



هدف این تمرین آشنایی با پیاده‌سازی الگوریتم‌های بهینه‌سازی در شبکه‌های عصبی، Backpropagation، رگرسیون و دسته‌بندی به کمک شبکه‌های عصبی چند لایه است.

مهلت تحویل این تمرین پایان روز جمعه ۱۷ آذر ۱۴۰۲ خواهد بود.

۱- الگوریتم‌های بهینه‌سازی و Backpropagation

۱-۱: برای هر یک از توابع زیر ابتدا Contour plot را رسم کنید. سپس با شروع از نقطه مشخص شده، ۵ قدم از الگوریتم کاهش گرادیان را با نرخ یادگیری ۰/۰۱ طی کنید (با استفاده از کد و نه به صورت دستی). همچنین مسیر (trajectory) طی شده توسط کاهش گرادیان را بر روی Contour plot مشخص کنید.

الف) تابع $4x_1^2 - x_1x_2 + 7 - x_1 - x_2$ با نقطه‌ی اولیه $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

ب) تابع $x_1^7 + x_1x_2 - x_1^3x_2^2$ با نقطه‌ی اولیه $\mathbf{x}_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

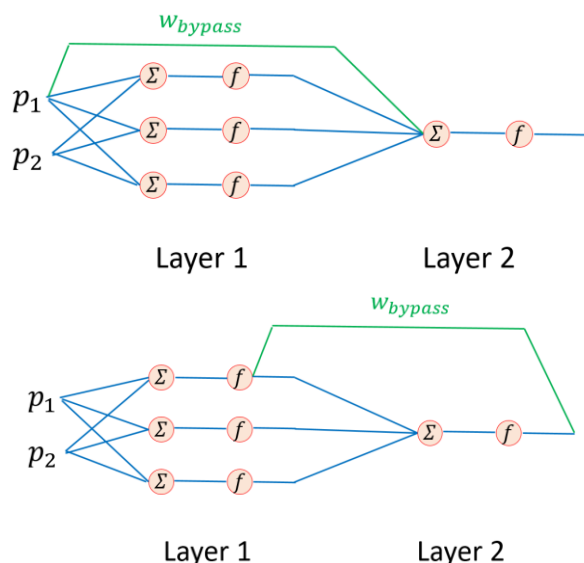
پ) تابع $[1 + (x_1 + x_2 - 5)^3][1 + (5x_1 - 2x_2)^4]$ با نقطه‌ی اولیه $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$

ت) تابع $3.5x_1^2 - 9x_1x_2 - 12.5x_2^2 + 7x_1 + 8x_2$ با نقطه‌ی اولیه $\mathbf{x}_0 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

۲-۱: با مشتق‌گیری از تابع خطای binary cross entropy و توابع فعال‌سازی tanh و Sigmoid کد متد backward آنها را در کلاس مربوطه بنویسید و سپس یک شبکه عصبی دو لایه تشکیل دهید که خطای آن binary cross entropy است و در لایه اول ۳ نورون با تابع فعال‌سازی tanh و در لایه دوم ۱ نورون با تابع فعال‌سازی Sigmoid دارد. تعداد ورودی‌های این شبکه هم ۲ تا است. همه پارامترهای این شبکه را به صورت تصادفی با توزیع یکنواخت از -۲ تا ۲ مقداردهی و گزارش کنید. در ادامه عملیات forward propagation و backward propagation را برای یک batch دلخواه متشکل از ۱۰ داده اعمال کنید. پارامترها را به روزرسانی و گزارش کنید.

۳-۱: مثال اسلاید ۳۰ از فصل ۶ را مطابق با آن چه در فرمولاسیون Backpropagation در درس خوانده‌اید، به طور کامل پیاده‌سازی کنید.

