

دانشگاه صنعتی خواجه نصیرالدین طوسی
دانشکده مهندسی برق - گروه مهندسی کنترل

یادگیری ماشین پاسخ تکلیف امتیازی ۱

نام و نام خانوادگی	سید محمد حسینی
شماره دانشجویی	۹۹۰۱۳۹۹
تاریخ	مهرماه ۱۴۰۲



فهرست مطالب

۶	۱	سوال اول
۶	۱.۱	فرض کنید در یک مساله‌ی طبقه‌بندی دو کلاسه، دو لایه‌ی انتهایی شبکه‌ی شکار فعال ساز ReLU و سیگموید است. چه اتفاقی می‌افتد.
۶	۲.۱	یک جایگزین برای ReLU در معادله ۱ آورده شده. ضمن محاسبه‌ی گرادیان آن حداقل یک مزیت آن نسبت به ReLU را توضیح دهید.
۶	۳.۱	به کمک یک نورون ساده یا پرسپرون یا نورون McCulloch-Pitts شبکه‌ی طراحی کنید که بتواند ناحیه‌های شورزده داخل مثلثی که در نمودار شکل ۱(آ) نشان داده شده را از سایر نواحی تفکیک کند. پس از انجام مرحله طراحی شبکه (که می‌تواند به صورت دستی انجام شود، برنامه‌ای که در این دفترچه کد و در کلاس برای نورون McCulloch-Pitts آماده شده را به گونه‌ای توسعه دهید که ۲۰۰۰ نقطه رندوم تولید کند و آن‌ها را به عنوان ورودیه شبکه طراحی شده توسط شما دهد و نقاطی که خروجی «۱» تولید می‌کنند را با رنگ سبز و نقاطی که خروجی «۰» تولید می‌کنند را با رنگ قرمز نشان دهد. خروجی تولید شده توسط برنامه شما باید به صورتی داده شده است باشد) به محدوده عددی محورهای x و y دقت کنید. اثر اضافه کردن دو تابع فعال ساز مختلفه فرآیند تصمیم‌گیری را هم بررسی کنید.
۷	۲	سوال دوم
۹	۱.۲	دیتاست CWRU که Bearing در «مینی پروژه شماره یک» با آن آشنا شدید را به خاطر آورید. علاوه بر دو کلاسی که در آن مینی پروژه در نظر گرفتید، با مراجعه به صفحه داده‌های عیب در حالت دو ۱۲k، کلاس دیگر نیز از طریق فایل‌های ۲B_۰۰۷_X و ۲B_۰۰۷_X@۶_۰۰۷_OR کنید. با انجام این کار یک کلاس داده سالم و سه کلاس از داده‌های دارای سه عیب متفاوت خواهید داشت. در مورد این که هر فایل مربوط به چه نوع عیبی است به صورت کوتاه توضیح دهید. سپس در ادامه، تمام کارهایی که در بخش «۲ سوال دوم» مینی پروژه یک «برای استخراج ویژگی و آماده‌سازی دیتا انجام داده بودید را روی دیتاست جدید خود پیاده‌سازی کنید. در قسمت تقسیم‌بندی داده‌ها، یک بخش برای «اعتبارسنجی» به بخش‌های «آموزش» و «آزمون» اضافه کنید و توضیح دهید که کاربرد این بخش چیست.
۹	۲.۲	یک مدل (Perceptron Multi-Layer ساده MLP) با ۲ لایه پنهان یا بیش‌تر بسازید. بخشی از داده‌های آموزش را برای اعتبارسنجی کنار بگذارید و با انتخاب بهینه‌ساز و تابع اتلاف مناسب، مدل را آموزش دهید. نمودارهای اتلاف و مربوط Accuracy به آموزش و اعتبارسنجی را رسم و نتیجه را تحلیل کنید. نتیجه تست مدل روی داده‌های آزمون را با استفاده ماتریس درهم‌ریختگی و نشان classification_report داده و نتایج به صورت دقیق تحلیل کنید.
۹	۳.۲	فرآیند سوال قبل را با یک بهینه‌ساز و تابع اتلاف جدید انجام داده و نتایج را مقایسه و تحلیل کنید. بررسی کنید که آیا تغییر تابع اتلاف می‌تواند در نتیجه اثرگذار باشد؟
۹	۴.۲	یچ Cross-validation K-Fold Stratified و Cross-validation K-Fold. در مورد ۴ دهید. سپس با ذکر دلیل، یکی از این روش‌ها را انتخاب کرده و بخش «۲ این سوال را با آن پیاده‌سازی کنید و نتایج خود را تحلیل کنید
۹	۳	عنوان سوال سوم



