- MAE (Mean Absolute Error)
- MSE (Mean Squared Error)
- RMSE (Root Mean Squared Error)
- R^2 (Coefficient of Determination)
- Cross-Entropy Loss

۱-۱ هر یک از معیارهای فوق را به‌صورت ریاضی تعریف کنید.

۱-۲ برای هر تابع خطا، مزایا و معایب آن را توضیح دهید. به‌طور خاص به موارد زیر توجه کنید:

- حساسیت به نقاط پرت(Outliers)
- هموار بودن و مشتق‌پذیری
- قابلیت استفاده در رگرسیون یا طبقه‌بندی

۱-۳ یک جدول مقایسه‌ای تهیه کنید که در آن ویژگی‌های کلیدی این توابع/معیارها (نوع داده مناسب، حساسیت، حوزه‌ی کاربرد، مزایا و معایب) به‌طور خلاصه بیان شده باشد.

۱-۴ مثال بزنید که در چه مسائل واقعی هر معیار بیش‌تر استفاده می‌شود (مثلاً پیش‌بینی قیمت خانه با MSE/RMSE، مسائل مقاوم در برابرoutlier ها باMAE ، مسائل طبقه‌بندی دودویی و چندکلاسه با-Cross Entropy Loss، ارزیابی مدل‌های رگرسیونی با R^2).

---

۲. (۴۰٪) [پیاده‌سازی: پرسپترون چندلایه برای فشرده‌سازی تصویر] 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.

h.veisi@ut.ac.ir

وَنام خدا

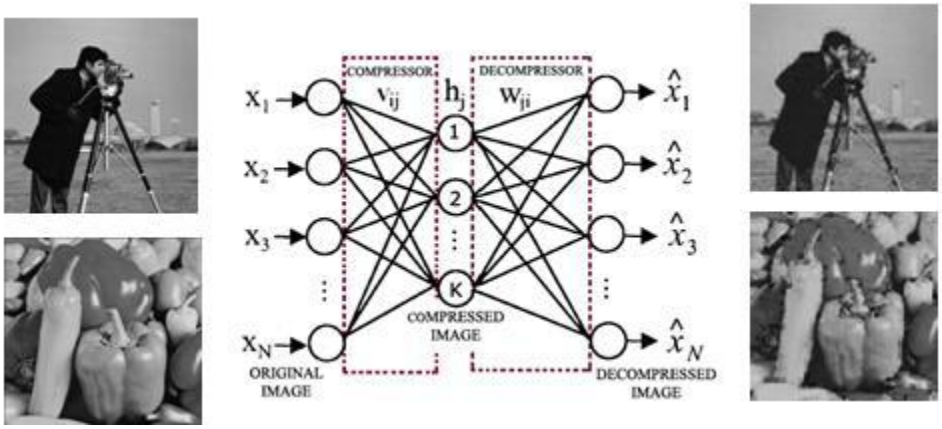
دانشکده سامانه‌های هوشمند

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

نیم‌سال اول ۱۴۰۵-۱۴۰۴

تاریخ تحویل: ۱۴۰۴/۰۹/۰۵

تمرین شماره ۲



شکل ۱ معماری مدل

- **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.

۲-۱ [۲۰٪] **[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.(از بلوک ۸*۸ در این بخش استفاده کنید)

۲-۲ [۱۰٪] **[تعداد بلوک‌های مختلف]** شبکه را برای سه اندازه بلوک ۴*۴ و ۸*۸ و ۱۶*۱۶ آموزش داده و در هر حالت، برای مجموعه آزمون، میانگین مقدار معیارPSNR را به همراه تصاویر اصلی و بازسازی شده ارائه دهید. در کدام حالت تصویر بازسازی شده بهتری تولید شده است؟ به

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

h.veisi@ut.ac.ir

نظر شما چرا؟

۲-۳ [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.(از بلوک ۸*۸ در این بخش استفاده کنید)