۱-۴: شکل‌های زیر دو شبکه عصبی را نشان می‌دهد که از اتصال bypass برخوردارند. اتصال bypass اتصالی است که به صورت مستقیم از یک لایه‌ها در لایه‌های ابتدایی (و یا ورودی) به لایه‌های انتهایی متصل می‌شود و پارامتری به نام w_{bypass} دارد. با فرض تابع فعال‌سازی Linear در لایه دوم و ReLU در لایه اول و نیز پارامترهای اولیه دلخواه، یک batch شامل چهار ورودی $p_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ ، $p_2 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$ ، $p_3 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ و $p_4 = \begin{bmatrix} -1 \\ -4 \end{bmatrix}$ با برچسب‌های به ترتیب ۲/۲، ۵/۷، ۲/۹ و ۱/۸ را به شبکه اعمال و خطای MSE شبکه را به دست بیاورید. در ادامه این خطا را برای به روزرسانی پارامترهای شبکه (از جمله w_{bypass}) به کار ببرید و مقادیر نهایی پارامترها را گزارش دهید.



۲- رگرسیون

۱-۲: سوال ۴-۱ تمرین اول را یکبار دیگر با استفاده از بهینه ساز SGD انجام دهید و خطای MSE را گزارش کنید.

۲-۲: مجموعه داده housing که برای پیش‌بینی قیمت مسکن به کار می‌رود در پیوست داده شده است. این مجموعه داده از ۱۳ ستون به عنوان ویژگی‌ها (features) و یک ستون به عنوان برچسب تشکیل شده است. هدف، آموزش یک شبکه عصبی است که بتواند نگاشت مناسبی از فضای ویژگی ۱۳ بعدی به فضای برچسب داشته باشد. طبیعی است که شبکه طراحی شده باید ۱۳ ورودی و ۱ خروجی داشته باشد. هر یک از سه شبکه عصبی زیر را به تعداد ۵۰ epoch با اندازه batch برابر ۱۶ بر روی مجموعه آموزشی (housing_train.txt) آموزش دهید و نمودار تغییرات خطای هر مدل بر روی مجموعه آموزشی را بر حسب تعداد epoch رسم کنید. در انتها خطای مدل‌ها را بر روی مجموعه آزمایشی (housing_test.txt) به دست آورید.

الف) یک شبکه عصبی تک لایه

ب) یک شبکه عصبی دو لایه با ۵ نورون ReLU در لایه مخفی

پ) یک شبکه عصبی دو لایه با ۱۰ نورون Sigmoid در لایه مخفی

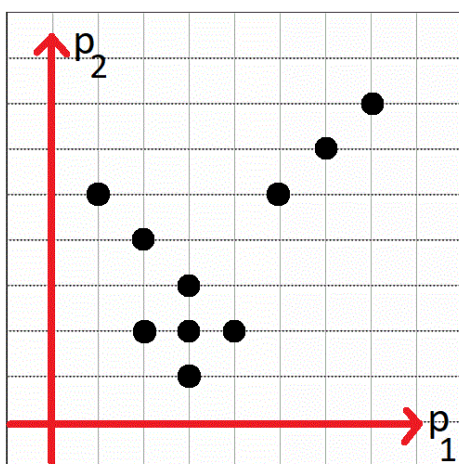
ت) یک شبکه عصبی چهار لایه با ۱۰ نورون ReLU در لایه مخفی اول و ۵ نورون Sigmoid در لایه مخفی دوم

۳-۲: ابتدا تعداد ۲۰۰ داده آموزشی را از تابع $\sin\left(\frac{i\pi}{4}p\right) - 1$ در بازه $-3 \leq p \leq 3$ به صورت تصادفی با توزیع یکنواخت نمونه‌برداری کنید تا مجموعه آموزشی تشکیل شود. سپس یک شبکه عصبی سه لایه با تعداد ۴ نورون در لایه مخفی و تابع فعال‌سازی Sigmoid تشکیل دهید و آن را بر روی مجموعه آموزشی به تعداد ۱۰ epoch آموزش دهید. در انتها نمودار پاسخ شبکه آموزش داده شده را به ازای $-4 \leq p \leq 4$ رسم کنید. تمام مراحل را به ازای $i = 1, 2, 4, 8$ انجام دهید و تحلیل کنید که چگونه پیچیدگی تابعی که به دنبال تقریب آن هستیم باعث افزایش یا کاهش خطای تقریب می‌شود.

