GAD-cnn-lstm-gru

September 18, 2024

1 model 1

sara

2 ————————————————————————————————-

3 imports

[1]:
```
import tensorflow as tf
import keras
from keras.models import load_model
from keras.callbacks import ModelCheckpoint, EarlyStopping
from keras_tqdm import TQDMNotebookCallback
import numpy as np
np.random.seed(1337)
from keras_tqdm import TQDMNotebookCallback
import nltk
import xml.etree.ElementTree as ET
import pandas as pd
import os
import string
from nltk.tokenize import TreebankWordTokenizer
from numpy.random import random_sample
import re
import pickle
from sklearn.metrics import accuracy_score
from sklearn.metrics import classification_report

from keras.layers import Embedding, Flatten,LSTM, GRU
from keras.layers.convolutional import Conv2D, MaxPooling2D
from tensorflow.keras.utils import to_categorical
from tensorflow.keras.models import Sequential, Model
from keras.layers import Dense, Dropout, Activation,  Input,␣
 ↪merge,Conv1D,MaxPooling1D,GlobalMaxPooling1D,Convolution1D
from keras import regularizers
from sklearn.metrics import precision_recall_fscore_support
from sklearn.model_selection import StratifiedKFold
import matplotlib.pyplot as plt
from keras.layers import Concatenate, concatenate
from keras import backend as K
from keras.layers import multiply
from keras.layers import merge
from keras.layers.core import *
from keras.layers.recurrent import LSTM
from keras.models import *
random_seed=1337
```

```
Using TensorFlow backend.
```

3.0.1 Define Callback functions to generate Measures

[2]:
```
from keras import backend as K

def f1(y_true, y_pred):
    def recall(y_true, y_pred):
        true_positives = K.sum(K.round(K.clip(y_true * y_pred, 0, 1)))
        possible_positives = K.sum(K.round(K.clip(y_true, 0, 1)))
        recall = true_positives / (possible_positives + K.epsilon())
        return recall

    def precision(y_true, y_pred):
        true_positives = K.sum(K.round(K.clip(y_true * y_pred, 0, 1)))
        predicted_positives = K.sum(K.round(K.clip(y_pred, 0, 1)))
        precision = true_positives / (predicted_positives + K.epsilon())
        return precision
    precision = precision(y_true, y_pred)
    recall = recall(y_true, y_pred)
    return 2*((precision*recall)/(precision+recall+K.epsilon()))
```

4 Experiments to reproduce the results of Table 9

4.0.1 Load pre procssed Data

[3]:
```
with open('../data/pickles/befree_3class_crawl-300d-2M.pickle', 'rb') as handle:
    gene_id_list = pickle.load(handle)
    gene_symbol_list = pickle.load(handle)
    disease_id_list = pickle.load(handle)
    X_train = pickle.load(handle)
    distance1_vectors = pickle.load(handle)
    distance2_vectors = pickle.load(handle)
    Y_train = pickle.load(handle)
    word_list = pickle.load(handle)
    word_vectors = pickle.load(handle)
    word_dict = pickle.load(handle)
    distance1_dict = pickle.load(handle)
    distance2_dict = pickle.load(handle)
    label_dict = pickle.load(handle)
    MAX_SEQUENCE_LENGTH = pickle.load(handle)
print ("word_vectors",len(word_vectors))
```

```
word_vectors 6766
```

position embedding

[4]:
```
import keras
from keras_pos_embd import TrigPosEmbedding

model = keras.models.Sequential()
model.add(TrigPosEmbedding(
    input_shape=(None,),
    output_dim=20,                      # The dimension of embeddings.
    mode=TrigPosEmbedding.MODE_EXPAND,  # Use `expand` mode
    name='Pos-Embd',
))
model.compile('adam', keras.losses.mae, {})


d1_train_embedded=model.predict(distance1_vectors)

d1_train_embedded.shape

d2_train_embedded=model.predict(distance2_vectors)

d2_train_embedded.shape
```

