

Real-World Applications of Machine Learning

2.2: Complex Machine Learning Models and Keras Part 1

Keras Convolutional Neural Network (CNN) model

Starting hyperparameters used:

epochs = 10

Batch_size = 16

n_hidden = 32

kernal_size = 2

Ending hyperparameters used:

5. Build and run CNN keras model

```
In [38]: 1 # Adjust model hyperparameters
          2 epochs = 50
          3 batch_size = 256
          4 n_hidden = 256
          5 kernel_size = 3 # Adjust the kernel size here
          6
```

Final accuracy and loss:

```
Epoch 1/50
72/72 ————— 3s 29ms/step - accuracy: 0.0878 - loss: 32.8681
Epoch 2/50
72/72 ————— 2s 28ms/step - accuracy: 0.0794 - loss: 444.0713
Epoch 3/50
72/72 ————— 2s 28ms/step - accuracy: 0.0903 - loss: 6295.6978
Epoch 4/50
72/72 ————— 2s 28ms/step - accuracy: 0.0842 - loss: 14114.4316
Epoch 5/50
72/72 ————— 2s 28ms/step - accuracy: 0.0979 - loss: 30875.8047
Epoch 6/50
72/72 ————— 2s 27ms/step - accuracy: 0.0923 - loss: 33370.5938
Epoch 7/50
72/72 ————— 2s 27ms/step - accuracy: 0.0874 - loss: 75583.1328
Epoch 8/50
72/72 ————— 2s 27ms/step - accuracy: 0.1267 - loss: 71762.4531
Epoch 9/50
72/72 ————— 2s 29ms/step - accuracy: 0.1130 - loss: 114932.4531
Epoch 10/50
```

Keanu G.

2.2-complex-machine-learning-models-pt-1

```
Epoch 42/50
72/72 ————— 2s 29ms/step - accuracy: 0.1228 - loss: 926242.8125
Epoch 43/50
72/72 ————— 2s 29ms/step - accuracy: 0.1405 - loss: 893793.7500
Epoch 44/50
72/72 ————— 2s 29ms/step - accuracy: 0.1262 - loss: 1038898.1875
Epoch 45/50
72/72 ————— 2s 28ms/step - accuracy: 0.1113 - loss: 893272.0000
Epoch 46/50
72/72 ————— 2s 28ms/step - accuracy: 0.1299 - loss: 1342451.2500
Epoch 47/50
72/72 ————— 2s 29ms/step - accuracy: 0.1394 - loss: 1102487.8750
Epoch 48/50
72/72 ————— 2s 30ms/step - accuracy: 0.1409 - loss: 995745.4375
Epoch 49/50
72/72 ————— 2s 29ms/step - accuracy: 0.1118 - loss: 1312391.8750
Epoch 50/50
72/72 ————— 2s 29ms/step - accuracy: 0.1104 - loss: 1273658.3750

Out[38]: <keras.src.callbacks.history.History at 0x1b235f70290>
```

Partial screenshot of (ending parameters) confusion matrix (easier to see in script):

```
144/144 ————— 0s 2ms/step
Pred                                BELGRADE_pleasant_weather \
True
BASEL_pleasant_weather              265
BELGRADE_pleasant_weather            73
BUDAPEST_pleasant_weather            8
DEBILT_pleasant_weather              0
DUSSELDORF_pleasant_weather          0
HEATHROW_pleasant_weather            1
KASSEL_pleasant_weather              0
LJUBLJANA_pleasant_weather           4
MAASTRICHT_pleasant_weather          0
MADRID_pleasant_weather              5
MUNCHENB_pleasant_weather            0
OSLO_pleasant_weather               0
STOCKHOLM_pleasant_weather           1
VALENTIA_pleasant_weather            0

Pred                                DUSSELDORF_pleasant_weather \
True
BASEL_pleasant_weather              2
BELGRADE_pleasant_weather            0
BUDAPEST_pleasant_weather            0
DEBILT_pleasant_weather              0
DUSSELDORF_pleasant_weather          0
HEATHROW_pleasant_weather            0
KASSEL_pleasant_weather              0
LJUBLJANA_pleasant_weather           0
MAASTRICHT_pleasant_weather          0
MADRID_pleasant_weather              0
MUNCHENB_pleasant_weather            0
OSLO pleasant weather                0
```

Keanu G.

2.2-complex-machine-learning-models-pt-1

Notes:

- The model predicted for all weather stations.
- There is relatively low accuracy and exceptionally high loss for all hyperparameters used in CNN model.
- Accuracy improves slightly then plateaus or decreases while the loss increases with each epoch.
- The model may be too simple to capture the complex data patterns.
- When using `binary_crossentropy` with sigmoid, similar results occur, however, the accuracy increases while loss converges lower with each epoch to around .08 loss and 0.3 accuracy.

Questions to consider:

- Does the model use an unscaled data approach? (yes)
- Possible preprocessing issues?
- Possible data architectural issues?

Recommendations:

- Implement scaled weather data.
- A better look into data preprocessing and architectural issues.
- Consider more complex layering, a different activation method, or keras model.