۴-۲: ابتدا تعداد ۲۰۰ داده آموزشی را از تابع $\sin\left(\frac{3\pi}{2}p\right) - 1$ در بازه $-2 \leq p \leq 2$ به صورت تصادفی با توزیع یکنواخت نمونه‌برداری کنید تا مجموعه آموزشی تشکیل شود. سپس یک شبکه عصبی سه لایه با تعداد i نورون در لایه مخفی و تابع فعال‌سازی Sigmoid تشکیل دهید و آن را بر روی مجموعه آموزشی به تعداد ۱۰ epoch آموزش دهید. در انتها نمودار پاسخ شبکه آموزش داده شده را به ازای $-4 \leq p \leq 4$ رسم کنید. تمام مراحل را به ازای $i = 2, 3, 4, 5$ انجام دهید و تحلیل کنید که چگونه پیچیدگی مدل انتخابی باعث افزایش یا کاهش خطای تقریب می‌شود.

۳- دسته‌بندی (Classification)

۱-۳: یک شبکه عصبی چند لایه طراحی کنید که عملیات دسته‌بندی در داده‌های زیر را انجام دهد. وزن‌ها و بایاس‌های مناسبی برای این شبکه بدون انجام backpropagation به دست آورید (همه توابع فعال‌سازی باید Unit Step باشد). فرض کنید که داده‌های کلاس ۱ دایره‌های مشکی رنگ و داده‌های کلاس ۲ بقیه نقاط صفحه که مختصات عدد صحیح دارند، باشد.



۲-۳: بر روی مجموعه داده MNIST که برای دسته‌بندی ارقام دستنویس ۰ تا ۹ به کار می‌رود و در پیوست این تمرین موجود است، سه شبکه عصبی زیر را به تعداد ۱۰۰ epoch آموزش دهید و صحت (Accuracy) آن‌ها بر روی مجموعه

آزمایشی را با هم مقایسه کنید. برای هر یک نمودار تغییرات Accuracy مجموعه آموزشی بر حسب تعداد update پارامترها را نیز رسم کنید.

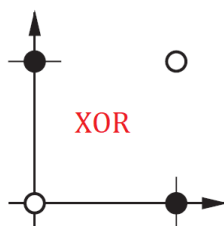
الف) یک شبکه تک لایه

ب) یک شبکه ۲ لایه که لایه مخفی آن متشکل از ۳۰ نورون و تابع فعال‌سازی ReLU است.

پ) یک شبکه ۳ لایه که لایه مخفی اول آن متشکل از ۳۰ نورون ReLU و لایه مخفی دوم آن متشکل از ۱۰ نورون Sigmoid است.

۳-۳: مسئله ۲-۲ تمرین اول را آموزش دهید. نمودار هزینه-تکرار را برای هر کدام از بخش‌های مورد نظر رسم کنید. در نهایت، تحلیل کنید و بهترین گزینه را برای این مسئله انتخاب کنید.

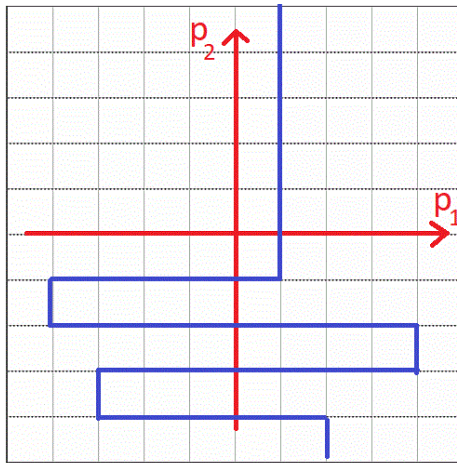
۳-۴: یک شبکه دو لایه برای دسته‌بندی ۴ داده زیر (XOR) طراحی کنید. تعداد نورون‌های لایه مخفی را دلخواه و تابع فعال‌سازی در لایه مخفی و آخر را به ترتیب ReLU و Softmax در نظر بگیرید.



الف) این شبکه را به تعداد ۱ epoch بر روی این چهار داده با $\text{batch size}=1$ و نرخ یادگیری ۰/۱ آموزش دهید و سپس Accuracy آن را بر روی این چهار داده به دست آورید.

ب) تعداد epoch های آموزش را به ۱۰ افزایش دهید. آیا تغییری در Accuracy مدل مشاهده می‌کنید؟ چرا؟

۳-۵: فرض کنید که هدف ما دسته‌بندی دو کلاسی داده‌های زیر باشد که در صفحه مختصات قرار دارند. این داده‌ها به گونه‌ای در صفحه قرار گرفته‌اند که داده‌های دسته ۱ همگی در سمت چپ خط آبی رنگ و داده‌های دسته ۲ همگی در سمت راست این خط هستند. خط آبی در واقع مرز تصمیم واقعی و ایده آل بین داده‌های دو دسته است و خطی است که هم از بالا و هم از پایین تا بینهایت به صورت مستقیم ادامه دارد و فقط در بخشی از صفحه که در شکل نشان داده شده است، دچار شکستگی می‌شود.



الف) تعداد ۱۰۰ داده آموزشی به صورت تصادفی به گونه‌ای تولید کنید که همگی در صفحه نشان داده شده در بالا قرار بگیرند. سپس برچسب واقعی این داده‌ها را مشخص و همگی را در صفحه با دو رنگ مجزا نشان دهید.

ب) دو شبکه عصبی مختلف یکی دو لایه با ۲ نورون ReLU در لایه مخفی و دیگری دو لایه با ۶ نورون ReLU در لایه مخفی تشکیل دهید. شبکه اول را بر روی کل مجموعه داده به اندازه ۱۰ epoch آموزش دهید.

پ) شبکه اول را بر روی کل مجموعه آموزشی تست کنید و آن داده‌هایی که توسط شبکه اول به درستی دسته‌بندی نمی‌شوند به همراه ۱۰ تا از داده‌هایی که به درستی دسته‌بندی می‌شوند (به صورت تصادفی) را به عنوان مجموعه آموزشی به شبکه دوم بدهید و شبکه دوم را به تعداد ۱۰ epoch بر روی مجموعه حاصل آموزش دهید و Accuracy آن را گزارش کنید.

ت) تحقیق کنید که این ایده به چه منظوری در یادگیری ماشین استفاده می‌شود؟ آیا این ایده به افزایش دقت دسته‌بندی نهایی کمک می‌کند؟ بهتر نبود که از همان ابتدا فقط شبکه دوم بر روی همه داده‌ها آموزش ببیند و اصلاً از شبکه اول استفاده نشود؟ توضیح دهید.

مواردی که باید تحویل داده شود:

- فایل (یا فایل‌های) کد برنامه مورد استفاده برای پیاده‌سازی تمرین با فرمت ipynb. در یک پوشه به نام Code
 - استفاده از هر گونه کتابخانه آماده (به جز Numpy و Pandas و Matplotlib) ممنوع است و در صورت استفاده هیچ نمره‌ای به تمرین تعلق نمی‌گیرد.
- فایل گزارش با نام Doc.pdf

فایل‌های کد و گزارش را به صورت یک فایل فشرده در قالب zip و با نام NN_Name_Family_HW2 در سایت کوئرا بارگذاری کنید (به جای Name نام و به جای Family نام خانوادگی خود را قرار دهید).

مهلت تحویل این تمرین تا پایان روز جمعه ۱۷ آذر خواهد بود.

موفق باشید