[4]:
```
(5330, 81, 20)
```

4.0.2 Prepare Word Embedding Layer

[5]:
```
EMBEDDING_DIM=word_vectors.shape[1]
print("EMBEDDING_DIM=",EMBEDDING_DIM)
embedding_matrix=word_vectors

def create_embedding_layer(l2_reg=0.01,use_pretrained=True,is_trainable=False):

    if use_pretrained:
        return Embedding(len(word_dict)␣
 ↪,EMBEDDING_DIM,weights=[embedding_matrix],input_length=MAX_SEQUENCE_LENGTH,trainable=is_tra
 ↪l2(l2_reg))
```

```
    else:
        return Embedding(len(word_dict)␣
↪,EMBEDDING_DIM,input_length=MAX_SEQUENCE_LENGTH)
```

```
EMBEDDING_DIM= 300
```

attention

[6]:
```
INPUT_DIM = 2
TIME_STEPS = MAX_SEQUENCE_LENGTH
def attentionNew(inputs):
    inputs = Lambda(lambda x: tf.keras.backend.tanh(x))(inputs)
    input_dim = int(inputs.shape[2])
    a = Permute((2, 1))(inputs)
    a = Dense(TIME_STEPS, activation='softmax')(a)
    a_probs = Permute((2, 1))(a)
    output_attention_mul = multiply([inputs, a_probs])
    output_attention_mul = Lambda(lambda x: tf.keras.backend.
↪tanh(x))(output_attention_mul)
    return output_attention_mul
```

4.0.3 Create the Model

[7]:
```
# set parameter for metric calculation, 'macro' for multiclass classification
param='macro'
def build_model():

    sequence_input = Input(shape=(MAX_SEQUENCE_LENGTH,), dtype='int32')
   ␣
↪embedding_layer=create_embedding_layer(use_pretrained=True,is_trainable=False)
    embedded_sequences = embedding_layer(sequence_input)


    pos_embedd_1=Input(shape=(MAX_SEQUENCE_LENGTH,20), dtype='float32')
    pos_embedd_2=Input(shape=(MAX_SEQUENCE_LENGTH,20), dtype='float32')

    embedded_sequences =␣
↪concatenate([embedded_sequences,pos_embedd_1,pos_embedd_2])


    x = Conv1D(32, 5, activation='relu')(embedded_sequences)
    x = MaxPooling1D(3)(x)
    x = Dropout(0.1)(x)
    conv_sequence_w5=GlobalMaxPooling1D()(x)     #x = Flatten()(x)
    x = Conv1D(64, 3, activation='relu')(embedded_sequences)
    x = MaxPooling1D(3)(x)
    x = Dropout(0.1)(x)
    conv_sequence_w4=GlobalMaxPooling1D()(x)    #x = Flatten()(x)



    x = Conv1D(128, 3, activation='relu')(embedded_sequences)
    x = MaxPooling1D(3)(x)
    x = Dropout(0.1)(x)
    conv_sequence_w3=GlobalMaxPooling1D()(x)    #x = Flatten()(x)

    forward = GRU(100, recurrent_dropout=0.
↪05,return_sequences=True)(embedded_sequences)
    backward = LSTM(100, go_backwards=True,recurrent_dropout=0.
↪05,return_sequences=True)(embedded_sequences)
    attention_forward=attentionNew(forward)
    attention_backward=attentionNew(backward)
    lstm_sequence = concatenate([attention_forward,attention_backward])

    lstm_sequence = Flatten()(lstm_sequence)
    merge =␣
↪concatenate([conv_sequence_w5,conv_sequence_w4,conv_sequence_w3,lstm_sequence])
    x = Dense(64, activation='relu', kernel_regularizer=regularizers.l2(0.
↪1))(merge)
    x = Dropout(0.1)(x)
    preds = Dense(3, activation='softmax')(x)
    model = Model(inputs=[sequence_input,␣
↪pos_embedd_1,pos_embedd_2],outputs=preds)
    opt=tf.keras.optimizers.Adam()
    model.
↪compile(loss='categorical_crossentropy',optimizer=opt,metrics=['acc',f1])
    return model
```

