

به نام خدا

تمرین ۳: دسته بندی مجموعه داده iris با استفاده از شبکه عصبی **fully connected**

(به وسیله کتابخانه **keras**)

گرد آورنده: ساجده لشگری

در این تمرین داده‌های iris که مربوط به ۳ نوع گل virginica, versicolor و setosa می‌باشند، در نظر گرفته شده است. این داده‌ها شامل ۴ ویژگی (طول کاسبرگ، عرض کاسبرگ، طول گلبرگ و عرض گلبرگ) و ۱۵۰ نمونه هستند و هدف دسته بندی داده‌ها به ۳ کلاس اشاره شده، می‌باشد.

(۱- الف)

ابتدا با روش one hot encoding متغیرهای اسمی را به عددی تبدیل کرده، سپس با استفاده از روش MinMaxScaler نرمالسازی داده‌ها انجام شده و ۲۰٪ از داده‌ها، به عنوان داده‌های آزمایش و مابقی به عنوان داده‌های آموزش در نظر گرفته می‌شوند.

شبکه عصبی fully connected برای این داده‌ها با ۳ لایه پنهان که هر کدام شامل ۳ نرون می‌باشند و در مجموع ۵۱ پارامتر و تابع فعال‌سازی لایه آخر softmax در نظر گرفته شده است.

چند تابع فعال‌سازی برای لایه‌های پنهان، ۳ تابع هدف برای محاسبه loss و ۲ روش بهینه سازی adam و nesterov برای بدست آوردن پارامترها، روی این شبکه اعمال شده است.

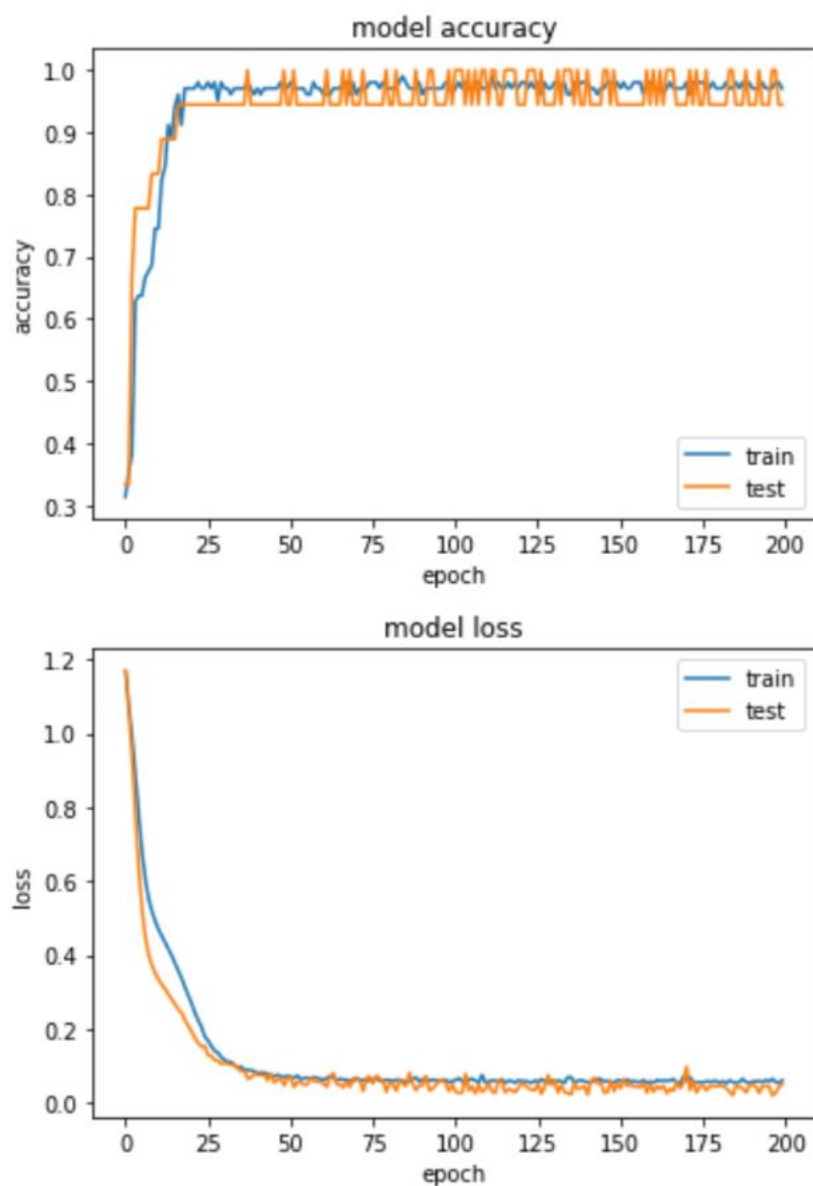
برای epochs=200 و validation_split=0.15 نتایج روی داده‌های آزمایش به صورت زیر می‌باشد.