۱۱	۴	عنوان سوال چهارم
۱۱	۵	عنوان سوال پنجم
۱۱	۶	عنوان سوال ششم
۱۱	۱.۶	عنوان بخش اول سوال ششم
۱۱	۲.۶	عنوان بخش دوم سوال ششم
۱۱	۷	عنوان سوال هفتم
۱۲	۸	عنوان سوال هشتم
۱۳	۱.۸	عنوان بخش اول سوال هشتم
۱۳	۲.۸	عنوان بخش دوم سوال هشتم
۱۳	۳.۸	عنوان بخش سوم سوال هشتم
۱۳	۹	ضمیمه



فهرست تصاویر

۷	شکل شماره ۱	۱
۸	شکل شماره ۲	۲
۹	شکل شماره ۳	۳
۱۰	شکل شماره ۴	۴



فهرست جداول

۱	جدول شماره ۱	۱۱
---	--------------	----



فهرست برنامه‌ها

۱۲ (Python) Caption My	۱
۱۲ (MATLAB) Caption My	۲
۱۲ (C++) Caption My	۳
۱۲ (C) Caption My	۴



۱ سوال اول

اگر سوال بخش بندی شده نباشد، پاسخ آن در این قسمت نوشته می شود.

۱.۱ فرض کنید در یک مساله ی طبقه بندی دو کلاسه ، دو لایه ی انتهایی شبکه ی شكا فعال ساز ReLU و سیگموید است. چه اتفاقی می افتد.

-خروجی Relu نامحدود است (از ۰ تا مثبت بی نهایت) این در حالی است که ما برای طبقه بندی خروجی احتمالاتی (۰ تا ۱) مطلوب ماست.

-شبکه ممکن است در حین آموزش مشکل همگرایی داشته باشد زیرا ReLU می تواند مقادیر نامحدود مثبت تولید کند. این می تواند منجر به شیب های بزرگ شود.

-خروجی لایه نهایی باید به عنوان احتمال کلاس تفسیر شود. استفاده از ReLU در لایه نهایی، تفسیر خروجی را دشوار می کند، زیرا می تواند هر مقدار غیر منفی تولید کند.

-برخی توابع اتلاف مثل bce خرجی احتمالاتی نیاز دارند و این با relu سازگار نمی باشد.

۲.۱ یک جایگزین برای ReLU در معادله ۱ آورده شده. ضمن محاسبه ی گرادیان آن حداقل یک مزیت آن نسبت به ReLU را توضیح دهید.

$$\text{ELU}(x) = \begin{cases} x & \text{if } x \geq 0 \\ \alpha(e^x - 1) & \text{if } x < 0 \end{cases} \quad (۱)$$