[8]:
```
model = build_model()
model.summary()
```

```
Model: "model_1"
________________________________________________________________________________
__________________
Layer (type)                    Output Shape         Param #     Connected to
================================================================================
==================
input_1 (InputLayer)            (None, 81)           0
________________________________________________________________________________
```

```
__________________
embedding_1 (Embedding)          (None, 81, 300)      2029800     input_1[0][0]
________________________________________________________________________________
__________________
input_2 (InputLayer)             (None, 81, 20)       0
________________________________________________________________________________
__________________
input_3 (InputLayer)             (None, 81, 20)       0
________________________________________________________________________________
__________________
concatenate_1 (Concatenate)      (None, 81, 340)      0
embedding_1[0][0]
                                                                  input_2[0][0]
                                                                  input_3[0][0]
________________________________________________________________________________
__________________
gru_1 (GRU)                      (None, 81, 100)      132300
concatenate_1[0][0]
________________________________________________________________________________
__________________
lstm_1 (LSTM)                    (None, 81, 100)      176400
concatenate_1[0][0]
________________________________________________________________________________
__________________
lambda_1 (Lambda)                (None, 81, 100)      0           gru_1[0][0]
________________________________________________________________________________
__________________
lambda_3 (Lambda)                (None, 81, 100)      0           lstm_1[0][0]
________________________________________________________________________________
__________________
permute_1 (Permute)              (None, 100, 81)      0           lambda_1[0][0]
________________________________________________________________________________
__________________
permute_3 (Permute)              (None, 100, 81)      0           lambda_3[0][0]
________________________________________________________________________________
__________________
dense_1 (Dense)                  (None, 100, 81)      6642        permute_1[0][0]
________________________________________________________________________________
__________________
dense_2 (Dense)                  (None, 100, 81)      6642        permute_3[0][0]
________________________________________________________________________________
__________________
permute_2 (Permute)              (None, 81, 100)      0           dense_1[0][0]
________________________________________________________________________________
__________________
permute_4 (Permute)              (None, 81, 100)      0           dense_2[0][0]
________________________________________________________________________________
__________________
```

```
conv1d_1 (Conv1D)               (None, 77, 32)       54432
concatenate_1[0][0]

--------------------------------------------------------------------------------
------------------
conv1d_2 (Conv1D)               (None, 79, 64)       65344
concatenate_1[0][0]

--------------------------------------------------------------------------------
------------------
conv1d_3 (Conv1D)               (None, 79, 128)      130688
concatenate_1[0][0]

--------------------------------------------------------------------------------
------------------
multiply_1 (Multiply)           (None, 81, 100)      0            lambda_1[0][0]
                                                                  permute_2[0][0]

--------------------------------------------------------------------------------
------------------
multiply_2 (Multiply)           (None, 81, 100)      0            lambda_3[0][0]
                                                                  permute_4[0][0]

--------------------------------------------------------------------------------
------------------
max_pooling1d_1 (MaxPooling1D)  (None, 25, 32)       0            conv1d_1[0][0]

--------------------------------------------------------------------------------
------------------
max_pooling1d_2 (MaxPooling1D)  (None, 26, 64)       0            conv1d_2[0][0]

--------------------------------------------------------------------------------
------------------
max_pooling1d_3 (MaxPooling1D)  (None, 26, 128)      0            conv1d_3[0][0]

--------------------------------------------------------------------------------
------------------
lambda_2 (Lambda)               (None, 81, 100)      0
multiply_1[0][0]

--------------------------------------------------------------------------------
------------------
lambda_4 (Lambda)               (None, 81, 100)      0
multiply_2[0][0]

--------------------------------------------------------------------------------
------------------
dropout_1 (Dropout)             (None, 25, 32)       0
max_pooling1d_1[0][0]

--------------------------------------------------------------------------------
------------------
dropout_2 (Dropout)             (None, 26, 64)       0
max_pooling1d_2[0][0]

--------------------------------------------------------------------------------
------------------
dropout_3 (Dropout)             (None, 26, 128)      0
max_pooling1d_3[0][0]

--------------------------------------------------------------------------------
```

```
__________________
concatenate_2 (Concatenate)     (None, 81, 200)      0           lambda_2[0][0]
                                                                 lambda_4[0][0]

________________________________________________________________________________
__________________
global_max_pooling1d_1 (GlobalM (None, 32)           0           dropout_1[0][0]
________________________________________________________________________________
__________________
global_max_pooling1d_2 (GlobalM (None, 64)           0           dropout_2[0][0]
________________________________________________________________________________
__________________
global_max_pooling1d_3 (GlobalM (None, 128)          0           dropout_3[0][0]
________________________________________________________________________________
__________________
flatten_1 (Flatten)             (None, 16200)        0
concatenate_2[0][0]

________________________________________________________________________________
__________________
concatenate_3 (Concatenate)     (None, 16424)        0
global_max_pooling1d_1[0][0]
global_max_pooling1d_2[0][0]
global_max_pooling1d_3[0][0]
                                                                 flatten_1[0][0]

________________________________________________________________________________
__________________
dense_3 (Dense)                 (None, 64)           1051200
concatenate_3[0][0]

________________________________________________________________________________
__________________
dropout_4 (Dropout)             (None, 64)           0           dense_3[0][0]
________________________________________________________________________________
__________________
dense_4 (Dense)                 (None, 3)            195         dropout_4[0][0]
================================================================================
==================
Total params: 3,653,643
Trainable params: 1,623,843
Non-trainable params: 2,029,800

________________________________________________________________________________
__________________
```

