

COMPENG 4SL4

Assignment 4

Neural Network Classifier With Two Hidden Layers

Submitted by:

Snigdha Majeti (Std No.: 400263044)

November 24, 2023

Neural Network Classifier:

For the splitting ratio, 80% of the dataset is used for training, 10% of the dataset is used for testing, and 10% of the dataset is used for validation. 3044, which are the last four digits of my student number, is the random seed used to split the dataset. The hidden layers use ReLU activation function, and the output layer uses logistic sigmoid activation function to compute the output. Stochastic Gradient Descent or SGD is used to train the neural network with a learning rate of 0.001. The early stopping condition for SGD is done by fixing the number of epochs to 250 and computing the training cross-entropy loss and validation cross-entropy loss after each epoch. After computing the losses and running the SGD algorithm for all 250 epochs the weights of the epoch that achieve the lowest validation cross-entropy loss are chosen.

The weights are initialized using uniform distribution for each layer from $(-\sqrt{\frac{6}{m+n}}, \sqrt{\frac{6}{m+n}})$, where m is the number of inputs of a fully connected layer and n is the number of outputs of a fully connected layer. The biases for each layer are always initialized to zero. For each, (n_1, n_2) configuration, where n_1 is the number of units for hidden layer 1 and n_2 is the number of units for hidden layer 2, SGD is used to train the neural network for a fixed number of epochs of 250. As discussed above, the training cross-entropy loss and the validation cross-entropy loss is computed after each epoch and the weights that achieve the lowest validation cross-entropy loss are chosen for each (n_1, n_2) configuration. Since, the weight initialization affects the model performance and the learning curves, each (n_1, n_2) are run three times with three different initialized weights, while keeping the bias initialization to zero for all the three runs .

The tables starting from the next page shows the three different initialized biases and weights, training loss, validation loss, and trained biases and weights for each (n_1, n_2) configurations. There are ten tables for ten different (n_1, n_2) configurations, which are listed below and chosen for comparison. In the tables below, the initialized biases and weights and trained biases and weights are formatted into lists with each list containing three numpy arrays. The first numpy array are biases and the weights corresponding to the first hidden layer, the second numpy array are biases and the weights corresponding to the second hidden layer, and the third numpy array are biases and the weights corresponding to the output layer or final layer. The first column in each numpy array corresponds to the bias of a given layer. The training loss and validation loss are the cross-entropy losses at the epoch where the validation loss is at its minimum or where the trained biases and weights are chosen from. Also, for all the model configurations the activation function for hidden layers is ReLU, and the activation function for the output layer is logistic sigmoid.

Table for (1, 1) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , 0.56107789, 0.98786005, -0.13352947, -1.07082455]]), array([[0. , -0.74321411]]), array([[0. , 0.6186633]])]	0.6887350246	0.6726559552	[array([[0. , 0.56107789, 0.98786005, -0.13352947, -1.07082455]]), array([[0. , -0.74321411]]), array([[-0.19259008, 14.14558375]])]
[array([[0. , -0.76914913, 0.85393992, 0.36354999, -0.57541701]]), array([[0. , -0.68144276]]), array([[0. , -1.60000756]])]	0.6887356376	0.6726109065	[array([[0. , -0.76914913, 0.85393992, 0.36354999, -0.57541701]]), array([[0. , -0.68144276]]), array([[-0.193137 , -4.20145506]])]
[array([[0. , 0.71077872, -0.17052105, -0.60827485, 0.61958766]]), array([[0. , 0.19252532]]), array([[0. , -1.06008347]])]	0.05361375352	0.05517885589	[array([[0.95877881, 1.14113426, 1.31420953, 1.16525473, -0.01671473]]), array([[-0.00393487, 2.45791887]]), array([[2.70461496, -2.26293673]])]

Table for (1, 2) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , 0.77727722, 0.02843436, -0.14888747, -0.50806095]]), array([[0. , -0.67003216], [0. , -0.8021183]]), array([[0. , -0.59163905, 0.57564262]])]	0.6887359915	0.6725871226	[array([[0. , 0.77727722, 0.02843436, -0.14888747, -0.50806095]]), array([[0. , -0.67003216], [0. , -0.8021183]]), array([[-0.19342612, 5.16314616, 7.46489182]])]
[array([[0. , 0.69403056, -1.03402901, 0.64216093, 0.71524929]]), array([[0. , -1.28308573], [0. , -0.3938469]]), array([[0. , -0.23584695, -0.23254942]])]	0.6887338975	0.6727596747	[array([[0. , 0.69403056, -1.03402901, 0.64216093, 0.71524929]]), array([[0. , -1.28308573], [0. , -0.3938469]]), array([[-0.19133426, -4.17770604, -1.44251453]])]
[array([[0. , 0.58363333, -0.78160044, 0.95661603, 0.87193963]]), array([[0. , 0.01049547], [0. , -0.02578116]]), array([[0. , -0.68865214, -0.82249401]])]	0.024522834	0.02722496771	[array([[0.02305449, -1.02220805, -1.13796322, -1.04852306, -0.06458115]]), array([[1.84844737, -2.53461438], [0. , -0.02578116]]), array([[4.2986811 , -7.7228897 , -1.10919547]])]

