

دانشگاه صنعتی خواجه نصیرالدین طوسی
دانشکده مهندسی برق

یادگیری ماشین

پاسخ مینی پروژه‌ی شماره ۲

نام و نام خانوادگی	مهدی خداپنده لو
شماره دانشجویی	۴۰۱۰۵۳۱۴
تاریخ	مهرماه ۱۴۰۳



فهرست مطالب

۵	۱ سوال اول
۵	۱.۱ فرض کنید در یک مساله‌ی طبقه‌بندی دو کلاسه، دو لایه‌ی انتهایی شبکه‌ی شکار فعال ساز ReLU و سیگموید است. چه اتفاقی می‌افتد.
۵	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-Pits شبکه‌ی طراحی کنید که ...
۶	۲ سوال دوم
۶	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 ...
۶	۲.۲ یک مدل ساده MLP با ۲ لایه پنهان یا بیش تر بسازید.
۷	۳.۲ فرآیند سوال قبل را با یک بهینه ساز 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؟
۷	۳ سوال چهارم



فهرست تصاویر

۶	خروجی شبکه یا فعال ساز پله	۱
۷	خروجی شبکه با فعال ساز سیگموئید	۲
۸	خروجی شبکه با فعال ساز رلو	۳
۹	مقایسه ی توابع فعال ساز مختلف	۴
۹	نمودار خطا برای شبکه عصبی با فعال ساز سیگموئید و بهینه ساز گرایان نزولی	۵
۱۰	نمودار خطا برای شبکه عصبی با فعال ساز رلو و بهینه ساز گرایان نزولی	۶
۱۰	نمودار خطا برای شبکه عصبی با فعال ساز سیگموئید و بهینه ساز آدام	۷
۱۱	نمودار خطا برای شبکه عصبی با فعال ساز رلو و بهینه ساز آدام	۸
۱۲	هیستوگرام داده ها	۹
۱۳	نمودار هیت مپ	۱۰
۱۴	ماتریس درهم ریختگی	۱۱



فهرست جداول

۸ report classification	۱
---	-----------------------------	---



فهرست برنامه‌ها



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 شکار 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.

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

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

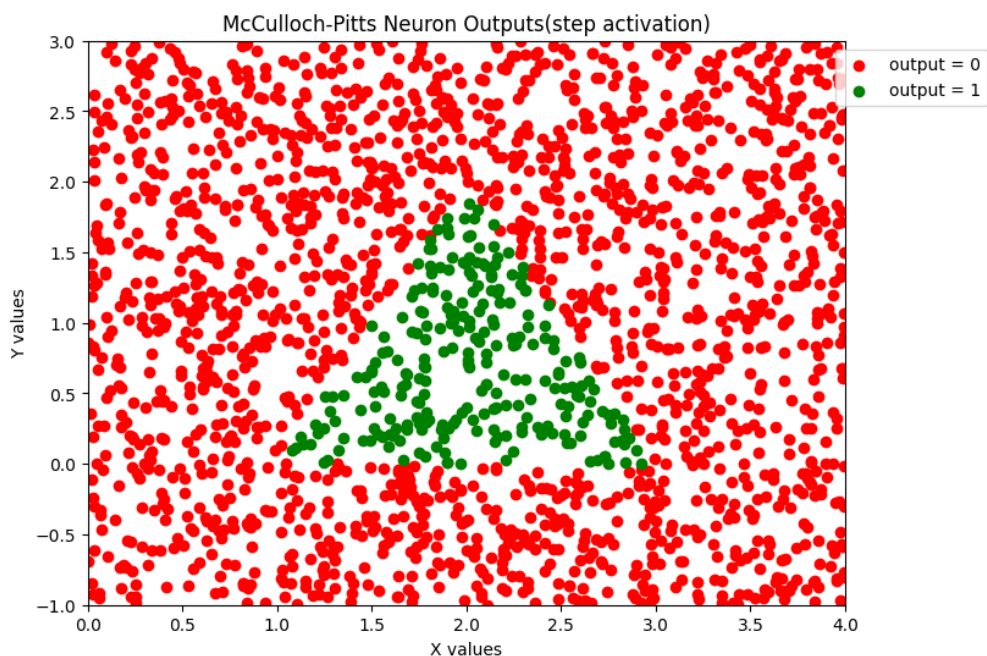
-از انجایی که خروجی ReLU به ازای مقادیر منفی صفر است در هنگام آموزش ممکن است وزن ها گیر کنند و تغییری در آن ها ایجاد نشود این در حالی است که ELU این مشکل را ندارد.