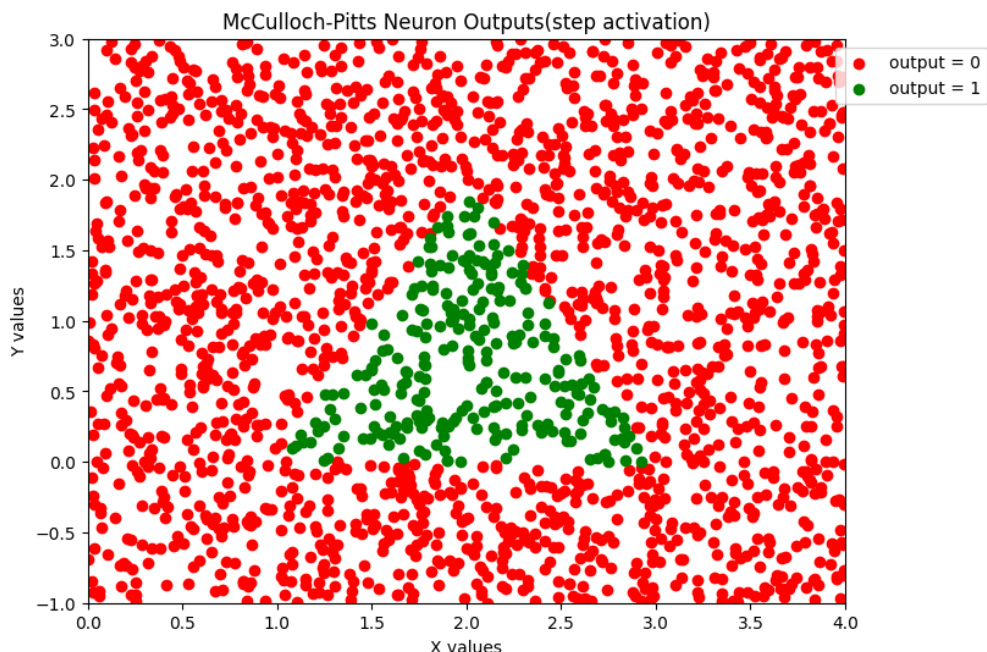
$$\frac{d}{dx}\text{ELU}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ \alpha e^x & \text{if } x < 0 \end{cases} \quad (۲)$$

-از انجایی که خروجی ReLU به ازای مقادیر منفی صفر است در هنگام آموزش ممکن است وزن ها گیر کنند و تغییری در آن ها ایجاد نشود این در حالی است که ELU این مشکل را ندارد.

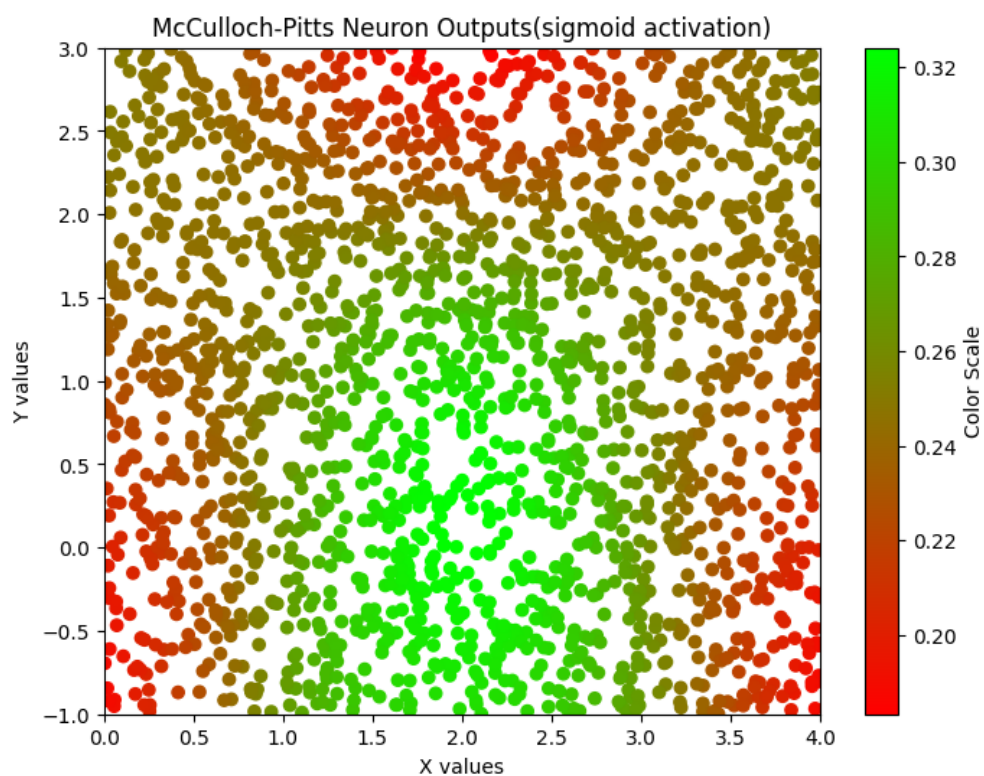
-از آن جایی که خروجی های ELU نسبت به ReLU به بیشتر حول صفر اند در هنگام آموزش نیاز نیست زیاد bias تغییر کند و همگرایی سریعتر انجام می شود.