---

۳. (۴۰٪) **[پیاده‌سازی: پرسپترون چندلایه برای دسته‌بندی]** در این تمرین با استفاده از یک شبکه‌ی پرسپترون چندلایه قرار است مدلی را آموزش دهیم که تصاویر لباس‌ها و اشیای پوشیدنی را از مجموعه داده Fashion-MNIST تشخیص دهد. این مجموعه داده شامل تصاویر مربعی با ابعاد ۲۸×۲۸ پیکسل در سطح خاکستری است که هر تصویر به یکی از ۱۰ کلاس در جدول۱ قراردارد تعلق دارد. در این مجموعه داده ۶۰ هزار تصویر برای بخش آموزش (Train) و ۱۰ هزار تصویر برای 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.

https://github.com/zalandoresearch/fashion-mnist

h.veisi@ut.ac.ir

وَنام خدا

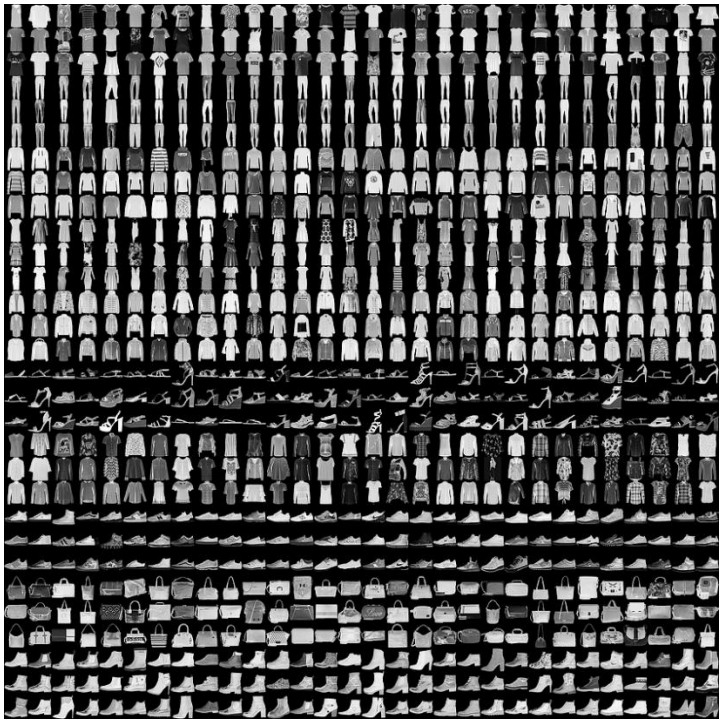
دانشکده سامانه‌های هوشمند

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

نیم‌سال اول ۱۴۰۵-۱۴۰۴

تمرین شماره ۲

تاریخ تحویل: ۱۴۰۴/۰۹/۰۵



شکل ۲ نمونه تصاویر از مجموعه دادگان Fashion-MNIST

جدول ۱ کلاس‌های موجود در مجموعه دادگان Fashion-MNIST

Description	Label
T-shirt/top	0
Trouser	1
Pullover	2
Dress	3
Coat	4
Sandal	5
Shirt	6
Sneaker	7
Bag	8
Ankle boot	9

h.veisi@ut.ac.ir

۳-۱ [۲۰٪] [آموزش auto-encoder] در این بخش به کمک شبکه‌ی پرسپترون چندلایه یک شبکه auto-encoder (مشابه با مثال فشرده‌سازی تصویر) برای همه تصاویر در مجموعه داده‌ی Fashion-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.**