-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-Pits شبکه ای طراحی کنید که ...

شکل ۱ نتیجه‌ی تفکیک دو ناحیه با استفاده از نورون McCulloch-Pits را نمایش می دهد. همان طور که از تصویر مشخص است ناحیه‌ی مثلثی به خوبی تفکیک شده است.



شکل ۱: خروجی شبکه یا فعال ساز پله

تصویر ۲ خروجی همان شبکه است با این تفاوت که تابع فعال ساز سیگموئید استفاده شده. همان طوط مه از تصویر مشخص است تفکیک نتیجه‌ی مطلوبی ندارد.

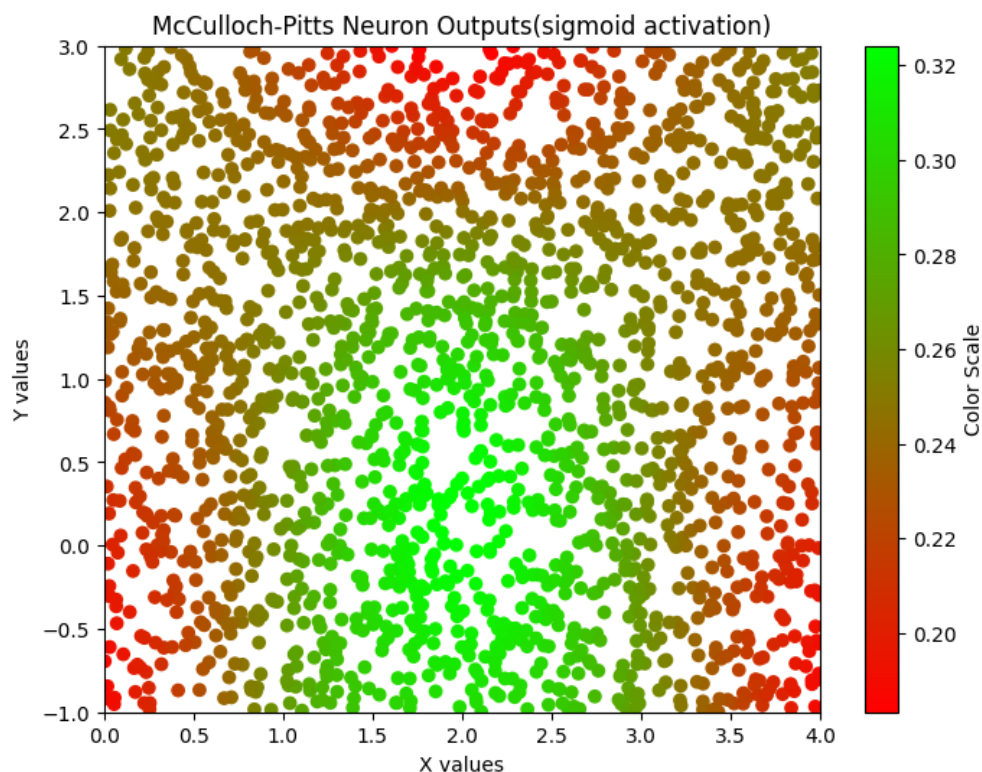
تصویر ۳ نیز خروجی شبیه با تابع فعال ساز رلو را نمایش می دهد که خروجی این تابع نیز مطلوب نمی باشد.

۲ سوال دوم

۱.۲ دیتاست CWRU که Bearing در مینی پروژه شماره یک با آن آشنا شدید را به خاطر آورید. علاوه بر دوکلاسی که در آن مینی پروژه در نظر ...

۲.۲ یک مدل ساده MLP با ۲ لایه پنهان یا بیش تر بسازید.

تصویر ۵ loss را برای یک شبکه‌ی عصبی با دولایه‌ی پنهان نمایش می دهد. در این شبکه در هر لایه‌ی پنهان ۲۵ نورون وجود دارد و از بهینه ساز گرادینت کاهشی استفاده شده همچنین تابع فعال ساز لایه‌های پنهان سیگموئید و فعال ساز لایه‌ی آخر softmax است. Accuracy در این حالت 46.5% است. نرخ یادگیری در این حالت برابر با 10-5 است.