Elu	Relu	Tanh	Tanh-relu-tanh	selu	
acc : 96.67% loss : 3.95%	acc : 26.67% loss : 114.44%	acc : 96.67% loss : 4.15%	acc : 96.67% loss : 4.09%	acc : 96.67% loss : 3.50%	Adam / cross entropy
acc : 96.67% loss : 1.01%	acc : 26.67% loss : 23.19%	acc : 96.67% loss : 1.07%	acc : 96.67% loss : 0.97%	acc : 96.67% loss : 1.10%	Adam / MSE
acc : 93.33% loss : 4.17%	acc : 26.67% loss : 48.58%	acc : 93.33% loss : 4.33%	acc : 93.33% loss : 3.32%	acc : 96.67% loss : 2.28%	Adam / MAE
acc : 96.67% loss : 1.64%	acc : 26.67% loss : 23.08%	acc : 96.67% loss : 1.52%	acc : 96.67% loss : 1.42%	acc : 96.67% loss : 1.66%	Nesterov / MSE

جدول ۱: نمایش دقت و مقدار loss برای توابع مختلف

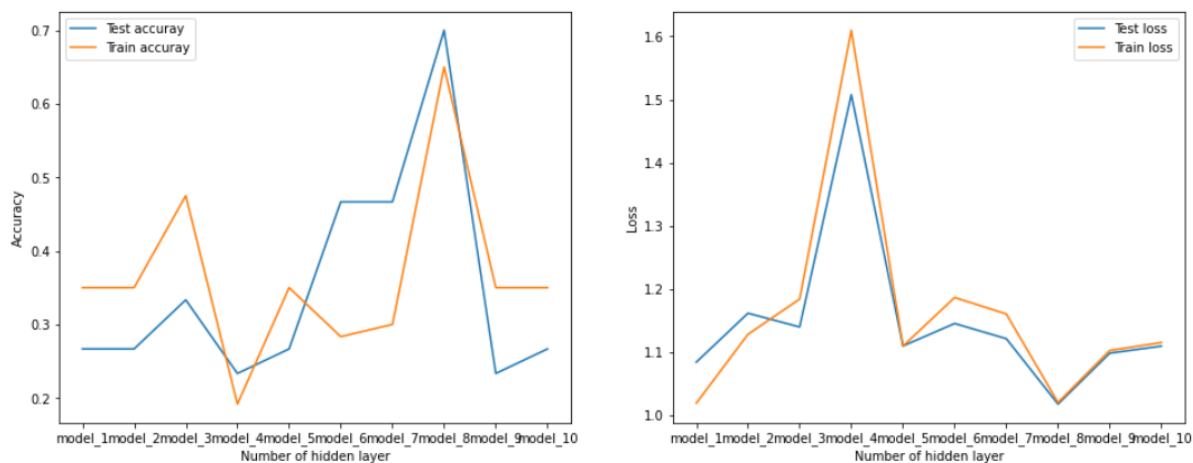
همانطور که در جدول ۱ مشاهده می‌شود بهترین تابع فعال‌سازی برای لایه‌های پنهان در بین این توابع، elu می‌باشد که دارای بیشترین دقت و کمترین مقدار loss در داده‌های آزمایش و آموزش است. همچنین در بین ۳ تابع هدف در نظر گرفته شده، بهترین نتیجه مربوط به کمترین مربعات خطا (MSE) می‌باشد.