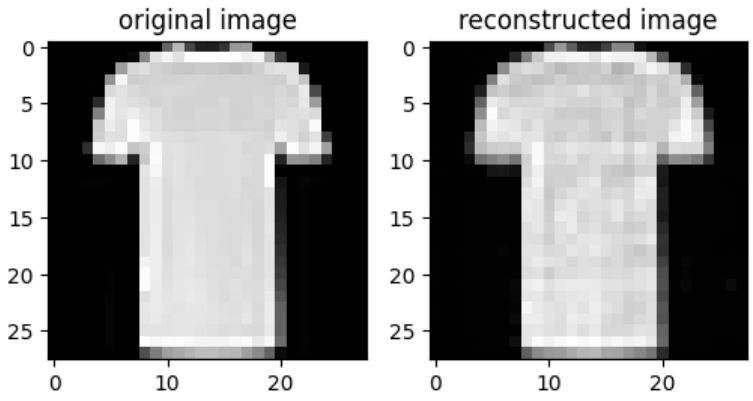
با داشتن الگوریتم پرسپترون چندلایه می‌توان از شبکه‌ای با ۷۸۴ نرون ورودی (پیکسل‌های یک تصویر) و لایه‌ی میانی با تعداد نرون کمتر (مثلاً ۶۴) و لایه‌ای خروجی با ۷۸۴ نرون (تصویر بازسازی شده) به عنوان auto-encoder استفاده کرد. با آموزش چنین شبکه‌ای می‌توانیم تصویر ورودی را در لایه‌ی وسط شبکه encode کنیم و 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 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 **Error! Reference source not found.**).



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

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.

h.veisi@ut.ac.ir

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 **Error! Reference source not found.** 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).

۳-۲ [۲۰٪] [دسته‌بندی لباس‌ها] با استفاده بهترین auto-encoder آموزش داده شده در سوال قبل، می‌خواهیم یک دسته‌بند مبتنی بر MLP برای دسته‌بندی لباس‌ها آموزش دهیم. برای این کار، داده‌های لباس‌ها 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 برچسب لباس یک نرون) و تعداد نرون دلخواه. این شبکه را با داده آموزش در بخش دادگان، آموزش داده و با داده آزمون ارزیابی کنید.

الف) شبکه MLP را با حداقل دو تعداد مختلف نرون در لایه مخفی این شبکه 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؟

h.veisi@ut.ac.ir

وَنام خدا

دانشکده سامانه‌های هوشمند

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

نیم‌سال اول ۱۴۰۵-۱۴۰۴

تمرین شماره ۲

تاریخ تحویل: ۱۴۰۴/۰۹/۰۵

---

۴. (۱۰٪) [**بدست آوردن قانون یادگیری**] در این قسمت قصد داریم، به روش جاری در شبکه‌های مدرن در عملیات پس انتشار خطا (Back-Propagation) در راستای به روزرسانی پارامترهای شبکه اشاره کنیم.

۴-۱ گراف پردازش (Computation-Graph) که در در شبکه‌های عمیق امروزی به کار می‌رود، چه استفاده‌ای دارد و مزیت استفاده از آن در مقایسه با روش‌های پیشین چیست؟ چگونگی استفاده از قانون زنجیر (Chain-Rule) را در این گراف شرح دهید.

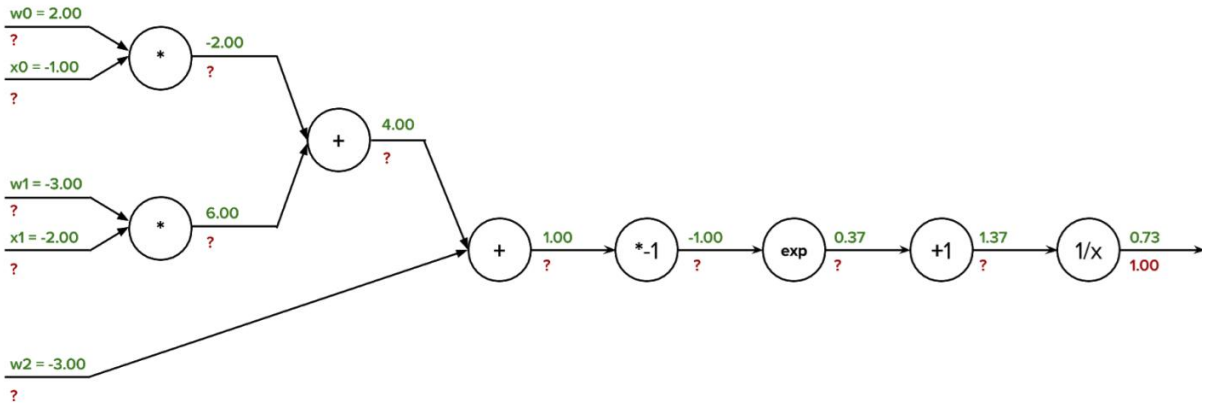
۴-۲ یک نورون منفرد از یک شبکه‌ی عصبی با تابع فعال‌سازی سیگموید را در نظر بگیرید. پیش‌فعال‌سازی این نورون به صورت زیر تعریف می‌شود:

$$z = w_0 x_0 + w_1 x_1 + w_2$$

و مقدار فعال‌سازی را به صورت زیر تعریف می‌کنیم:

$$a = \sigma(z) = \frac{1}{1 + e^{-z}}$$

نمودار محاسباتی این نورون در شکل زیر نشان داده شده است. در این نمودار، مقادیر فعال‌سازی جلورو (forward activations) بالای یال‌ها درج شده و همچنین گرادیان بالادستی (یعنی $\frac{\partial L}{\partial a}$ ، گرادیان تابع خطا نسبت به خروجی نورون) مشخص شده است.



h.veisi@ut.ac.ir

از اطلاعات داده‌شده استفاده کنید و مقادیر گرادیان‌های مجهول (برچسب‌گذاری‌شده با علامت سؤال در نمودار) را محاسبه کنید.

در نهایت، مشتق خروجی نورون a نسبت به ورودی‌ها و وزن‌ها را به‌صورت نمادین به‌دست آورید:

1. $\frac{\partial a}{\partial x_0}$
2. $\frac{\partial a}{\partial w_0}$
3. $\frac{\partial a}{\partial x_1}$
4. $\frac{\partial a}{\partial w_1}$
5. $\frac{\partial a}{\partial w_2}$

۳-۴ فرض کنید شبکه عصبی دو لایه مانند زیر داریم:

$$z_1 = W_1 x^{(i)} + b_1$$
$$a_1 = ReLU(z_1)$$
$$z_2 = W_2 a_1 + b_2$$
$$\widehat{y^{(i)}} = \sigma(z_2)$$
$$L^{(i)} = y^{(i)} \cdot \log\left(\widehat{y^{(i)}}\right) + \left(1 - y^{(i)}\right) \cdot \log\left(1 - \widehat{y^{(i)}}\right)$$
$$J = -\frac{1}{m}\sum_{i=1}^{m} L^{(i)}$$

توجه کنید که $x^{(i)}$ نشان دهنده یک نمونه ورودی با ابعاد $D_x \times 1$ است. همچنین $y^{(i)}$ برچسب یک نمونه است و به صورت اسکالر می‌باشد. دیتاست شامل m نمونه است. همچنین z_1 ابعاد $D_{a1} \times 1$ دارد.

- ابعاد W_1, b_1, W_2, b_2 را بنویسید.

h.veisi@ut.ac.ir

وَنام خدا

- نتیجه $\partial J/\partial \widehat{y^{(i)}}$ را بدست آوردید و آن را با δ_1 نشان دهید.
- نتیجه $\partial \widehat{y^{(i)}}/\partial z_2$ را بدست آوردید و آن را با δ_2 نشان دهید.
- نتیجه $\partial z_2/\partial a_1$ را بدست آوردید و آن را با δ_3 نشان دهید.
- نتیجه $\partial a_1/\partial z_1$ را بدست آوردید و آن را با δ_4 نشان دهید.
- نتیجه $\partial z_1/\partial W_1$ را بدست آوردید و آن را با δ_5 نشان دهید.
- نتیجه $\partial J/\partial W_1$ را با استفاده از نتایج قبل بدست آورید.

h.veisi@ut.ac.ir