شکل ۲: خروجی شبکه با فعال ساز سیگموئید

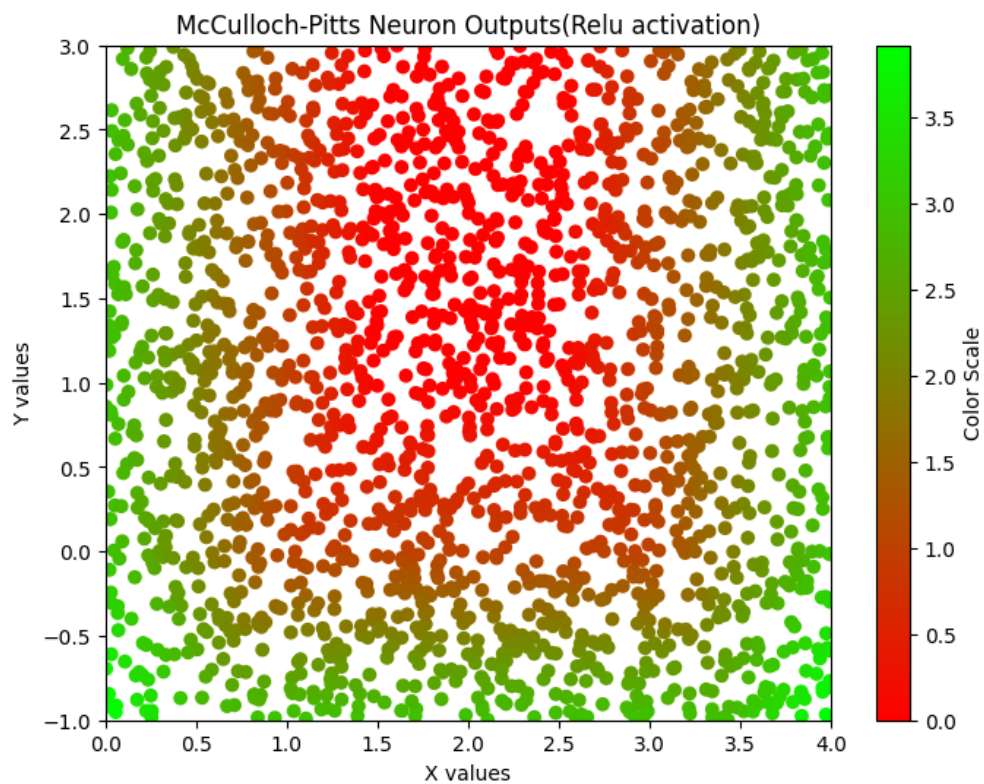
تصویر ۶ loss را برای همان شبکه نمایش می دهد با این تفاوت که در این جا تابع فعال ساز لایه های پنهان relu می باشد. Accuracy در این حالت 51% است. نرخ یادگیری در این حالت برابر با 5-10 است.

۳.۲ فرآیند سوال قبل را با یک بهینه ساز و تابع اتلاف جدید انجام داده و نتایج را مقایسه و تحلیل کنید. بررسی کنید که آیا تغییر تابع اتلاف می تواند در نتیجه اثرگذار باشد؟

تصویر ۷ loss را برای حالتی که از بهینه ساز آدام استفاده شده نمایش می دهد. در این شبکه فعال ساز لایه های پنهان سیگموئید و فعال ساز لایه ی آخر softmax است. Accuracy در این حالت 47% است. نرخ یادگیری در این حالت برابر با 0.01 است. تصویر ۸ نیز loss را برای حالتی که از بهینه ساز آدام استفاده شده نمایش می دهد با این تفاوت که فعال ساز لایه های پنهان رلو است. Accuracy در این حالت 51.5% است. نرخ یادگیری در این حالت برابر با 0.01 است.