در شکل ۱ نیز مشاهده می‌شود که با افزایش مقدار epoch، دقت افزایش و loss کاهش می‌یابد که این دلیلی بر خوب بودن مدل می‌باشد.



شکل ۱: نمودار مربوط به تابع فعال‌سازی elu با categorical cross entropy

۱- ب) در این بخش تعداد لایه‌های پنهان را ۱ تا ۱۰، batch size=8، epoch=100، تابع فعال‌سازی لایه‌های پنهان elu ، تابع هدف categorical cross entropy و روش بهینه‌سازی adam، در نظر گرفته شده و نتایج در شکل‌های زیر مشاهده می‌شود.



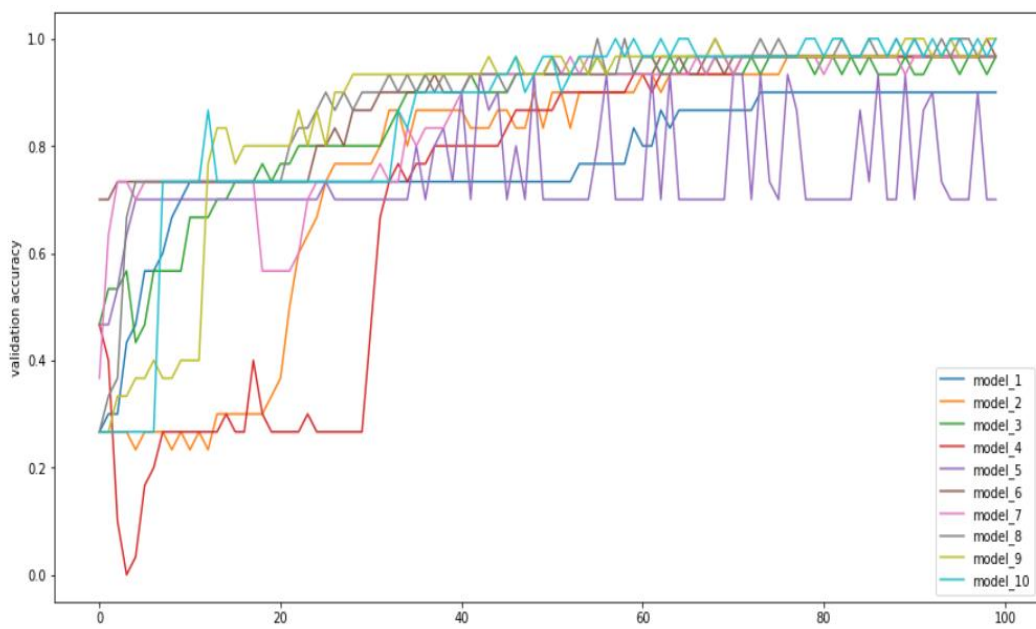
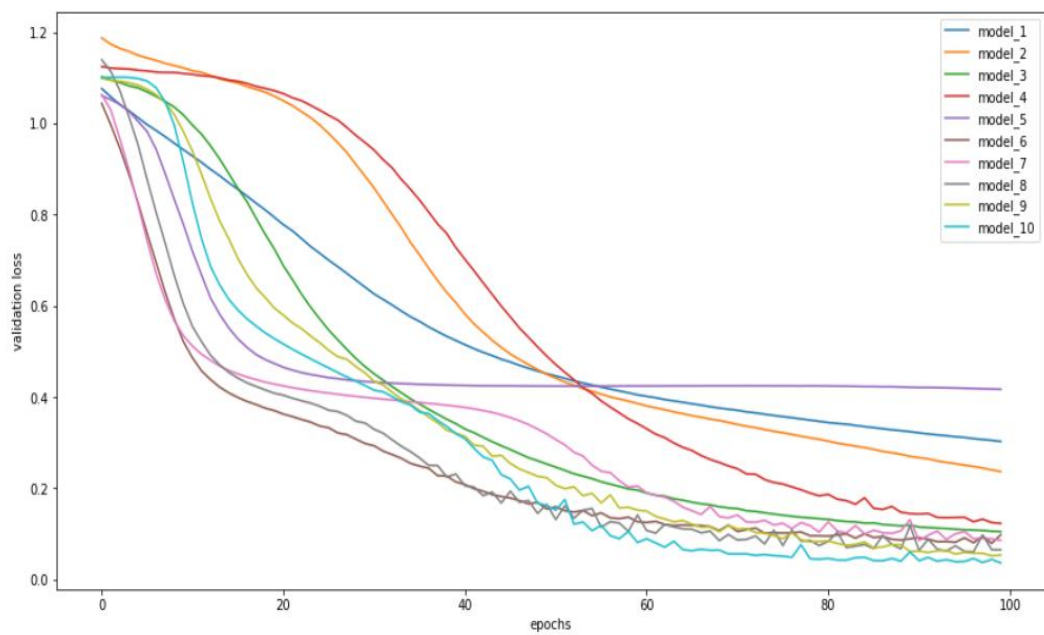
شکل ۲

همانطور که در شکل ۲ مشاهده می‌شود با افزایش تعداد لایه‌ها لزوماً نتیجه بهتری بدست نمی‌آید، میزان دقت تا لایه‌ی سوم افزایش، سپس در لایه چهارم کاهش، دوباره تا لایه هشتم افزایش و بعد از آن کاهش پیدا می‌کند.

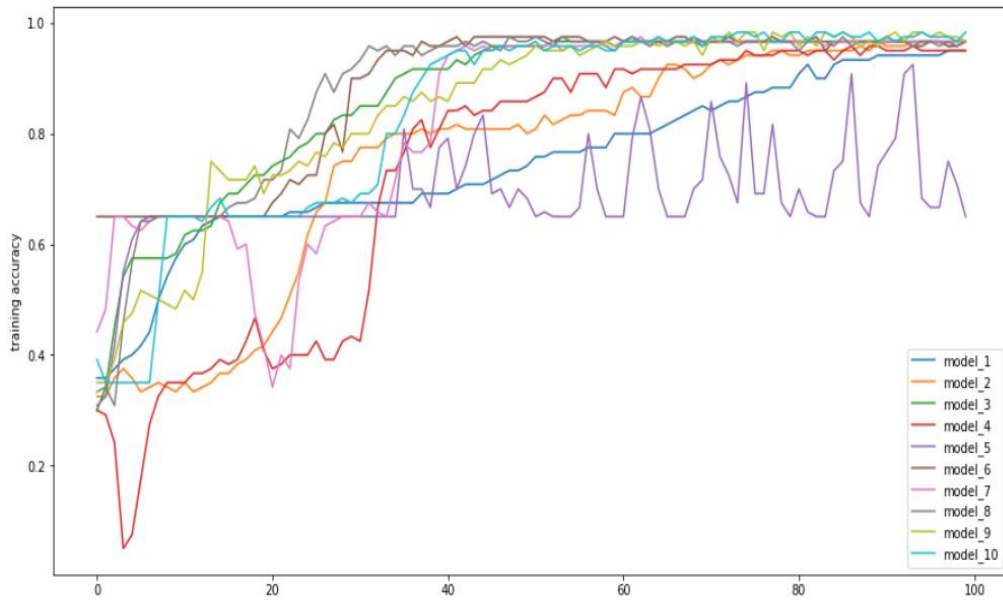
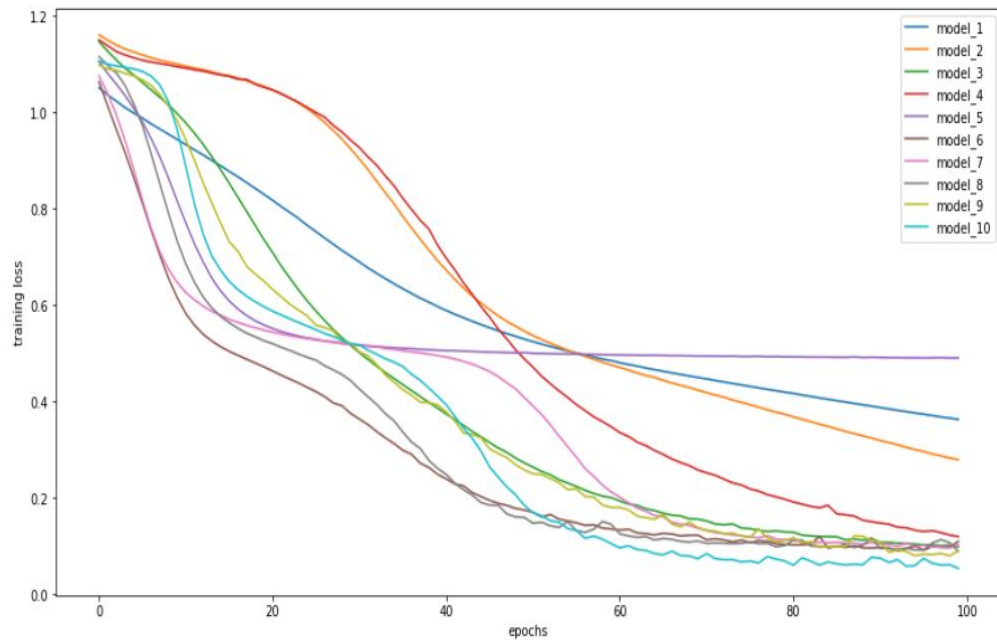
بیشترین میزان دقت مربوط به لایه هشتم و بعد از آن لایه سوم و کمترین مقدار آن مربوط به لایه چهارم و دهم می‌باشد.

همچنین مقدار loss در لایه سوم کمی کاهش پیدا کرده سپس بیشترین مقدار خود را در لایه چهارم بدست آورده است و کمترین مقدار آن مربوط به لایه هشتم می‌باشد.

با توجه به شکل ۲، ۳ و ۴ می‌توان گفت در این شرایط، بهترین تعداد برای لایه‌های پنهان ۸ و بعد از آن ۳ (که هزینه کمتری از لحاظ زمان نیز دارد) می‌باشد.