[9]:
```
validation_split_rate = 0.1
skf = StratifiedKFold(n_splits=5,shuffle=True, random_state=42)
Y = [np.argmax(y, axis=None, out=None) for y in Y_train]
for tr_index, te_index in skf.split(X_train,Y):
    test_index = te_index
    train_index = tr_index
```

```
trainRate = (len(train_index)/len(Y))*100
testRate = (len(test_index)/len(Y))*100
print ("TrainRate:{:.2f}% testRate:{:.2f}% validation:{:.2f}%  ".
 ↪format(trainRate,testRate, trainRate*validation_split_rate))
X_train, X_test = X_train[train_index], X_train[test_index]
pos_train1, pos_test1 = d1_train_embedded[train_index],␣
 ↪d1_train_embedded[test_index]
pos_train2, pos_test2 = d2_train_embedded[train_index],␣
 ↪d2_train_embedded[test_index]
y_train, y_test = Y_train[train_index], Y_train[test_index]



# # Saving the training data split as a pickle file
# training_data = {
#     'X_train': X_train,
#     'pos_train1': pos_train1,
#     'pos_train2': pos_train2,
#     'y_train': y_train
# }

# with open('training_data.pkl', 'wb') as f:
#     pickle.dump(training_data, f)

# # Saving the testing data split as a pickle file
# testing_data = {
#     'X_test': X_test,
#     'pos_test1': pos_test1,
#     'pos_test2': pos_test2,
#     'y_test': y_test
# }

# with open('testing_data.pkl', 'wb') as f:
#     pickle.dump(testing_data, f)
```

```
TrainRate:80.00% testRate:20.00% validation:8.00%
```

[10]:
```
# Load the training data from the pickle file
with open('training_data.pkl', 'rb') as f:
    train_data = pickle.load(f)

# Load the testing data from the pickle file
with open('testing_data.pkl', 'rb') as f:
    test_data = pickle.load(f)
```

```
# Extract data from the loaded dictionaries
X_train = train_data['X_train']
pos_train1 = train_data['pos_train1']
pos_train2 = train_data['pos_train2']
y_train = train_data['y_train']


X_test = test_data['X_test']
pos_test1 = test_data['pos_test1']
pos_test2 = test_data['pos_test2']
y_test = test_data['y_test']
print(X_train.shape)
print(X_test.shape)
```

```
(4264, 81)
(1066, 81)
```

4.0.4 Run the Evaluation on the test dataset

[11]:
```
MaxEpochs =50
batchsize =32
validation_split_rate = 0.1
history=model.fit([X_train,pos_train1,pos_train2],␣
 ↪y_train,validation_split=validation_split_rate ,epochs=MaxEpochs,␣
 ↪batch_size=batchsize,verbose=1)
```

```
Train on 3837 samples, validate on 427 samples
Epoch 1/50
3837/3837 [==============================] - 190s 50ms/step - loss: 2780.3218 -
acc: 0.4561 - f1: 0.1841 - val_loss: 2779.6724 - val_acc: 0.4075 - val_f1:
0.2241
Epoch 2/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.6294 -
acc: 0.4694 - f1: 0.2441 - val_loss: 2779.6408 - val_acc: 0.4333 - val_f1:
0.3851
Epoch 3/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.6003 -
acc: 0.4780 - f1: 0.2902 - val_loss: 2779.6190 - val_acc: 0.4660 - val_f1:
0.2290
Epoch 4/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.5662 -
acc: 0.5129 - f1: 0.3406 - val_loss: 2779.6011 - val_acc: 0.4848 - val_f1:
0.2184
Epoch 5/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.5442 -
acc: 0.5298 - f1: 0.3990 - val_loss: 2779.6117 - val_acc: 0.4918 - val_f1:
0.3542
Epoch 6/50
```

```
3837/3837 [==============================] - 17s 5ms/step - loss: 2779.5107 -
acc: 0.5549 - f1: 0.4490 - val_loss: 2779.6922 - val_acc: 0.4778 - val_f1:
0.3971
Epoch 7/50
3837/3837 [==============================] - 17s 4ms/step - loss: 2779.4939 -
acc: 0.5736 - f1: 0.4904 - val_loss: 2779.5927 - val_acc: 0.5222 - val_f1:
0.4128
Epoch 8/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2779.4441 -
acc: 0.5932 - f1: 0.5441 - val_loss: 2779.6225 - val_acc: 0.5386 - val_f1:
0.4823
Epoch 9/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.4104 -
acc: 0.6171 - f1: 0.5762 - val_loss: 2779.5837 - val_acc: 0.5129 - val_f1:
0.4833
Epoch 10/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.3748 -
acc: 0.6424 - f1: 0.6102 - val_loss: 2779.5785 - val_acc: 0.5246 - val_f1:
0.4884
Epoch 11/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2779.3336 -
acc: 0.6818 - f1: 0.6545 - val_loss: 2779.5945 - val_acc: 0.5059 - val_f1:
0.4616
Epoch 12/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2779.3164 -
acc: 0.6839 - f1: 0.6679 - val_loss: 2779.4921 - val_acc: 0.5550 - val_f1:
0.4665
Epoch 13/50
3837/3837 [==============================] - 17s 4ms/step - loss: 2779.3153 -
acc: 0.6982 - f1: 0.6843 - val_loss: 2779.6356 - val_acc: 0.5480 - val_f1:
0.5288
Epoch 14/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2779.2680 -
acc: 0.7232 - f1: 0.7066 - val_loss: 2779.7694 - val_acc: 0.5105 - val_f1:
0.4885
Epoch 15/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.2400 -
acc: 0.7407 - f1: 0.7352 - val_loss: 2779.6487 - val_acc: 0.5176 - val_f1:
0.5017
Epoch 16/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.2406 -
acc: 0.7352 - f1: 0.7263 - val_loss: 2779.6183 - val_acc: 0.5340 - val_f1:
0.5043
Epoch 17/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.1886 -
acc: 0.7649 - f1: 0.7612 - val_loss: 2779.7316 - val_acc: 0.5293 - val_f1:
0.5162
Epoch 18/50
```