۳ سوال چهارم

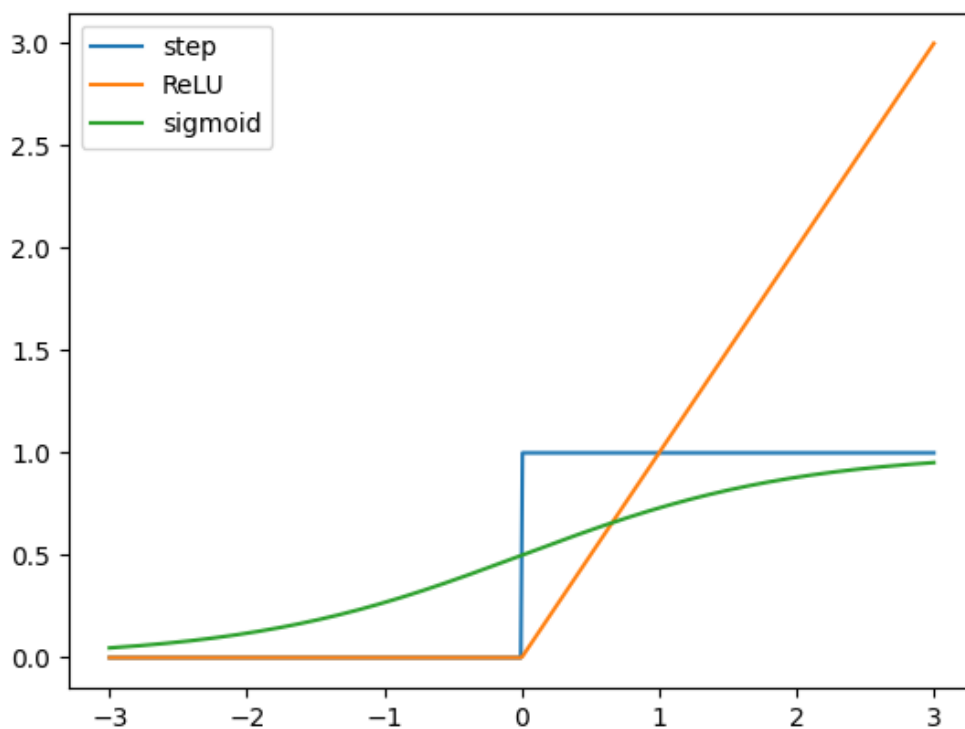
تصویر ۹ مربوط به هیستوگرام داده ها و تصویر ۱۰ نمودار هیت مپ می باشد. تصویر ۱۱ ماتریس در هم ریختگی را نمایش می دهد. accuracy این مدل ۷۴ درصد می باشد.