شکل ۳



شکل ۴

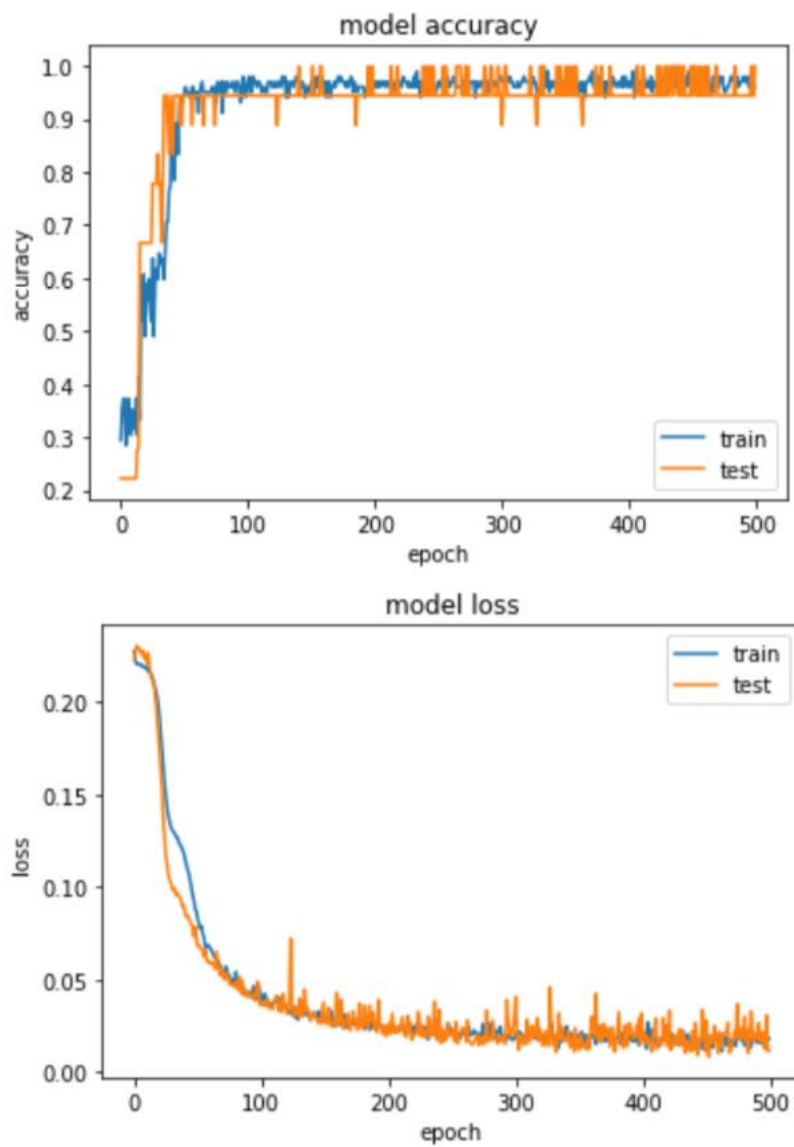
۲- برای هر ۳ لایه پنهان تابع فعال‌سازی softmax در نظر گرفته شده و از روش‌های بهینه‌سازی GD، SGD، Momentum (با پارامتر 0.9) و nesterov با دو تابع هدف MSE و categorical cross entropy استفاده شده و نتایج جدول ۲ بدست آمده است.

در اینجا epochs=500 و validation_split=0.15 در نظر گرفته شده است.

	GD	SGD	Momentum	Nesterov		
Test	acc: 46.67% loss: 21.81%	acc: 93.33% loss: 7.11%	acc: 93.33% loss: 1.89%	acc : 96.67% loss : 1.77%	MSE	
	acc : 26.67% loss : 112.66%	acc: 96.67% loss: 14.41%	acc: 93.33% loss: 30.89%	acc: 93.33% loss: 38.60%	Categorical-Crossentropy	
Train	acc: 30.00% loss: 22.22%	acc : 95.00% loss : 7.81%	acc: 97.50% loss: 1.26%	acc : 98.33% loss : 1.30%	MSE	
	acc : 35.00% loss : 108.87%	acc: 96.67% loss: 13.34%	acc: 97.50% loss: 6.47%	acc: 96.67% loss: 11.07%	Categorical-Crossentropy	

جدول ۲: نمایش دقت و مقدار loss برای توابع مختلف

همانطور که در جدول ۲ مشاهده می‌شود، تقریباً می‌توان گفت روش Nesterov نسبت به روش‌های دیگر عملکرد بهتری دارد. همچنین با استفاده از تابع هدف حداقل مربعات، مقدار loss کمتر از حالت cross entropy می‌باشد و معمولاً مقدار دقت در برخی مواقع بهبود و در برخی مواقع ثابت می‌ماند.



شکل ۵: نمودار مربوط به روش nesterov با تابع هدف MSE