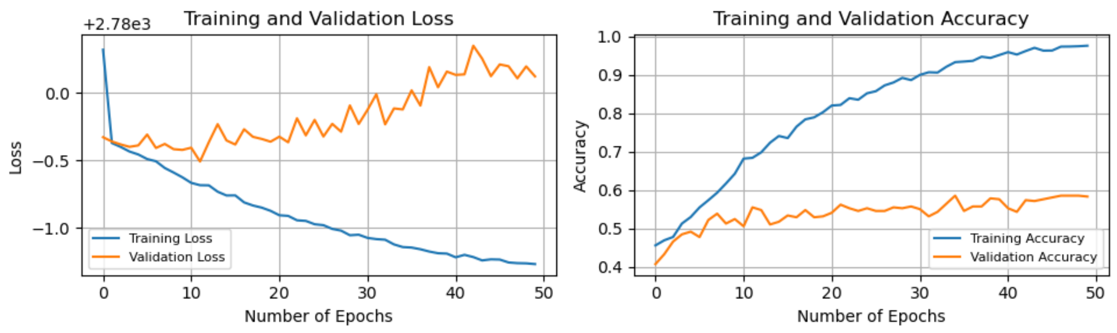
```
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.1657 -
acc: 0.7839 - f1: 0.7783 - val_loss: 2779.6754 - val_acc: 0.5480 - val_f1:
0.5324
Epoch 19/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2779.1498 -
acc: 0.7892 - f1: 0.7853 - val_loss: 2779.6597 - val_acc: 0.5293 - val_f1:
0.5220
Epoch 20/50
3837/3837 [==============================] - 19s 5ms/step - loss: 2779.1260 -
acc: 0.8024 - f1: 0.7961 - val_loss: 2779.6392 - val_acc: 0.5316 - val_f1:
0.5174
Epoch 21/50
3837/3837 [==============================] - 19s 5ms/step - loss: 2779.0931 -
acc: 0.8199 - f1: 0.8163 - val_loss: 2779.6766 - val_acc: 0.5410 - val_f1:
0.5332
Epoch 22/50
3837/3837 [==============================] - 19s 5ms/step - loss: 2779.0885 -
acc: 0.8215 - f1: 0.8189 - val_loss: 2779.6347 - val_acc: 0.5621 - val_f1:
0.5452
Epoch 23/50
3837/3837 [==============================] - 20s 5ms/step - loss: 2779.0558 -
acc: 0.8389 - f1: 0.8386 - val_loss: 2779.8127 - val_acc: 0.5527 - val_f1:
0.5441
Epoch 24/50
3837/3837 [==============================] - 20s 5ms/step - loss: 2779.0513 -
acc: 0.8350 - f1: 0.8325 - val_loss: 2779.6860 - val_acc: 0.5457 - val_f1:
0.5459
Epoch 25/50
3837/3837 [==============================] - 20s 5ms/step - loss: 2779.0265 -
acc: 0.8517 - f1: 0.8496 - val_loss: 2779.8010 - val_acc: 0.5527 - val_f1:
0.5416
Epoch 26/50
3837/3837 [==============================] - 20s 5ms/step - loss: 2779.0195 -
acc: 0.8572 - f1: 0.8569 - val_loss: 2779.6776 - val_acc: 0.5457 - val_f1:
0.5533
Epoch 27/50
3837/3837 [==============================] - 19s 5ms/step - loss: 2778.9918 -
acc: 0.8723 - f1: 0.8693 - val_loss: 2779.7718 - val_acc: 0.5457 - val_f1:
0.5462
Epoch 28/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2778.9790 -
acc: 0.8799 - f1: 0.8758 - val_loss: 2779.7135 - val_acc: 0.5550 - val_f1:
0.5534
Epoch 29/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2778.9448 -
acc: 0.8918 - f1: 0.8901 - val_loss: 2779.9082 - val_acc: 0.5527 - val_f1:
0.5410
Epoch 30/50
```

```
3837/3837 [==============================] - 18s 5ms/step - loss: 2778.9488 -
acc: 0.8861 - f1: 0.8872 - val_loss: 2779.7708 - val_acc: 0.5574 - val_f1:
0.5534
Epoch 31/50
3837/3837 [==============================] - 17s 4ms/step - loss: 2778.9251 -
acc: 0.8997 - f1: 0.8996 - val_loss: 2779.8753 - val_acc: 0.5504 - val_f1:
0.5361
Epoch 32/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.9161 -
acc: 0.9064 - f1: 0.9066 - val_loss: 2779.9919 - val_acc: 0.5316 - val_f1:
0.5213
Epoch 33/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.9116 -
acc: 0.9057 - f1: 0.9041 - val_loss: 2779.7680 - val_acc: 0.5433 - val_f1:
0.5389