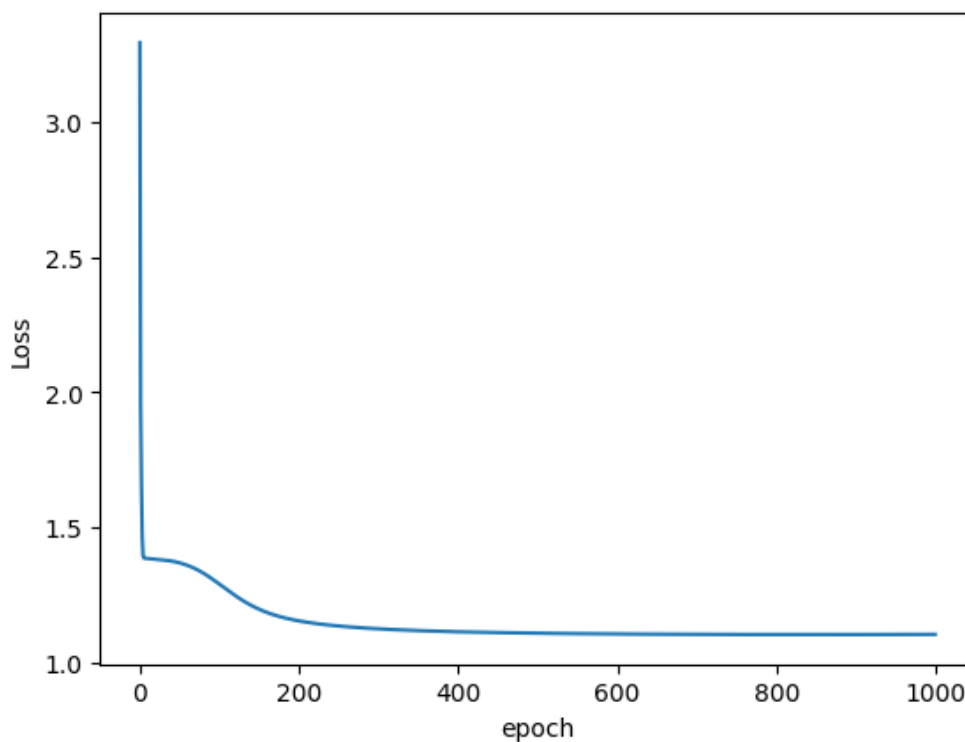
شکل ۳: خروجی شبکه با فعال ساز رلو

جدول ۱: report classification

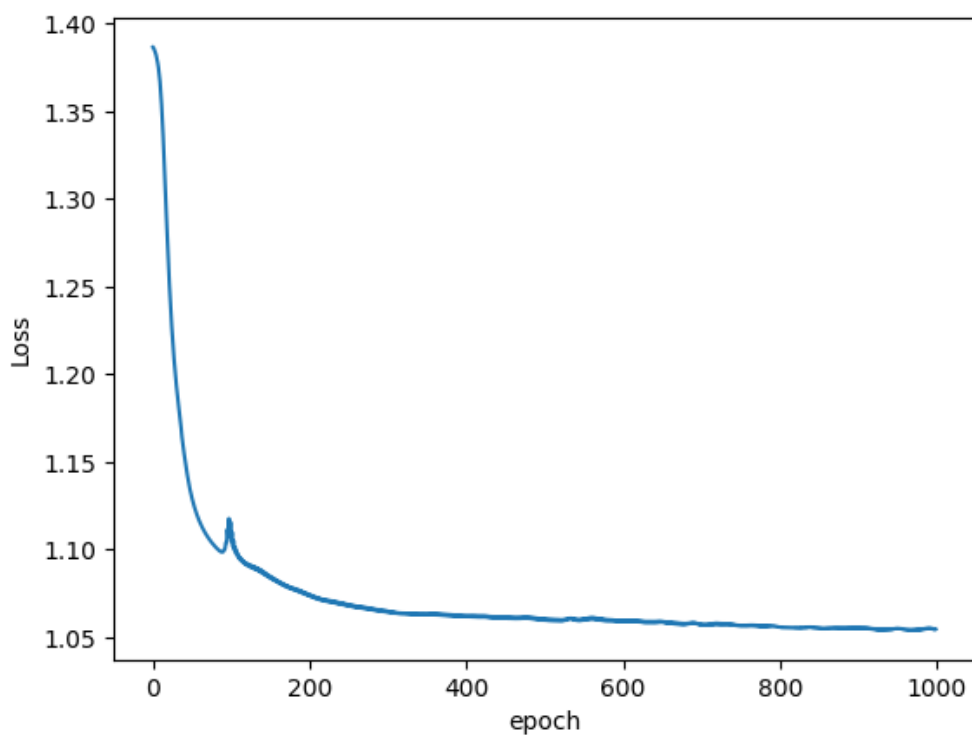
	precision	recall	f1-score	support
class 0	0.74	0.70	0.72	96
class 1	0.75	0.78	0.76	109
accuracy			0.74	205
macro avg	0.74	0.74	0.74	205
weighted avg	0.74	0.74	0.74	205