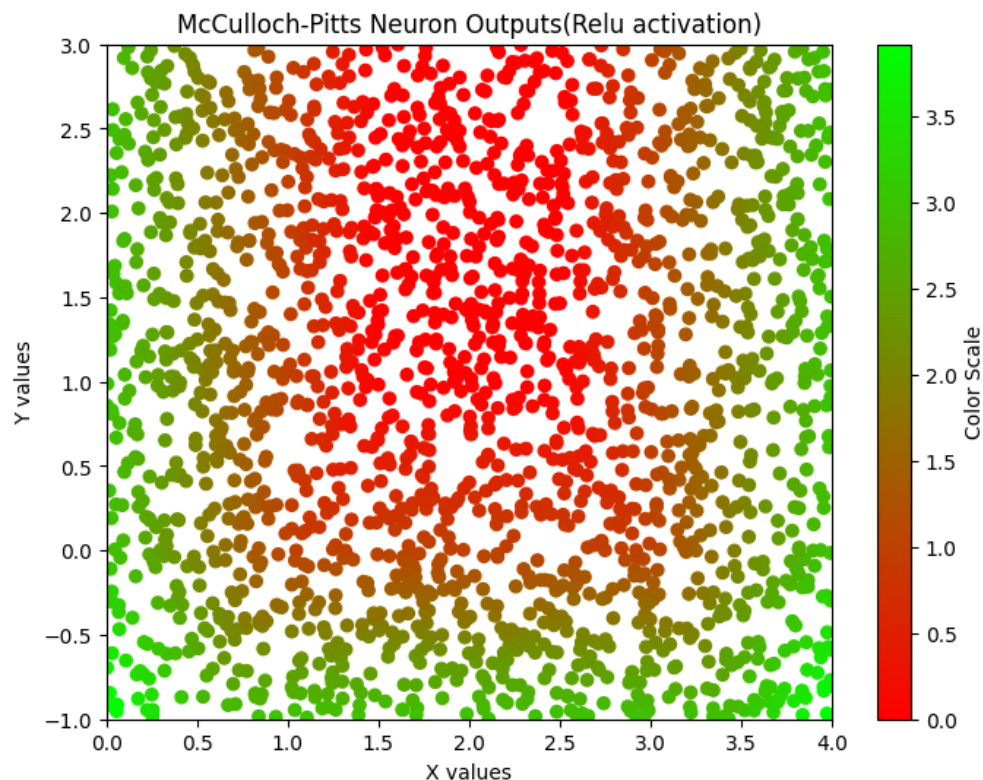
۳.۱ به کمک یک نورون ساده یا پرسپرون یا نورون McCulloch-Pitts شبکه ای طراحی کنید که بتواند ناحیه هاشورزده داخل مثلثی که در نمودار شکل ۱(آ) نشان داده شده را از سایر نواحی تفکیک کند. پس از انجام مرحله طراحی شبکه (که می تواند به صورت دستی انجام شود)، برنامه ای که در این دفترچه کد و در کلاس برای نورون آموخته McCulloch-Pitts اید را به گونه ای توسعه دهید که ۲۰۰۰ نقطه رندوم تولید کند و آن ها را به عنوان ورودیه شبکه طراحی شده توسط شما دهد و نقاطی که خروجی «۱» تولید می کنند را با رنگ سبز و نقاطی که خروجی «۰» تولید می کنند را با رنگ قرمز نشان دهد. خروجی تولید شده توسط برنامه شما باید به صورتی داده شده است باشد (به محدوده عددی محورهای x و y دقت کنید). اثر اضافه کردن دو تابع فعال ساز مختلفه فرآیند تصمیم گیری را هم بررسی کنید.