Epoch 34/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2778.8768 -
acc: 0.9208 - f1: 0.9211 - val_loss: 2779.8861 - val_acc: 0.5644 - val_f1:
0.5558
Epoch 35/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.8571 -
acc: 0.9328 - f1: 0.9315 - val_loss: 2779.8787 - val_acc: 0.5855 - val_f1:
0.5735
Epoch 36/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.8533 -
acc: 0.9343 - f1: 0.9356 - val_loss: 2780.0203 - val_acc: 0.5457 - val_f1:
0.5377
Epoch 37/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.8409 -
acc: 0.9361 - f1: 0.9369 - val_loss: 2779.9072 - val_acc: 0.5574 - val_f1:
0.5551
Epoch 38/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.8239 -
acc: 0.9468 - f1: 0.9461 - val_loss: 2780.1927 - val_acc: 0.5574 - val_f1:
0.5498
Epoch 39/50
3837/3837 [==============================] - 17s 4ms/step - loss: 2778.8113 -
acc: 0.9437 - f1: 0.9434 - val_loss: 2780.0432 - val_acc: 0.5785 - val_f1:
0.5753
Epoch 40/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.8086 -
acc: 0.9513 - f1: 0.9510 - val_loss: 2780.1598 - val_acc: 0.5761 - val_f1:
0.5664
Epoch 41/50
3837/3837 [==============================] - 17s 4ms/step - loss: 2778.7810 -
acc: 0.9588 - f1: 0.9597 - val_loss: 2780.1352 - val_acc: 0.5527 - val_f1:
0.5479
Epoch 42/50
```

```
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.7994 -
acc: 0.9526 - f1: 0.9527 - val_loss: 2780.1391 - val_acc: 0.5433 - val_f1:
0.5245
Epoch 43/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2778.7829 -
acc: 0.9617 - f1: 0.9606 - val_loss: 2780.3517 - val_acc: 0.5738 - val_f1:
0.5672
Epoch 44/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.7580 -
acc: 0.9700 - f1: 0.9691 - val_loss: 2780.2563 - val_acc: 0.5714 - val_f1:
0.5603
Epoch 45/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.7659 -
acc: 0.9627 - f1: 0.9628 - val_loss: 2780.1257 - val_acc: 0.5761 - val_f1:
0.5695
Epoch 46/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2778.7643 -
acc: 0.9630 - f1: 0.9636 - val_loss: 2780.2128 - val_acc: 0.5808 - val_f1:
0.5688
Epoch 47/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.7427 -
acc: 0.9732 - f1: 0.9724 - val_loss: 2780.1994 - val_acc: 0.5855 - val_f1:
0.5759
Epoch 48/50

3837/3837 [==============================] - 17s 5ms/step - loss: 2778.7378 -
acc: 0.9734 - f1: 0.9738 - val_loss: 2780.1098 - val_acc: 0.5855 - val_f1:
0.5759
Epoch 49/50
3837/3837 [==============================] - 17s 5ms/step - loss: 2778.7363 -
acc: 0.9742 - f1: 0.9745 - val_loss: 2780.1983 - val_acc: 0.5855 - val_f1:
0.5806
Epoch 50/50
3837/3837 [==============================] - 18s 5ms/step - loss: 2778.7308 -
acc: 0.9755 - f1: 0.9759 - val_loss: 2780.1241 - val_acc: 0.5831 - val_f1:
0.5781
```