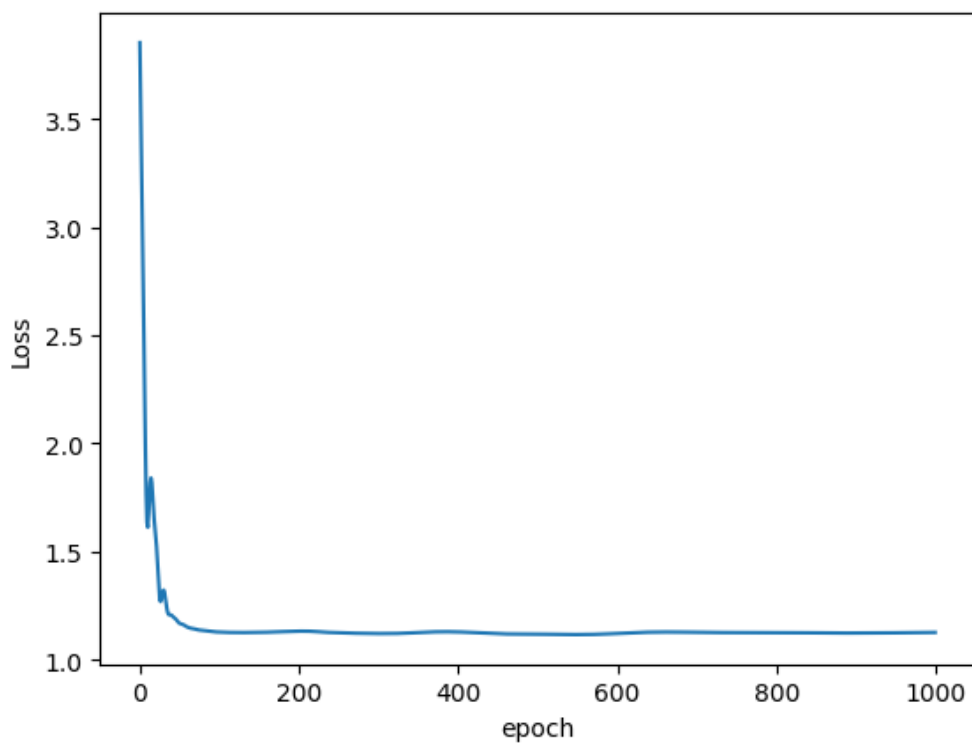
شکل ۴: مقایسه ی توابع فعال ساز مختلف



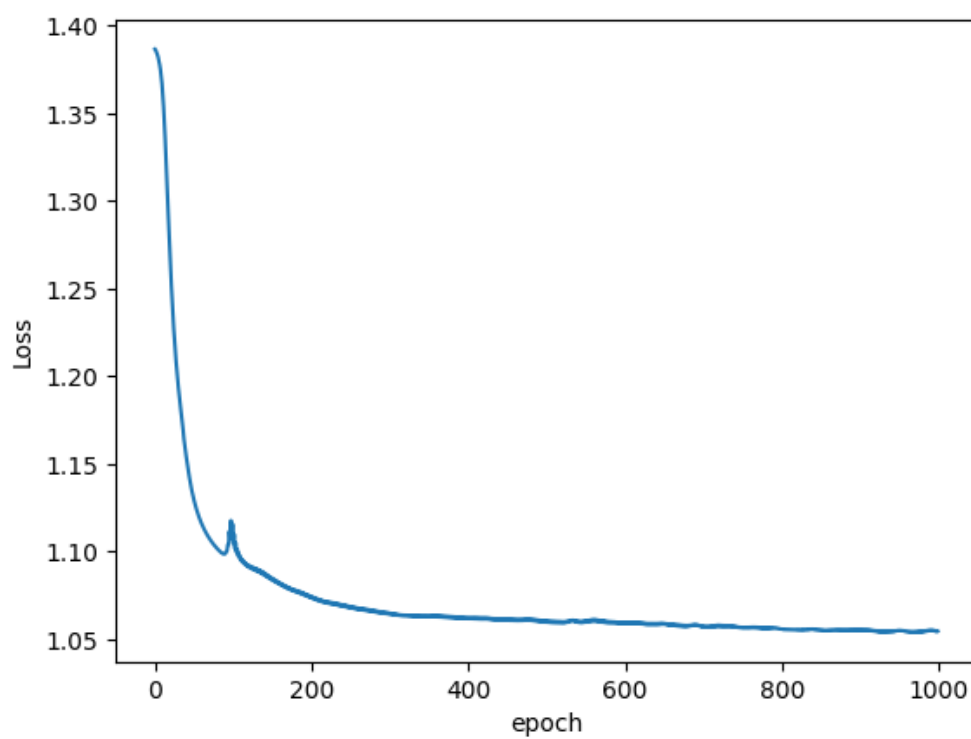
شکل ۵: نمودار خطا برای شبکه عصبی با فعال ساز سیگموید و بهینه ساز گرایان نزولی