شکل ۱: شکل شماره ۱



شکل ۲: شکل شماره ۲

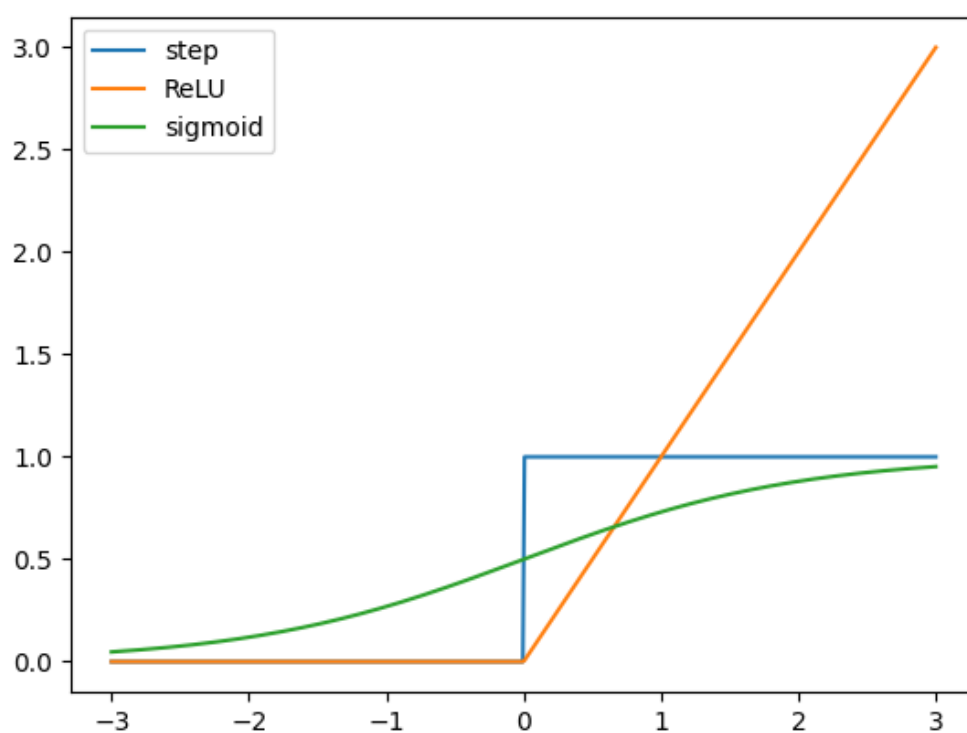


شکل ۳: شکل شماره ۳

۲ سوال دوم

۱.۲ دیتاست CWRU که Bearing در «مینی پروژه شماره یک» با آن آشنا شدید را به خاطر آورید. علاوه بر دو کلاسی که در آن مینی پروژه در نظر گرفتید، با مراجعه به صفحه داده های عیب در حالت دو کلاس، کلاس دیگر نیز از طریق فایل های $2B_{0.07_X}$ اضافه $6_X@0.07_OR$ کنید. با انجام این کار یک کلاس داده سالم و سه کلاس از داده های دارای سه عیب متفاوت خواهید داشت. در مورد این که هر فایل مربوط به چه نوع عیبی است به صورت کوتاه توضیح دهید. سپس در ادامه، تمام کارهایی که در بخش «۲ سوال دوم» مینی پروژه یک» برای استخراج ویژگی و آماده سازی دیتا انجام داده بودید را روی دیتاست جدید خود پیاده سازی کنید. در قسمت تقسیم بندی داده ها، یک بخش برای «اعتبارسنجی» به بخش های «آموزش» و «آزمون» اضافه کنید و توضیح دهید که کاربرد این بخش چیست

۲.۲ یک مدل (Perceptron Multi-Layer ساده MLP) با ۲ لایه پنهان یا بیش تر بسازید. بخشی از داده های آموزش را برای اعتبارسنجی کنار بگذارید و با انتخاب بهینه ساز و تابع اتلاف مناسب، مدل را آموزش دهید. نمودارهای اتلاف و مربوط Accuracy به آموزش و اعتبارسنجی را رسم و نتیجه را تحلیل کنید. نتیجه تست مدل روی داده های آزمون را با استفاده ماتریس درهم ریختگی و نشان classification_report داده و نتایج به صورت دقیق تحلیل کنید



شکل ۴: شکل شماره ۴