[12]:
```python
import matplotlib.pyplot as plt

# Training & Validation accuracy
train_loss = history.history['loss']
val_loss = history.history['val_loss']
train_acc = history.history['acc']
val_acc = history.history['val_acc']
epochs = len(train_loss)

xc = range(epochs)
```

```
plt.figure(figsize=(10, 3))

# Loss subplot
plt.subplot(1, 2, 1)
plt.plot(xc, train_loss, label='Training Loss')
plt.plot(xc, val_loss, label='Validation Loss')
plt.xlabel('Number of Epochs', fontsize=10)
plt.ylabel('Loss', fontsize=10)
plt.title('Training and Validation Loss', fontsize=12)
plt.legend(fontsize=8)
plt.grid(True)

# Accuracy subplot
plt.subplot(1, 2, 2)
plt.plot(xc, train_acc, label='Training Accuracy')
plt.plot(xc, val_acc, label='Validation Accuracy')
plt.xlabel('Number of Epochs', fontsize=10)
plt.ylabel('Accuracy', fontsize=10)
plt.title('Training and Validation Accuracy', fontsize=12)
plt.legend(fontsize=8, loc='lower right')  # Change position to lower right
plt.grid(True)

plt.tight_layout()
plt.show()
```



[13]:
```
import torch
from sklearn.metrics import accuracy_score, classification_report,␣
↪precision_recall_fscore_support

# Predict on the training dataset
train_predicted = np.argmax(model.predict([X_train, pos_train1, pos_train2]),␣
↪axis=1)
y_train_to_label = np.argmax(y_train, axis=1)
```

```
# Calculate accuracy, precision, recall, and F1-score for the training data
train_accuracy = accuracy_score(y_train_to_label, train_predicted)
train_prec, train_reca, train_fscore, _ =␣
 ↪precision_recall_fscore_support(y_train_to_label, train_predicted,␣
 ↪average=param)

# Print the classification report for the training data
print("Training Classification Report:")
print(classification_report(y_train_to_label, train_predicted))

# Print the precision, recall, and F1-score for the training data
print("Training Accuracy: {:.2f}%".format(train_accuracy * 100))
print("Training Precision: {:.2f}%".format(train_prec * 100))
print("Training Recall: {:.2f}%".format(train_reca * 100))
print("Training F1 Score: {:.2f}%".format(train_fscore * 100))
```