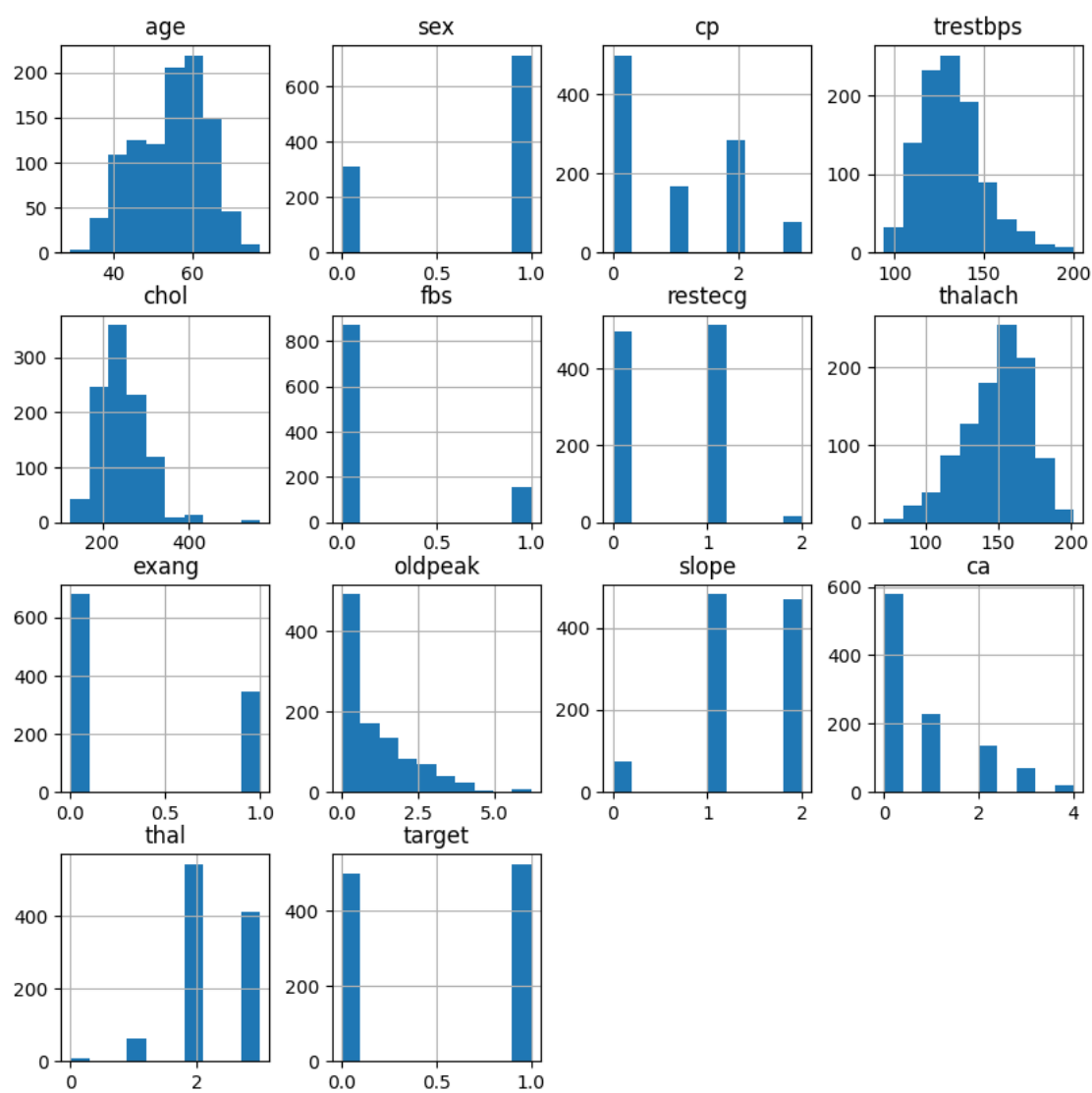
شکل ۶: نمودار خطا برای شبکه عصبی با فعال ساز رلو و بهینه ساز گرایان نزولی



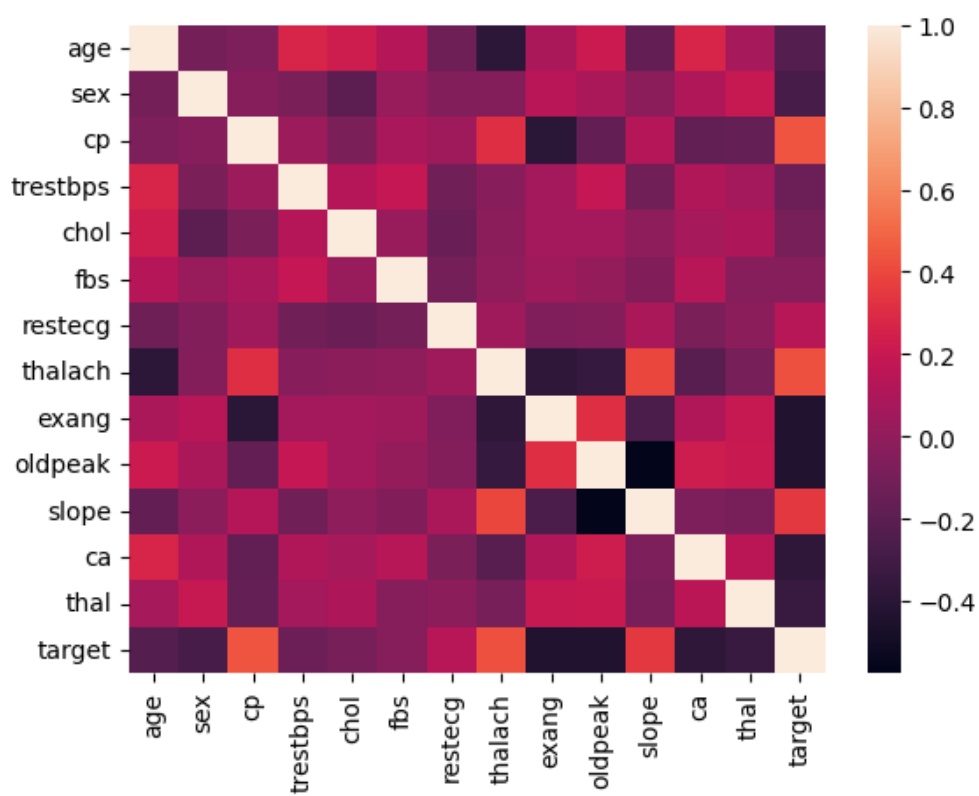
شکل ۷: نمودار خطا برای شبکه عصبی با فعال ساز سیگموئید و بهینه ساز آدام



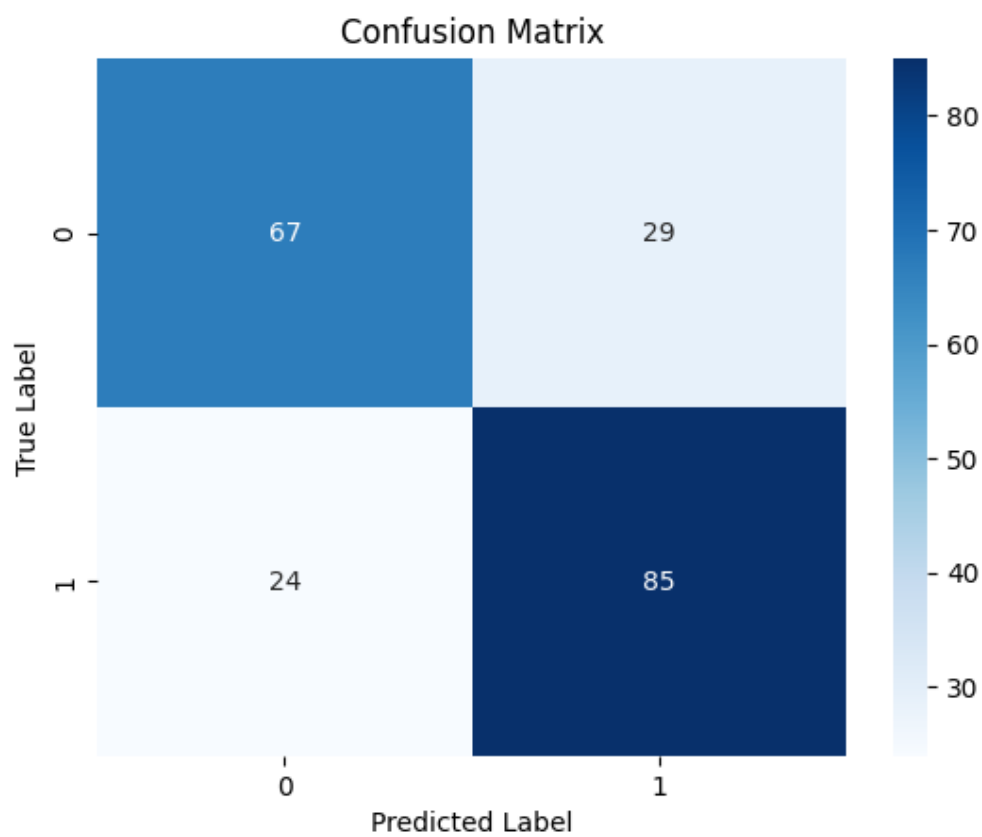
شکل ۸: نمودار خطا برای شبکه عصبی با فعال ساز رلو و بهینه ساز آدام



شکل ۹: هیستوگرام داده ها



شکل ۱۰: نمودار هیت مپ



شکل ۱۱: ماتریس درهم ریختگی