Table for (2, 2) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , -0.95464345, 0.57280377, 0.67205988, -0.82005157], [0. , 0.39726901, -0.2568889 , 0.74647635, -0.57593757]]), array([[0. , 0.63542515, 0.80697399], [0. , -0.79420666, -0.66688375]]), array([[0. , 0.12317749, -0.2626865]])]	0.02975738831	0.04071983342	[array([[-0.00583463, -1.55439028, -1.77919206, -1.58632083, -0.03672725], [-0.30702431, 0.03751734, -0.42946655, 0.64097604, -0.42293335]]), array([[-0.2480461 , 2.72663456, 0.02245207], [0. , -0.79420666, -0.66688375]]), array([[-4.11164489, 2.57021279, -1.36225206]])]
[array([[0. , -0.0016433 , -0.82662265, -0.23167737, 0.61763535], [0. , 0.64419962, -0.25109687, -0.80277133, 0.55504281]]), array([[0. , 0.71045492, 0.25742774], [0. , -0.73310474, 1.03711667]]), array([[0. , 1.3973697 , -0.27532468]])]	0.01838029315	0.02434681787	[array([[0.09523581, -1.14615224, -1.5464211 , -1.3203776 , 0.04306066], [0.82439955, 1.18533867, -0.29045227, 0.26081373, 0.58708466]]), array([[-0.00276393, 2.32451376, -0.28661274], [1.67719404, -2.07108248, 0.86946989]]), array([[-0.88550276, 2.42965382, -2.32271052]])]
[array([[0. , -0.41487373, -0.50229174, 0.91731761, 0.92932473], [0. , 0.2369193 , -0.91511224, -0.22948963, 0.85046618]]), array([[0. , -0.7243318 , 1.13705229], [0. , 0.44823566, -0.17105302]]), array([[0. , -1.28979343, 0.44923351]])]	0.019053289	0.02391118834	[array([[-0.12461123, -1.0678171 , -1.26659317, -1.10747743, -0.00465073], [0.65534313, 1.05095455, -0.6309167 , 0.24647493, 0.9275761]]), array([[1.59473423, -3.94110252, 0.64810561], [0.09984634, 2.62051374, 0.15902375]]), array([[0.01664639, -3.41414774, 1.76224449]])]

Table for (3, 1) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , 0.10872305, 0.92100501, 0.03256916, -0.11139706], [0. , -0.50963798, -0.31678195, -0.52237596, 0.30083506], [0. , 0.11473761, 0.31134311, -0.69114721, 0.24068492]]), array([[0. , 0.22505639, 1.01509473, 0.82198459]]), array([[0. , 0.66932632]])]	0.01995481327	0.02432608676	[array([[-0.00281626, -0.02047345, 1.14572638, 0.25412286, 0.13556178], [-0.0499745 , -1.36484656, -1.30002979, -1.15688532, 0.06141271], [-0.00842866, 0.0042672 , -0.14216644, -1.23269792, 0.02376182]]), array([[-0.38932944, -0.74565414, 2.56811871, 0.75729204]]), array([[-3.95112116, 3.32578413]])]
[array([[0. , 0.41827948, 0.54420439, 0.34945364, -0.35887509], [0. , 0.36074784, 0.54290892, -0.11680501, 0.30940794], [0. , -0.08324701, 0.77797727, 0.81268513, -0.43684566]]), array([[0. , -0.31455113, -0.35975852, -0.22936953]]), array([[0. , 0.77042285]])]	0.6887359986	0.6725866627	[array([[0. , 0.41827948, 0.54420439, 0.34945364, -0.35887509], [0. , 0.36074784, 0.54290892, -0.11680501, 0.30940794], [0. , -0.08324701, 0.77797727, 0.81268513, -0.43684566]]), array([[0. , -0.31455113, -0.35975852, -0.22936953]]), array([[-0.19343171, 13.12673218]])]
[array([[0. , -0.43888063, -0.48279404, 0.08767892, 0.86725138], [0. , 0.69178289, -0.32917548, -0.90856972, 0.0402525], [0. , 0.77813435, 0.51871911, 0.7626803 , 0.87371217]]), array([[0. , 1.11359186, 0.12565057, 0.39121601]]), array([[0. , 0.36831026]])]	0.03004712517	0.03560482603	[array([[0.05066319, -1.06418303, -1.24881225, -1.08355509, 0.09340732], [0.04349991, 0.72718602, -0.33075242, -1.03303393, 0.01106024], [0.4382791 , 1.03882204, 0.43360024, 0.64709364, 0.5239984]]), array([[-0.311747 , 2.46273437, 0.20434509, -1.0568911]]), array([[-3.46027373, 4.24056813]])]

Table for (2, 4) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , 0.82604039, -0.39769247, 0.07252918, 0.72768515], [0. , -0.88681293, -0.45409457, -0.97592358, 0.55760686]]), array([[0. , 0.83007822, 0.87417225], [0. , 0.96164911, 0.23973004], [0. , -0.06701569, 0.37703045], [0. , 0.68311724, -0.78728026]]), array([[0. , -0.35086681, -0.93590099, 0.98022298, 0.84459776]])]	0.01910428148	0.02486495775	[array([[0.72444008, 1.38971452, 0.08506859, 0.64759285, 0.402049], [0.20205199, -1.16918551, -1.62118797, -1.33691232, 0.0476599]]), array([[1.60192020e-02, 1.03931856e+00, 5.80286267e-01], [1.53988983e+00, 9.87357765e-01, -1.46264183e+00], [-1.54198101e-03, -3.32552015e-01, 2.19200275e+00], [-2.43491814e-01, 4.34386586e-01, -6.23295507e-01]]), array([[-0.65871178, -0.19266509, -2.34864341, 2.15933256, 0.10556628]])]
[array([[0. , 0.2499661 , 0.18346337, -0.09650993, 0.5478509], [0. , -0.84106352, 0.50298515, 0.58230411, 0.01707839]]), array([[0. , 0.01402062, -0.26409208], [0. , 0.50409673, -0.58785931], [0. , 0.13976026, 0.2020186], [0. , -0.83049976, -0.61080877]]), array([[0. , 0.96649754, 0.20628184, -0.95617257, 0.72068087]])]	0.04538158308	0.04753772746	[array([[0.91617189, 1.00552888, 1.23936878, 1.11483554, -0.00802457], [-0.12717739, -0.93306866, 0.19640117, 0.58713753, 0.20243517]]), array([[-1.03019239e-02, 4.13329680e-03, -2.50518806e-01], [1.40040679e-04, 1.20270402e+00, -1.59546869e-01], [-2.15434517e-03, 1.90356316e+00, -1.57100107e-01], [0.00000000e+00, -8.30499760e-01, -6.10808775e-01]]), array([[2.81863172, 0.15409857, -1.79848592, -2.36364798, -0.31313392]])]
[array([[0. , -0.67364447, 0.44253728, -0.73185191, -0.11477104], [0. , 0.84267831, -0.48270016, -0.94593114, -0.78163725]]), array([[0. , 0.202254 , 0.10759452], [0. , -0.57141976, 0.78137587], [0. , 0.58371157, 0.46631267], [0. , 0.12024102, 0.9594998]]), array([[0. , 0.84306208, 0.75037775, -0.66937334, 1.02166152]])]	0.02232335559	0.02739681718	[array([[0.46534089, -0.8622277 , -1.23338856, -1.29444233, -0.03294617], [1.1251391 , 1.03061388, 0.86394426, 0.68617864, -0.0654154]]), array([[0.94999152, 0.50591259, -1.94309872], [-0.28067125, -0.34154868, 0.55306026], [0.17364587, -0.42188393, 2.64608462], [-0.27452272, 1.82481232, 0.01997794]]), array([[-1.32615063e-03, 1.71805427e+00, -1.02441038e-01, -1.50809129e+00, 1.73297707e+00]])]

Table for (3, 3) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , 0.10723722, -0.58246244, -0.53967001, -0.2635963], [0. , 0.38422868, -0.38724923, -0.82873229, 0.28959379], [0. , 0.72625126, 0.25185618, 0.40507633, 0.06509694]]), array([[0. , -0.86387413, 0.53907567, 0.75962855], [0. , 0.14971013, -0.39696723, 0.33793079], [0. , 0.12416128, 0.74862431, 0.22583319]]), array([[0. , -0.47178318, -1.1956673 , -0.13392968]])]	0.03856046929	0.0387050612	[array([[-0.04615853, 0.23547205, -0.51686129, -1.07423736, -0.0668423], [-0.06575861, 0.41371235, -0.34118509, -0.7036675 , 0.15126195], [1.01835315, 1.11622884, 1.14203055, 0.96762054, -0.0942861]]), array([[-1.66896098e-03, -9.42142157e-01, 6.79205370e-01, 1.62568745e+00], [-3.19312438e-03, -3.12578667e-01, -2.49092342e-01, 1.70058438e+00], [-2.12861078e-02, 7.94224802e-02, 7.21459892e-01, 3.21678137e-01]]), array([[2.86374252, -1.72922583, -2.11856053, 0.14992633]])]
[array([[0. , 0.14924344, -0.79231436, -0.01288788, -0.07143975], [0. , 0.77298542, -0.48415698, 0.63945537, 0.14528029], [0. , 0.59353801, 0.81942833, -0.42274837, 0.35826811]]), array([[0.00000000e+00, -6.37227356e-04, 3.05922105e-01, 1.02272960e-01], [0.00000000e+00, -4.44814414e-01, -7.13182237e-02, 9.75877481e-01], [0.00000000e+00, -6.02859411e-01, 1.40380231e-01, 8.09180186e-01]]), array([[0. , 0.84064071, 0.89658428, -0.55234656]])]	0.01891897533	0.01638028445	[array([[-0.05530826, -1.10099344, -0.98091229, -0.13224354, 0.05039631], [0.66549619, 1.2284614 , -0.36298485, 0.83465055, -0.16837896], [0.31270191, 0.49217145, 1.35513683, 0.75695454, 0.08490823]]), array([[1.15533632, 1.03980352, -0.89423002, -1.65090936], [-0.21378335, -0.24850843, -0.07649171, 0.63585668], [0.16604087, -1.07703148, 1.50100522, 2.27829049]]), array([[1.1315117 , 1.90279542, -0.08008958, -2.33570889]])]
[array([[0. , 0.29605008, -0.09811087, -0.68220617, 0.3987858], [0. , -0.82921667, -0.15051907, 0.02859666, 0.02363924], [0. , -0.36163173, 0.19781599, -0.69882878, -0.12631162]]), array([[0. , 0.75729032, 0.0994972 , 0.31055328], [0. , 0.82886151, 0.46285758, 0.88303738], [0. , 0.68955374, -0.23515609, 0.81573478]]), array([[0. , -0.98547809, -0.89346138, -0.78340634]])]	0.0434786036	0.0418292611	[array([[0.72996605, 0.86428653, 0.8168921 , 0.89658004, -0.01467143], [-0.12791258, -0.86496191, -0.34937466, -0.13037078, 0.0980733], [-0.1484362 , 0.06346077, 1.07786578, 0.17769385, 0.14144109]]), array([[-0.00271948, 1.10242654, -0.35172459, 0.49250896], [-0.11116723, 1.65421473, -0.34078815, 0.98218541], [0.07615267, 1.13975593, -0.58825432, 0.78124868]]), array([[2.77334879, -1.72576759, -1.7766719 , -2.03151944]])]

Table for (2, 5) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , -0.38358387, 0.97537382, -0.49952964, 0.95732796], [0. , -0.31028284, -0.89070697, 0.90148064, 0.38838422]]), array([[0. , 0.88306796, -0.80274452], [0. , -0.84384441, -0.12842932], [0. , 0.74653466, -0.24295249], [0. , -0.86559364, -0.39985785], [0. , 0.57543276, -0.16695087]]), array([[0. , -0.84435048, 0.36892406, 0.17771847, -0.18016212, 0.85870786]])]	0.03393010208	0.04157077298	[array([[0.38895901, 0.12059843, 1.90130972, 0.66688777, 0.40801201], [0.06362066, -1.26751887, -0.96093749, -1.03915778, 0.03578168]]), array([[1.04936566, 1.24518469, -2.3786223], [0. , -0.84384441, -0.12842932], [0.17663148, 0.72743472, -0.2845977], [0. , -0.86559364, -0.39985785], [-0.00814357, 0.08315317, 1.87975595]]), array([[0.08048348, -3.63374563, 0.60609965, -0.51352263, -0.30246452, 1.66949776]])]
[array([[0. , 0.87641944, -0.20741613, -0.23546669, -0.88494066], [0. , -0.36117214, 0.09807229, 0.69256158, -0.01988752]]), array([[0. , 0.70042101, -0.23089042], [0. , -0.12651382, 0.88739659], [0. , 0.175626 , -0.58092406], [0. , -0.32968628, -0.32452854], [0. , 0.91127032, 0.2943613]]), array([[0. , -0.64414556, 0.70883954, -0.52881639, -0.88626219, -0.12493831]])]	0.01315766738	0.01610634459	[array([[1.21049035, 0.95045945, 1.15127881, 1.41447465, -0.08849792], [0.07870888, -1.31903974, -0.71428911, 0.35224967, 0.11841206]]), array([[-8.39713757e-04, 1.66856649e+00, -1.14289146e-01], [1.30064525e+00, -1.34374482e+00, 1.16508063e+00], [2.24942392e-02, 8.79010883e-01, -2.83118768e-01], [0.00000000e+00, -3.29686282e-01, -3.24528542e-01], [-4.15980835e-02, 1.44384837e+00, 8.54969070e-02]]), array([[0.5497152 , -1.54464348, 2.29331277, -1.1366648 , -0.85270255, -0.85869276]])]
[array([[0. , -0.4071985 , -0.19071826, -0.12478121, 0.48565458], [0. , -0.84782398, -0.71251037, 0.40284333, -0.02965351]]), array([[0. , -0.51380458, 0.31012524], [0. , -0.34653565, -0.31738447], [0. , 0.51758183, -0.32375462], [0. , -0.72104113, -0.31152356], [0. , 0.85548172, 0.61185007]]), array([[0. , 0.33721913, -0.62498857, 0.5306304 , 0.24610955, 0.90119014]])]	0.02641999741	0.02918013661	[array([[-1.16993484e-03, -8.42188073e-01, -1.09040689e+00, -1.47173240e+00, 1.72910805e-01], [-2.50922312e-01, -1.31511323e+00, -8.93939834e-01, -2.17147909e-01, 8.71689590e-02]]), array([[-0.05956755, -0.45790298, 0.36200049], [0. , -0.34653565, -0.31738447], [-0.1127253 , 1.06035767, 0.20485745], [0. , -0.72104113, -0.31152356], [-0.36006183, 2.08271037, 1.33631518]]), array([[-3.47062965, 0.29658498, -0.96975061, 1.11786375, -0.3404079 , 2.57255618]])]

Table for (4, 3) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , -0.23135663, -0.38924803, -0.09266551, -0.74182912], [0. , 0.14522036, -0.16681688, -0.21054762, 0.01509666], [0. , 0.04103173, 0.38754162, 0.47683055, -0.0892612], [0. , -0.38706609, 0.32308608, -0.66921615, 0.14681234]]), array([[0. , 0.27441028, -0.01580607, -0.76629565, 0.3442677], [0. , -0.34366824, 0.73594878, 0.10831493, -0.17322786], [0. , -0.56459324, -0.04656301, -0.26794632, 0.67607342]]), array([[0. , 0.7968329 , 0.17493398, -0.24306826]])]	0.008111331133	0.009917745456	[array([[0.23149742, -1.02055679, -1.18103808, -0.82087084, -0.1486276], [-0.02330085, 0.16953864, -0.15931627, -0.21355396, 0.01049491], [1.13188055, 0.87657915, 1.01764314, 0.9699241 , -0.12514612], [0.87401978, 0.21779662, 0.31833404, -1.56649667, 0.29518744]]), array([[3.97267790e-01, 1.67384221e+00, 1.36511581e-03, -1.83879416e+00, 7.60558519e-01], [6.68967030e-02, -3.47814265e-01, 7.36605859e-01, 2.37320935e-01, -1.34932472e-01], [1.02121356e+00, -1.08536574e+00, -8.70665986e-02, 1.14888554e+00, 6.31408795e-01]]), array([[-1.86962673, 3.70020982, -0.32601257, -1.45971483]])]
[array([[0. , -0.32320882, -0.49148416, -0.25042711, -0.03464125], [0. , -0.6450352 , 0.69045898, -0.42842338, 0.62336389], [0. , -0.26497411, 0.32904267, 0.54785383, -0.32390273], [0. , 0.52077468, 0.19327997, -0.44864597, -0.14419259]]), array([[0. , -0.42619051, 0.5179254 , 0.7782328 , 0.57152072], [0. , 0.75178804, -0.27314109, -0.31663117, 0.34205761], [0. , -0.29684716, 0.79985392, 0.11141238, -0.12073191]]), array([[0. , 0.94849948, -1.17056555, 0.29616781]])]	0.005087114771	0.005653787916	[array([[-0.28535112, -0.26336202, -0.39454593, -0.10666801, 0.03606775], [0.19353966, -0.89039462, -1.10105016, -1.60439738, 0.25421847], [-0.17832182, -1.14941612, -0.36704863, 0.59158242, 0.08123796], [1.04812986, 0.66358304, 0.81375987, 0.58399072, 0.06470764]]), array([[-0.04849029, 0.31486109, 1.61000239, 0.99729065, -0.4014424], [1.25547258, -0.30210101, -1.25718178, -0.38112807, 1.51096467], [0.08384748, -0.03702894, 1.30671617, 0.18702178, -0.43533961]]), array([[-0.36541559, 1.92609512, -2.46161313, 1.1033552]])]
[array([[0. , 0.63508362, -0.63305109, 0.73568756, -0.01496114], [0. , 0.05971041, -0.60846881, -0.14719865, 0.37949282], [0. , 0.73846436, -0.72700944, -0.12419221, 0.85367475], [0. , 0.68419715, 0.81929858, -0.162456 , 0.17125376]]), array([[0. , 0.19309994, -0.12499107, 0.48502426, 0.11871538], [0. , 0.828549 , 0.83198909, 0.4595852 , 0.61873851], [0. , 0.3947747 , -0.89021903, -0.58953667, -0.61017717]]), array([[0. , 0.0802085 , -0.02421601, 1.20935564]])]	0.01415768256	0.01596017368	[array([[0.08267031, 1.11127324, -0.18012041, 0.40608738, -0.05590489], [-0.42481796, 0.43148096, -0.25417642, -0.08707069, 0.09835625], [-0.25849174, 1.29604586, -0.27105349, 0.38996918, -0.04454326], [0.88561686, 0.87708772, 1.30166503, 1.02245639, 0.1062838]]), array([[1.92245599e-03, 1.77580418e-01, -1.60196752e-01, 4.62612224e-01, 3.71431938e-01], [1.02339676e-02, 7.49745376e-01, 6.01813606e-01, 3.85855219e-01, 2.21492903e+00], [1.72811220e+00, -3.60968793e-01, -1.88733879e+00, -2.70063951e+00, -2.90386426e+00]]), array([[0.16550156, -0.30657118, -1.62461894, 3.69184184]])]

Table for (2, 6) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , 0.76453235, -0.05733203, 0.46117688, -0.96908986], [0. , 0.02750987, 0.237925 , -0.1894247 , 0.72397237]]), array([[0. , -0.23189434, 0.34939123], [0. , -0.53807123, 0.58001815], [0. , 0.49281186, -0.79533306], [0. , -0.16790428, -0.71639109], [0. , 0.44449202, 0.85709985], [0. , -0.35719689, 0.38666827]]), array([[0. , -0.80210993, 0.83705236, 0.41692017, 0.52363473, -0.70393632, 0.89860735]])]	0.02534259271	0.03417660916	[array([[0.93837575, 1.05942221, 1.10703764, 1.01799725, -0.01666809], [0.32264896, 0.29750688, 0.7765665 , 0.47108411, 0.2573464]]), array([[-4.23236329e-04, 6.72087616e-01, 6.95602632e-01], [1.21948349e+00, -1.60104220e+00, -1.17763469e-01], [-8.02846121e-02, 3.40552738e-01, -8.77717320e-01], [0.00000000e+00, -1.67904277e-01, -7.16391088e-01], [-6.58378264e-04, 1.70740519e+00, 1.36486783e+00], [1.24729649e+00, -1.47489057e+00, -3.65400016e-01]]), array([[0.56790362, -1.04466137, 1.80684558, 0.2672932 , 0.76854429, -1.63275474, 1.82022619]])]
[array([[0. , 0.22953868, 0.56536762, -0.85473827, 0.74775231], [0. , -0.9279414 , 0.02551461, 0.69525658, 0.65451852]]), array([[0. , -0.07786227, -0.0843151], [0. , 0.66549723, 0.49432753], [0. , -0.51354543, 0.08386406], [0. , 0.06259645, 0.28225865], [0. , -0.57734364, 0.16318119], [0. , 0.63354981, -0.53294359]]), array([[0. , -0.41512803, 0.28430117, -0.37904799, 0.91318896, 0.14004283, -0.51161598]])]	0.01829907171	0.02410089878	[array([[0.86114247, 0.42586981, 0.70626424, 0.57407045, 0.72412625], [0.03500775, -1.48203448, -1.56492264, -1.45363768, 0.10845883]]), array([[0. , -0.07786227, -0.0843151], [-0.1108845 , 0.17770117, 1.22452833], [-0.03085501, -0.46006231, 0.01058792], [-0.01767322, -0.40807696, 1.91783557], [0.06649718, -0.97690715, 0.58521174], [1.65127436, 1.08184408, -2.21054738]]), array([[-0.76283689, -0.46759877, 0.89120915, -0.12973342, 1.96229639, 0.72122649, -2.65504482]])]
[array([[0. , -0.2334822 , -0.06250813, -0.9909093 , -0.60872672], [0. , 0.19666968, -0.11948805, 0.37947263, -0.3423889]]), array([[0. , -0.0089158 , 0.16417409], [0. , -0.84794252, 0.0376765], [0. , -0.43756713, 0.11141125], [0. , 0.03786805, -0.45015225], [0. , 0.8048924 , 0.08064609], [0. , 0.61188416, 0.06145858]]), array([[0. , -0.61529399, -0.4587879 , -0.63699513, 0.80319541, -0.03735104, 0.80396262]])]	0.01956453886	0.02811682111	[array([[0.17677237, -1.07576425, -1.39605958, -1.19473978, -0.13715416], [1.09191116, 0.88206947, 0.77944005, 0.77809727, -0.07973603]]), array([[0.44897197, -0.627327 , 1.06758948], [0.71192171, -0.97450871, 0.73714044], [0.66737224, -0.84679591, 0.90153283], [0.68691852, 1.10155111, -0.92599272], [0.00180439, 0.94212523, -0.00158252], [0.37621478, 1.32176 , -0.48499559]]), array([[-0.29372346, -1.37561123, -1.33318658, -1.44577081, 1.74147789, 0.44508169, 1.57128234]])]

Table for (3, 5) Configuration:

Initialized Biases and Weights	Training Loss	Validation Loss	Trained Biases and Weights
[array([[0. , -0.45745902, 0.24058203, -0.12823107, 0.30934213], [0. , -0.51451322, 0.65921142, -0.12114722, 0.69739779], [0. , 0.2378663 , -0.11672259, -0.35134529, 0.7633815]]), array([[0. , -0.14506016, 0.66406188, -0.02211435], [0. , -0.8657176 , -0.84237695, 0.58852733], [0. , 0.05793561, 0.86533746, -0.77299856], [0. , -0.25830244, -0.3032291 , 0.43287475], [0. , 0.81944913, 0.78186696, 0.10676909]]), array([[0. , -0.33981772, -0.93468569, 0.90979972, -0.89719012, -0.53571308]])]	0.00437353426	0.006822249547	[array([[0.20594172, -0.51138545, -0.59144921, -0.95502425, 0.05421314], [-0.26558529, -1.28678611, -0.33807009, -0.05004738, 0.51780719], [0.96265355, 0.5732705 , 1.06871695, 0.96614316, 0.04767706]]), array([[-0.12798268, -0.26300765, 0.55736329, 0.03662103], [0.90395 , -1.13672458, -1.13986613, 1.34045187], [0.89775396, 1.08338816, 1.4398182 , -1.57570941], [0.47554276, -0.61046802, -0.58376448, 1.04224533], [-0.15082621, 0.64255203, 0.66367437, 0.35391383]]), array([[0.36115808, -0.28109708, -2.19306414, 2.23818092, -1.54844153, 0.05362499]])]
[array([[0. , 0.53466839, -0.14508919, 0.8350686 , 0.76227867], [0. , -0.86518393, 0.4877055 , 0.30057571, -0.66425496], [0. , -0.37111659, -0.11570314, -0.83402276, -0.12267373]]), array([[0. , 0.15675812, 0.84359863, -0.09821425], [0. , -0.73318395, -0.13273869, -0.13425124], [0. , 0.4334813 , 0.12684447, 0.14619765], [0. , 0.29888044, -0.43919301, 0.29873371], [0. , 0.3844131 , -0.33717737, 0.40217515]]), array([[0. , 0.883148 , 0.16902411, 0.77296752, 0.66535736, 0.52732807]])]	0.01926383424	0.02154462766	[array([[-0.2224881 , 0.08435299, -0.16626507, 0.73411116, 0.61733011], [-0.39328707, -1.27451794, -0.45249738, 0.33940947, 0.04221649], [0.07724245, -0.99097445, -1.27122464, -1.48647627, 0.14994984]]), array([[-0.39074495, -0.14321861, 0.94964724, 0.93498396], [0. , -0.73318395, -0.13273869, -0.13425124], [-0.43028754, -0.02801792, 0.48895954, 1.26901695], [-0.24464841, 0.03433731, 0.21519182, 1.19126451], [-0.30407915, 0.04722016, 0.25561117, 1.22498908]]), array([[-3.78279748, 1.91436195, 1.02259779, 2.10600169, 1.5940211 , 1.6421934]])]
[array([[0. , 0.87644098, 0.20846506, 0.55467896, 0.67698399], [0. , -0.25041455, -0.66205216, 0.79843416, 0.78288191], [0. , 0.12057947, 0.0071268 , -0.91186587, -0.27342602]]), array([[0. , 0.02309879, 0.34587565, 0.23560323], [0. , -0.268335 , 0.16883381, -0.25355995], [0. , -0.0991596 , 0.27189779, 0.01127241], [0. , 0.31417713, 0.29211281, 0.46137117], [0. , 0.25968116, 0.83473207, 0.50845438]]), array([[0. , 0.69339763, 0.05604293, 0.11118644, 0.03015911, 0.4961323]])]	0.006604889806	0.006557043869	[array([[1.10374319, 0.92964824, 1.08491032, 1.11635296, -0.03915278], [-0.03416041, -0.87478885, -0.58255979, -0.07991175, 0.67467475], [0.33388446, -0.61068959, -0.58395896, -1.39939544, -0.03113897]]), array([[0.15392538, -1.6472658 , 1.70981642, 1.19777563], [0.00701314, -0.30283625, 0.24936167, -0.22955814], [0.05619852, -0.49036221, 0.54575227, 0.28505875], [0.48265762, 1.82743923, -0.7150662 , -0.11377795], [-0.05792651, -0.82064152, 1.43812844, 0.98293649]]), array([[-0.94564605, 2.12837361, 0.47740605, 0.56627598, -1.47888981, 1.11508484]])]

To get the best (n_1, n_2) neural network configuration, the average validation losses for each (n_1, n_2) configuration are compared for three runs and the one with minimum average validation loss is chosen as the best model configuration. After deciding the best model configuration, the one with the minimum validation loss over the three models for that particular best model configuration is chosen as the final model and the weights are weights corresponding to that final model.

```
Best Model Configuration is: (4, 3)
Best Model Weights is: [array([[ -0.28535112, -0.26336202, -0.39454593, -0.10666801,  0.03606775],
 [ 0.19353966, -0.89039462, -1.10105016, -1.60439738,  0.25421847],
 [ -0.17832182, -1.14941612, -0.36704863,  0.59158242,  0.08123796],
 [ 1.04812986,  0.66358304,  0.81375987,  0.58399072,  0.06470764]]), array([[ -0.04849029,  0.31486109,  1.6100023
 9,  0.99729065, -0.4014424 ],
 [ 1.25547258, -0.30210101, -1.25718178, -0.38112807,  1.51096467],
 [ 0.08384748, -0.03702894,  1.30671617,  0.18702178, -0.43533961]]), array([[ -0.36541559,  1.92609512, -2.4616131
 3,  1.1033552 ]])]
```

Figure 4.1: Best Model Configuration and the Weights of the Final Model

As can be seen from Figure 4.1, the best model configuration is (4, 3) that is 4 units in the first hidden layer and 3 units in the second hidden layer. The best model configuration and weights for the final model is obtained using the approach mentioned above.

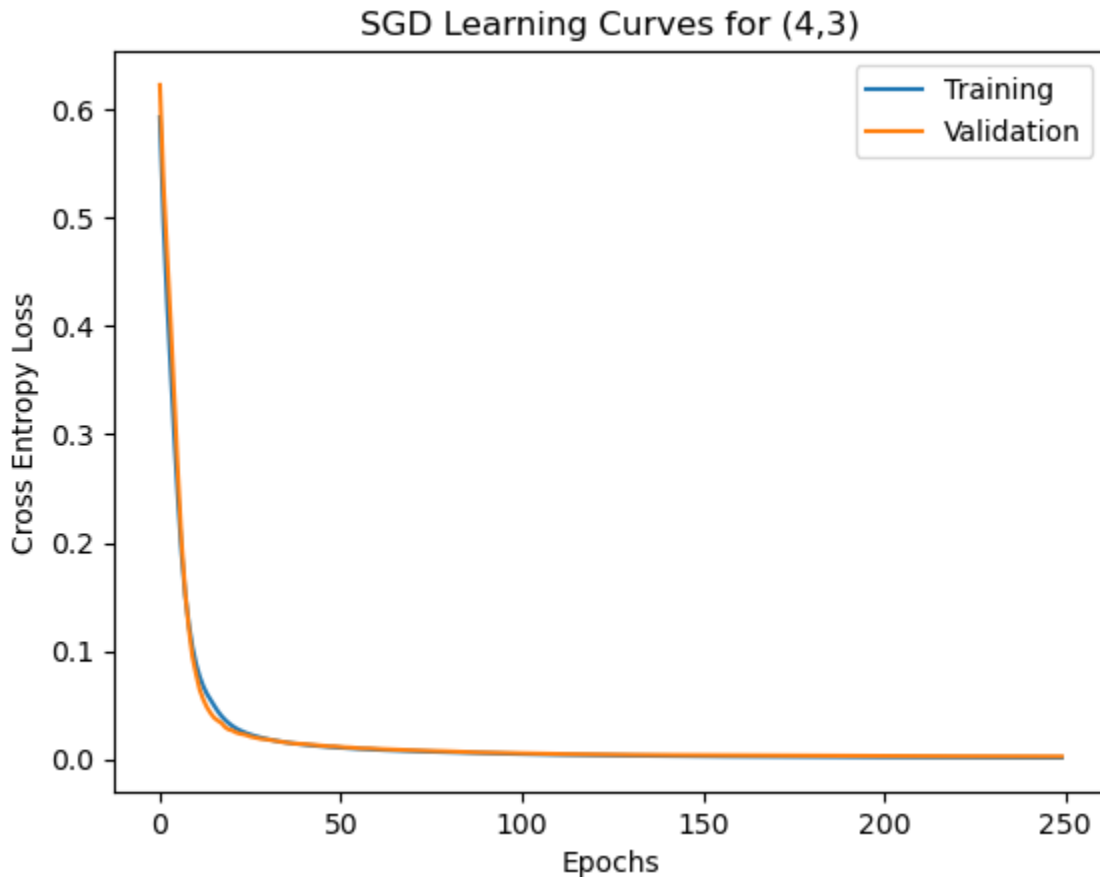


Figure 4.2: Training and Validation Learning Curves for the Final Model

```
Misclassification Error on Training: 0.0  
Misclassification Error on Testing: 0.0  
Misclassification Error on Validation: 0.0
```

Figure 4.3: Misclassification Error on Training, Testing, and Validation Using the Final Model

As can be seen from the above Figure 4.2, the training and validation curves are on top of each other. The U-shape could not be observed here, maybe because the learning rate is fixed for all epochs instead of decreasing gradually over epochs, or due to having less number of examples in the dataset, not using regularization or also could be due to the initialization of the weights.

Figure 4.3, shows the misclassification error on training, testing, and validation for the final model which are all at 0.0 meaning there are no FP's and FN's. The above values suggest the model is performing really well and this could be the case since the dataset has really few examples of around 1372.

For most of the configurations, the learning curves look similar, that is both training loss and validation loss plateau after a certain number of epochs. Also, in most cases, the validation loss is higher or greater than the training loss, which makes sense because the model is trained using the training dataset, therefore it should have lesser loss compared to validation loss. However, in some situations the validation loss is greater than the training loss maybe because of the way weights are initialized.

For the given dataset, the classifier performs very well since the misclassification error on training, testing, and validation are all zero. If there is more time, to get an asymmetric U-shaped validation learning curve and also to know the model or classifier performance better, I would have tried using this model or classifier on a different dataset that is similar to the given one and get the performance or try creating more examples using the dataset to test further. Also, try using regularization, trying different weight initialization techniques, using different activation functions for hidden layers, adding more layers instead of just two, initializing biases to non-zero values, and also having an adaptable learning rate instead of a constant one.