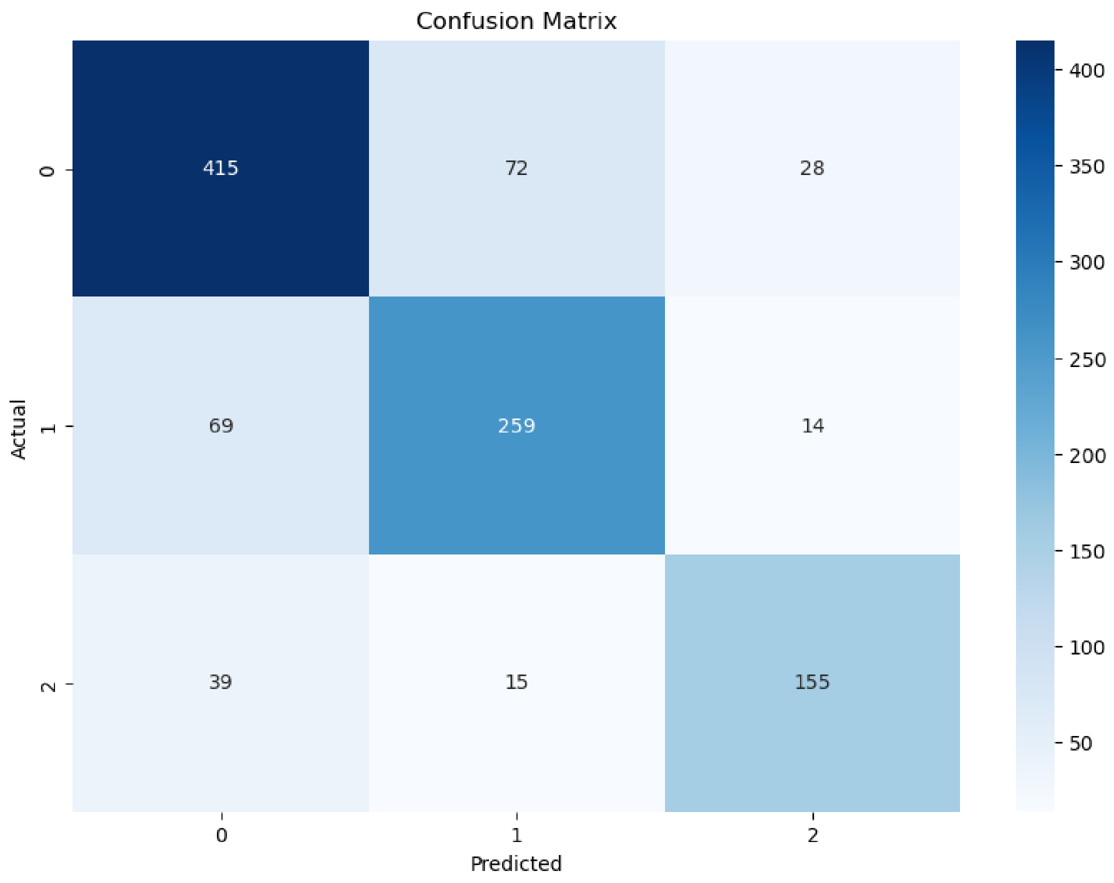
```
Training Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.95      0.95      2023
           1       0.93      0.95      0.94      1468
           2       0.94      0.91      0.92       773

    accuracy                           0.94      4264
   macro avg       0.94      0.94      0.94      4264
weighted avg       0.94      0.94      0.94      4264

Training Accuracy: 94.23%
Training Precision: 93.98%
Training Recall: 93.68%
Training F1 Score: 93.82%
```

[19]:
```
predicted = np.argmax(model.predict([X_test,pos_test1,pos_test2]), axis=1)
y_test_to_label = np.argmax(y_test, axis=1)
prec, reca, fscore, sup = precision_recall_fscore_support(y_test_to_label,␣
 ↪predicted, average=param)
# Generate the classification report as a dictionary
report_dict = classification_report(y_test_to_label, predicted,␣
 ↪output_dict=True)

# Create a new dictionary to hold the formatted values
formatted_report_dict = {}

# Iterate over the items in the report dictionary
for key, value in report_dict.items():
    if isinstance(value, dict):
        # Format the nested dictionary values
        formatted_report_dict[key] = {sub_key: f"{sub_value:.4f}" for sub_key,␣
↪sub_value in value.items()}
    else:
        # Format the top-level dictionary values
        formatted_report_dict[key] = f"{value:.4f}"

# Create a string representation of the formatted dictionary
formatted_report_str = classification_report(y_test_to_label, predicted,␣
↪digits=4)

# Print the formatted classification report
print(formatted_report_str)
print(" Precision:{:.2f}% Recall:{:.2f}% Fscore:{:.2f}% ".format(prec*100,␣
↪reca*100, fscore*100))
```

```
              precision    recall  f1-score   support

           0     0.7935    0.8058    0.7996       515
           1     0.7486    0.7573    0.7529       342
           2     0.7868    0.7416    0.7635       209

    accuracy                         0.7777      1066
   macro avg     0.7763    0.7683    0.7720      1066
weighted avg     0.7778    0.7777    0.7776      1066

 Precision:77.63% Recall:76.83% Fscore:77.20%
```

[20]:
```
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import classification_report, confusion_matrix,␣
↪precision_recall_fscore_support
from sklearn.model_selection import StratifiedKFold
# Calculate and visualize the confusion matrix
cm = confusion_matrix(y_test_to_label, predicted)
plt.figure(figsize=(10, 7))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=['0', '1','2'],␣
↪yticklabels=['0', '1','2'])
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix')
plt.show()

# Print precision, recall, and f-score
prec, reca, fscore, sup = precision_recall_fscore_support(y_test_to_label,␣
↪predicted, average=param)
```

```
print(" Precision:{:.2f}% Recall:{:.2f}% Fscore:{:.2f}% ".format(prec*100,␣
↪reca*100, fscore*100))
```



```
Precision:77.63% Recall:76.83% Fscore:77.20%
```

[ ]: