In []: %config IPCompleter.greedy=True %matplotlib inline from numpy import * from matplotlib.pyplot import * from IPython.display import * import warnings warnings.simplefilter('ignore', DeprecationWarning) from sklearn import * iris = datasets.load iris() X=iris.data; Y=iris.target 1 Performance Tuning of a Neural Net (8 points) 1.0 Baseline Performance SVM can reach an classification accuracy ~ 8x% correct for the HARD Iris problem. sss=model_selection.StratifiedShuffleSplit(n_splits=3,test_size=0.1) # (45+45+45) vs. (5+5+5) model=svm.SVC(C=10) acc=[] for train index, test index in sss.split(X, Y): # 3-fold cross-validation X_train, X_test = X[train_index], X[test_index] Y train, Y test = Y[train index], Y[test index] model.fit(X train[:,0:2],Y train) #training acc.append(model.predict(X_test[:,0:2])==Y_test) # testing print(np.mean(acc)) 0.8666666666666667 1.1 Tuning your ANN (4 points) Tune your model hyperparameters (# of layers, # of units in each layer, activation function, optimizer, epochs, batch_size, etc.) to see if you can push your ANN performance up to ~9x% correct for the HARD iris problem. In []: from keras.models import Sequential, clone model from keras.layers import Dense from keras.utils import to categorical import tensorflow as tf from keras.layers import Dropout from keras import regularizers from keras.callbacks import LearningRateScheduler def lr schedule(epoch): **return** 0.1 * (0.6 ** int(epoch / 15)) optimizer = tf.keras.optimizers.legacy.SGD(lr=0.02) # Adjust the learning rate based on experimentation model = Sequential() model.add(Dense(units=48, activation='relu', input_dim=2, kernel_regularizer=regularizers.12(0.02))) # Adjust units based on experimentation model.add(Dropout(0.5)) model.add(Dense(units=40, activation='relu')) model.add(Dense(units=3, activation='softmax')) acc=[] for train_index, test_index in sss.split(X, Y): # 3-fold cross-validation X train, X test = X[train index], X[test index] Y_train, Y_test = Y[train_index], Y[test_index] new_model=clone_model(model) # Otherwise the old model will keep learning new model.compile(loss='categorical crossentropy',optimizer=optimizer,metrics=['accuracy']) new_model.fit(X_train[:, 0:2], to_categorical(Y_train), epochs=55, batch_size=15, callbacks=[LearningRateScheduler(lr_schedule)]) acc.append(np.mean(argmax(new_model.predict(X_test[:,0:2]),1)==Y_test)) # testing print(acc, np.mean(acc)) Epoch 1/55 Epoch 2/55 Epoch 3/55 Epoch 5/55 Epoch 6/55 Epoch 7/55 Epoch 8/55 1/9 [==>.....] - ETA: 0s - loss: 0.9728 - accuracy: 0.6000 /Library/Frameworks/Python.framework/Versions/3.11/lib/python3.11/site-packages/keras/src/optimizers/legacy/gradient descent.py:114: UserWarning: The `lr argument is deprecated, use `learning rate` instead. super(). init (name, **kwargs) Epoch 9/55 1/9 [==>....... 0.4667Epoch 9/55 Epoch 11/55 Epoch 12/55 Epoch 14/55 Epoch 15/55 Epoch 17/55 Epoch 18/55 Epoch 20/55 Epoch 21/55 Epoch 23/55 Epoch 24/55 Epoch 26/55 Epoch 27/55 Epoch 29/55 Epoch 30/55 Epoch 31/55 Epoch 32/55 Epoch 33/55 Epoch 34/55 Epoch 35/55 Epoch 36/55 Epoch 37/55 Epoch 38/55 Epoch 39/55 Epoch 40/55 Epoch 41/55 Epoch 42/55 Epoch 44/55 Epoch 45/55 Epoch 47/55 Epoch 48/55 Epoch 50/55 Epoch 51/55 Epoch 53/55 Epoch 54/55 1/1 [=======] - 0s 20ms/step Epoch 2/55 Epoch 3/55 Epoch 5/55 Epoch 6/55 Epoch 8/55 Epoch 9/55 Epoch 11/55 Epoch 12/55 Epoch 14/55 Epoch 15/55 Epoch 16/55 Epoch 17/55 Epoch 18/55 Epoch 19/55 Epoch 20/55 Epoch 21/55 Epoch 22/55 Epoch 23/55 Epoch 24/55 Epoch 25/55 Epoch 26/55 Epoch 27/55 Epoch 28/55 Epoch 29/55 Epoch 30/55 Epoch 32/55 Epoch 33/55 Epoch 35/55 Epoch 36/55 Epoch 38/55 Epoch 39/55 Epoch 40/55 Epoch 41/55 Epoch 42/55 9/9 [====== =========] - 0s 440us/step - loss: 0.6703 - accuracy: 0.7037 - lr: 0.0360 Epoch 43/55 Epoch 44/55 Epoch 45/55 Epoch 47/55 Epoch 48/55 Epoch 50/55 Epoch 51/55 Epoch 53/55 Epoch 54/55 1/1 [========] - 0s 20ms/step Epoch 2/55 Epoch 3/55 Epoch 5/55 Epoch 6/55 Epoch 8/55 Epoch 9/55 Epoch 11/55 Epoch 12/55 Epoch 14/55 Epoch 15/55 Epoch 16/55 Epoch 17/55 Epoch 18/55 Epoch 19/55 Epoch 20/55 Epoch 21/55 Epoch 22/55 Epoch 23/55 Epoch 24/55 Epoch 25/55 Epoch 26/55 Epoch 27/55 Epoch 28/55 Epoch 29/55 Epoch 30/55 Epoch 31/55 Epoch 32/55 Epoch 33/55 Epoch 34/55 Epoch 35/55 Epoch 36/55 Epoch 37/55 Epoch 38/55 Epoch 39/55 Epoch 40/55 Epoch 41/55 Epoch 42/55 Epoch 43/55 Epoch 44/55 Epoch 45/55 Epoch 47/55 Epoch 48/55 Epoch 50/55 Epoch 51/55 Epoch 53/55 Epoch 54/55 1/1 [=======] - 0s 20ms/step 1.2 Is your (deep) network better than SVM? Why or why not? (4 points) After tuning model hyperparameters with different combinations, I observe that the accuracy of my network is only between 6x% to 8x%. Hence, my network is not better than SVM. Following explanations may be the reasons that my network is not better than the SVM: 1. Data Size and Complexity: The Iris dataset is small, consisting of only 150 samples. Deep networks often require large amounts of data to effectively learn complex patterns. If the dataset is not sufficiently complex, a simpler model like SVM may generalize better. 2. Model Complexity & Overfitting: The complexity of my network might be excessive for the simplicity of the Iris dataset. Increasing the number of layers and units might lead to overfitting, especially when the dataset is small. 3. Hyperparameter Tuning: Maybe there will be a combination of the hyperparameters which makes the accuracy higher than SVM, but through my experiences for over hundredes of times, I still haven't found it. 4. Regularization and Dropout: The use of regularization (e.g., L2 regularization) and dropout in my network may not be suitable for the given problem. Sometimes, too much regularization can hinder the learning process.

Psychoinformatics - Week 11 (Exercises)

by 徐舒庭 (b11705018@ntu.edu.tw)