

Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.01,	Batch Size: 32,	Test Loss: 0.6494
Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.01,	Batch Size: 32,	Test Loss: 0.6534
Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.01,	Batch Size: 64,	Test Loss: 0.6659
Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.001,	Batch Size: 32,	Test Loss: 0.6811
Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.001,	Batch Size: 32,	Test Loss: 0.6801
Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.001,	Batch Size: 64,	Test Loss: 0.6886
Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.0001,	Batch Size: 16,	Test Loss: 0.7750
Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.0001,	Batch Size: 32,	Test Loss: 0.7750
Hidden Layers: [4],	Activation: tanh,	Epochs: 10,	LR: 0.0001,	Batch Size: 64,	Test Loss: 0.8157
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.01,	Batch Size: 16,	Test Loss: 0.7471
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.01,	Batch Size: 32,	Test Loss: 0.6510
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.01,	Batch Size: 64,	Test Loss: 0.6524
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.001,	Batch Size: 16,	Test Loss: 0.6450
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.001,	Batch Size: 32,	Test Loss: 0.6773
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.001,	Batch Size: 64,	Test Loss: 0.6818
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.0001,	Batch Size: 16,	Test Loss: 0.7839
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.0001,	Batch Size: 32,	Test Loss: 0.7839
Hidden Layers: [4],	Activation: tanh,	Epochs: 25,	LR: 0.0001,	Batch Size: 64,	Test Loss: 0.9100
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.01,	Batch Size: 16,	Test Loss: 0.6529
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.01,	Batch Size: 32,	Test Loss: 0.6540
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.01,	Batch Size: 64,	Test Loss: 0.6358
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.001,	Batch Size: 16,	Test Loss: 0.6546
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.001,	Batch Size: 32,	Test Loss: 0.6546
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.001,	Batch Size: 64,	Test Loss: 0.6577
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.0001,	Batch Size: 16,	Test Loss: 0.6768
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.0001,	Batch Size: 32,	Test Loss: 0.6893
Hidden Layers: [4],	Activation: tanh,	Epochs: 50,	LR: 0.0001,	Batch Size: 64,	Test Loss: 0.7176

Dalam analisis hasil eksperimen dengan konfigurasi hyperparameter pada model MLP, terlihat bahwa variasi jumlah epoch, learning rate (LR), dan ukuran batch (batch size) memberikan pengaruh signifikan terhadap nilai test loss. Untuk fungsi aktivasi **linear**, dengan LR 0.01 dan batch size 32, nilai test loss mencapai 0.6969 setelah 10 epoch. Pengurangan learning rate (misalnya menjadi 0.001 atau 0.0001) cenderung meningkatkan test loss, yang menunjukkan bahwa penurunan learning rate dapat memperlambat proses konvergensi model tanpa memberikan peningkatan yang signifikan dalam akurasi.

Untuk fungsi aktivasi lainnya, seperti **sigmoid**, **ReLU**, **softmax**, dan **tanh**, hasil yang lebih baik terlihat pada konfigurasi dengan LR 0.01, terutama pada batch size yang lebih kecil (16 atau 32). Fungsi aktivasi **sigmoid**, misalnya, menghasilkan test loss 0.6501 dengan batch size 16 pada 10 epoch. Fungsi aktivasi ReLU, di sisi lain, cenderung mengalami peningkatan test loss yang lebih besar pada batch size yang lebih besar, yang menunjukkan bahwa fungsi aktivasi ini lebih sensitif terhadap perubahan batch size pada konfigurasi tertentu. Fungsi aktivasi **tanh** dan **softmax** juga menunjukkan bahwa model dengan epoch yang lebih banyak (misalnya 50 epoch) berpotensi untuk memperburuk hasil apabila learning rate terlalu kecil.

Secara keseluruhan, untuk konfigurasi ini, hasil terbaik tercapai dengan menggunakan fungsi aktivasi **sigmoid** pada LR 0.01 dan batch size 16, dengan test loss terendah di angka 0.6501. Dapat disimpulkan bahwa fungsi aktivasi dan learning rate memainkan peran penting dalam menentukan performa model, dengan fungsi aktivasi non-linear seperti **sigmoid** dan **ReLU** lebih efektif pada nilai LR yang lebih tinggi dan ukuran batch yang lebih kecil.

Dari hasil eksperimen dengan konfigurasi **8 hidden layers**, terdapat variasi performa berdasarkan kombinasi antara epoch, learning rate (LR), batch size, dan fungsi aktivasi yang digunakan. Pada fungsi aktivasi **linear**, hasil loss yang dicapai cenderung lebih tinggi dibandingkan dengan aktivasi lain, dengan nilai test loss terbesar mencapai 0.8609. Ini menunjukkan bahwa model dengan aktivasi linear mungkin tidak optimal untuk tugas ini, mengingat nilai loss yang lebih besar daripada yang menggunakan fungsi aktivasi non-linear seperti **sigmoid**, **ReLU**, **softmax**, atau **tanh**.

Fungsi aktivasi **sigmoid** menunjukkan hasil yang cukup baik, terutama pada epoch 25 dengan learning rate 0.01 dan batch size 64, dengan test loss mencapai 0.6299. Namun, untuk kombinasi lainnya, terutama pada epoch yang lebih sedikit atau learning rate yang sangat rendah, performanya cenderung memburuk, seperti pada nilai test loss yang sangat tinggi di bawah learning rate 0.0001. **ReLU** juga memberikan hasil yang kompetitif, dengan nilai test loss terendah mencapai 0.6389 pada epoch 25, learning rate 0.001, dan batch size 32, meskipun beberapa kombinasi menghasilkan nilai loss yang sedikit lebih buruk.

Fungsi aktivasi **softmax** dan **tanh** juga menunjukkan hasil yang mengesankan pada beberapa kombinasi. **Softmax** berhasil mencatatkan test loss terbaik di 0.6093 pada epoch 50, LR 0.01, dan batch size 64, menunjukkan bahwa untuk tugas klasifikasi, softmax dapat memberikan performa yang lebih stabil dan baik. Di sisi lain, **tanh** memperlihatkan test loss terendah 0.6272 pada epoch 25, LR 0.01, dan batch size 32. Secara keseluruhan, hasil menunjukkan bahwa pemilihan fungsi aktivasi dan pengaturan hyperparameter sangat mempengaruhi hasil model, dengan softmax dan sigmoid memberi hasil terbaik pada dataset ini.

Dari hasil eksperimen menggunakan konfigurasi dengan jumlah hidden layer sebanyak 16, terlihat bahwa pemilihan fungsi aktivasi dan parameter lainnya memiliki pengaruh signifikan terhadap performa model. Secara umum, fungsi aktivasi **ReLU** menunjukkan hasil yang lebih baik pada beberapa kombinasi parameter, terutama pada **epochs 50** dan **learning rate 0.01**, yang menghasilkan **test loss** terendah sekitar 0.6087 dengan **batch size 16**. Ini menunjukkan bahwa **ReLU** cenderung memberikan hasil yang lebih optimal untuk model ini.

Sebaliknya, fungsi aktivasi **sigmoid** dan **softmax** cenderung menghasilkan **test loss** yang lebih tinggi, terutama dengan learning rate rendah seperti **0.0001**. Misalnya, pada sigmoid dengan epoch 50 dan batch size 64, **test loss** mencapai 0.7561, yang lebih buruk dibandingkan dengan **ReLU** di konfigurasi serupa. Fungsi aktivasi ini mungkin mengalami kesulitan dalam konvergensi pada learning rate yang sangat kecil, yang dapat memengaruhi kemampuan model untuk menemukan solusi yang lebih baik.

Secara keseluruhan, hasil yang diperoleh mengindikasikan bahwa kombinasi **fungsi aktivasi ReLU**, **learning rate 0.01**, dan **epoch 50** lebih optimal dalam konteks ini, memberikan performa terbaik dalam hal **test loss**. Untuk hasil yang lebih stabil, disarankan untuk menghindari penggunaan learning rate yang terlalu kecil atau epoch yang sangat pendek, karena hal tersebut cenderung menurunkan performa model.

D. Hidden Layer [4,8]

[illegible][illegible]

Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.01,	Batch Size:	16,	Test Loss:	0.6493
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.01,	Batch Size:	32,	Test Loss:	0.6700
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.01,	Batch Size:	64,	Test Loss:	0.6661
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.001,	Batch Size:	16,	Test Loss:	0.6803
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.001,	Batch Size:	32,	Test Loss:	0.6835
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.001,	Batch Size:	64,	Test Loss:	0.6888
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.0001,	Batch Size:	16,	Test Loss:	0.7075
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.0001,	Batch Size:	32,	Test Loss:	0.7697
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	10,	LR:	0.0001,	Batch Size:	64,	Test Loss:	0.8993
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.01,	Batch Size:	16,	Test Loss:	0.6274
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.01,	Batch Size:	32,	Test Loss:	0.6345
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.01,	Batch Size:	64,	Test Loss:	0.6478
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.001,	Batch Size:	16,	Test Loss:	0.6318
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.001,	Batch Size:	32,	Test Loss:	0.6604
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.001,	Batch Size:	64,	Test Loss:	0.6669
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.0001,	Batch Size:	16,	Test Loss:	0.6862
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.0001,	Batch Size:	32,	Test Loss:	0.7061
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	25,	LR:	0.0001,	Batch Size:	64,	Test Loss:	0.7633
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.01,	Batch Size:	16,	Test Loss:	0.6471
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.01,	Batch Size:	32,	Test Loss:	0.6279
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.01,	Batch Size:	64,	Test Loss:	0.6293
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.001,	Batch Size:	16,	Test Loss:	0.6607
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.001,	Batch Size:	32,	Test Loss:	0.6459
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.001,	Batch Size:	64,	Test Loss:	0.6397
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.0001,	Batch Size:	16,	Test Loss:	0.6858
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.0001,	Batch Size:	32,	Test Loss:	0.6961
Hidden Layers:	[4, 8],	Activation:	tanh,	Epochs:	50,	LR:	0.0001,	Batch Size:	64,	Test Loss:	0.7268

Untuk konfigurasi hidden layers dengan [4,8], hasil menunjukkan bahwa penggunaan aktivasi **linear**, **sigmoid**, **ReLU**, **softmax**, dan **tanh** memberikan variasi performa yang signifikan tergantung pada kombinasi epoch, learning rate (LR), dan batch size. Pada umumnya, konfigurasi dengan aktivasi **sigmoid** dan **softmax** menunjukkan performa yang lebih baik pada epoch yang lebih tinggi dan learning rate yang lebih besar. Contohnya, pada epoch 50 dengan LR 0.01, batch size 64, **softmax** menghasilkan loss terendah, yakni 0.6293, yang lebih baik dibandingkan kombinasi lainnya. Sebaliknya, penggunaan aktivasi **linear** dan **sigmoid** dengan learning rate sangat kecil (0.0001) cenderung meningkatkan loss secara signifikan, menunjukkan ketidakefektifan kombinasi tersebut.

Dalam hal batch size, ukuran yang lebih kecil seperti 16 sering memberikan hasil yang lebih stabil, tetapi dengan waktu pelatihan yang lebih lama. Batch size yang lebih besar, seperti 64, memperlihatkan peningkatan loss pada konfigurasi tertentu, terutama dengan learning rate kecil. Pada epoch yang lebih tinggi (misalnya 50), loss yang tercatat cenderung lebih rendah, mencerminkan konvergensi yang lebih baik pada model, meskipun ada variasi pada konfigurasi aktivasi dan learning rate. Aktivasi **ReLU** dan **tanh**, meskipun kadang memberikan loss yang lebih tinggi, tetap menunjukkan potensi dalam konfigurasi yang lebih optimal di epoch dan LR yang lebih besar.

Secara keseluruhan, konfigurasi terbaik tampaknya melibatkan kombinasi aktivasi **softmax** atau **sigmoid** dengan epoch yang lebih tinggi dan learning rate moderat (0.01). Kombinasi ini memberikan performa terbaik pada beberapa uji, dengan loss yang lebih rendah di sebagian besar batch size. Ini menunjukkan pentingnya pemilihan parameter yang tepat dalam mencapai hasil yang optimal, terutama dalam klasifikasi dengan menggunakan MLP.

Pada konfigurasi **hidden layer** [8, 16], eksperimen dengan berbagai pengaturan hyperparameter menunjukkan hasil yang variatif, terutama dalam hal fungsi aktivasi. Ketika menggunakan **linear**, test loss menunjukkan hasil yang relatif lebih tinggi, berkisar antara 0.6860 hingga 0.8033 untuk berbagai kombinasi learning rate dan batch size. Hal ini mengindikasikan bahwa konfigurasi linear mungkin kurang optimal untuk dataset ini, dengan nilai test loss yang tidak cukup rendah, bahkan meskipun ada penambahan epoch.

Namun, saat menggunakan **sigmoid**, hasil test loss mengalami penurunan yang signifikan, terutama pada konfigurasi dengan learning rate yang lebih tinggi (0.01) dan batch size kecil (16 atau 32). Nilai test loss mencapai 0.6218 hingga 0.6573 pada epoch 50, menunjukkan kinerja yang lebih baik dibandingkan dengan linear. Penggunaan **ReLU** juga memberikan hasil yang baik, dengan nilai test loss antara 0.6170 hingga 0.6607 pada epoch yang lebih panjang, yang menunjukkan bahwa **ReLU** lebih efektif dalam menangani masalah ini.

Pada konfigurasi **softmax**, meskipun memberikan hasil yang relatif baik pada batch size lebih kecil dan learning rate yang lebih besar, nilai test loss tidak dapat bersaing dengan hasil dari aktivasi sigmoid atau **ReLU**, terutama pada epoch yang lebih panjang. Test loss berkisar antara 0.6044 hingga 0.9131 pada epoch 50, yang menunjukkan bahwa softmax mungkin tidak ideal dalam konfigurasi ini. Demikian pula, dengan **tanh**, hasilnya tidak jauh lebih baik dari softmax, dengan test loss berada di kisaran 0.6223 hingga 0.8347. Secara keseluruhan, **sigmoid dan ReLU** tampaknya memberikan performa terbaik pada tugas ini, dengan sigmoid cenderung menunjukkan hasil yang lebih stabil pada beberapa konfigurasi hyperparameter.

Untuk konfigurasi Hidden Layer [16, 32] dengan fungsi aktivasi yang berbeda, hasil evaluasi menunjukkan variasi yang cukup besar dalam nilai test loss, tergantung pada kombinasi hyperparameter yang digunakan. Dengan fungsi aktivasi **linear**, test loss menunjukkan nilai yang lebih tinggi, terutama pada pengaturan epoch yang lebih banyak dan learning rate yang lebih kecil, seperti pada epoch 50 dan learning rate 0.0001 yang menghasilkan test loss 0.6915. Namun, ketika menggunakan fungsi aktivasi **sigmoid**, **ReLU**, dan **softmax**, hasilnya menunjukkan kinerja yang lebih baik, terutama pada epoch yang lebih tinggi dengan learning rate yang lebih besar.

Konfigurasi dengan fungsi aktivasi **sigmoid** menunjukkan test loss yang lebih rendah pada epoch 10 dan 25, dengan learning rate 0.01, seperti pada kombinasi batch size 16 yang menghasilkan test loss 0.6242. Fungsi **ReLU** juga memberikan hasil yang cukup baik, dengan test loss yang lebih rendah pada kombinasi epoch 25 dan batch size 64, yang mencapai 0.5902. Meskipun fungsi aktivasi **softmax** dan **tanh** menunjukkan hasil yang lebih bervariasi, **softmax** cenderung menghasilkan nilai test loss yang lebih tinggi pada epoch dan learning rate yang lebih rendah, sedangkan **tanh** memberikan nilai test loss yang relatif stabil meskipun sedikit lebih tinggi pada beberapa pengaturan.

Secara keseluruhan, fungsi aktivasi **ReLU** dan **sigmoid** menunjukkan kinerja terbaik di antara semua pengaturan, dengan test loss yang lebih rendah pada berbagai kombinasi hyperparameter, terutama ketika batch size lebih besar dan epoch yang lebih tinggi digunakan. Meskipun ada beberapa fluktuasi pada beberapa kombinasi, seperti pada learning rate rendah, hasil ini menunjukkan pentingnya pemilihan fungsi aktivasi yang tepat dalam membangun model MLP untuk optimasi hasil yang lebih baik.

Dari hasil eksperimen dengan jumlah hidden layers [4, 8, 16] pada berbagai kombinasi fungsi aktivasi, epoch, learning rate, dan batch size, dapat disimpulkan bahwa konfigurasi dengan fungsi aktivasi **softmax** memberikan performa yang paling stabil. Pada epoch 50, learning rate 0.01, dan batch size 16, model menghasilkan test loss yang paling rendah, yaitu **0.6259**, menunjukkan bahwa pengaturan ini cenderung lebih efektif dalam mengoptimalkan model.

Sebaliknya, fungsi aktivasi **sigmoid** dan **tanh** menunjukkan performa yang lebih buruk dibandingkan dengan softmax, terutama pada learning rate yang lebih kecil (0.0001). Hal ini terlihat jelas dengan test loss yang lebih tinggi pada kombinasi seperti **0.9391** (sigmoid, epoch 10, lr 0.0001) dan **0.6376** (tanh, epoch 10, lr 0.01). Fungsi aktivasi sigmoid, dalam hal ini, tampaknya tidak cukup optimal untuk dataset ini, sedangkan tanh cenderung memberikan sedikit perbaikan dibandingkan sigmoid.

Secara umum, **ReLU** juga menunjukkan performa yang baik, terutama pada epoch lebih panjang (25-50) dengan test loss sekitar **0.6321** hingga **0.6563** untuk berbagai kombinasi. Namun, kombinasi dengan learning rate yang lebih rendah, seperti **0.0001**, menghasilkan test loss yang lebih buruk, mencapai sekitar **0.8595** pada epoch 10. Fungsi aktivasi **linear** menunjukkan hasil yang lebih variatif, dengan beberapa konfigurasi memberikan test loss yang cukup tinggi, seperti pada epoch 25 dengan learning rate 0.0001 yang menghasilkan **0.8171**.