COMPONENT THREE

Convolutional Neural Network

Abstract

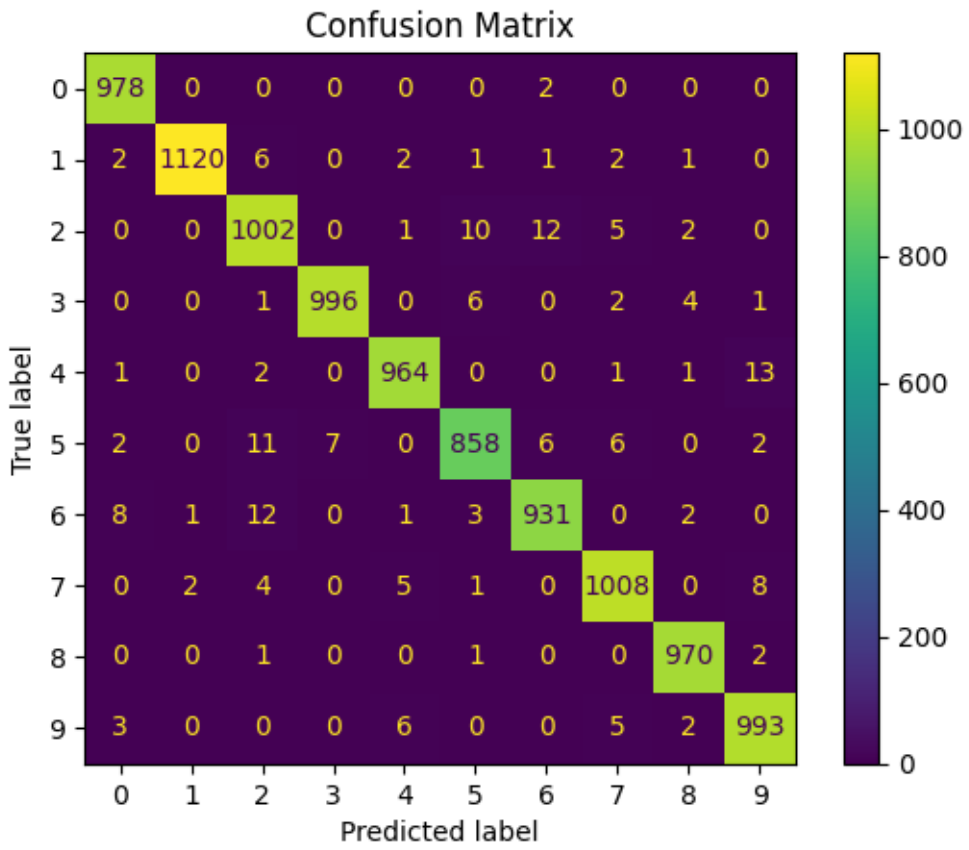
The aim of this report is to correctly classify handwritten-based user authentication. The task here is to correctly classify the handwritten digits in the MNIST dataset using Convolutional Neural Network (CNN).

Methodology

We used different regularization methods to affect the performance of the CNN model, examined how changes to the number of convolution blocks affect the performance of the model, checked the effect of varying learning rates on the performance of the CNN algorithm and also observed if there was overfitting in the model.

Question A

REGULARIZER	OPTIMIZER	LEARNING RATE	VALIDATION ACCURACY	CNN
Dropout 0.5 MaxPooling2D(pool_size=(2,2) Dense Layer Batch size (28) Epoch 10	SGD	0.001	0.98	3 CONVOLUTIONAL BLOCK
MaxPooling2D(pool_size=(2,2) Dense Layer Batch size(28) Epoch 10	SGD	0.001	0.96	2 CONVOLUTIONAL BLOCK
MaxPooling2D(pool_size=(2,2) Dense Layer Batch size(28) Epoch 10	SGD	0.0001	0.93	2 CONVOLUTIONAL BLOCK
Dropout 0.7 MaxPooling2D(pool_size=(2,2) Dense Layer Batch size(128) Epoch 5 Early Stopping	ADAM	0.001	0.98	3 CONVOLUTIONAL BLOCK



	precision	recall	f1-score	support
0	0.98	1.00	0.99	980
1	1.00	0.99	0.99	1135
2	0.96	0.97	0.97	1032
3	0.99	0.99	0.99	1010
4	0.98	0.98	0.98	982
5	0.97	0.96	0.97	892
6	0.98	0.97	0.97	958
7	0.98	0.98	0.98	1028
8	0.99	1.00	0.99	974
9	0.97	0.98	0.98	1009
accuracy			0.98	10000
macro avg	0.98	0.98	0.98	10000
weighted avg	0.98	0.98	0.98	10000

Regularization methods is used to reduce or prevent overfitting in training models, I trained multiple models with different regularization techniques and compared their performance. From these results, all the regularization methods improved the validation accuracy. Overall the result suggest that when we increased the drop out to 0.7, increased batch size to 128, changed optimizer to ADAM, reduced the number of epochs to 5, used early stopping , it still gave a good accuracy of 98% which shows that it is an effective regularization method for improving the performance of CNN models on the MNIST dataset.

Question B

To investigate how changes to the number of convolutional blocks affected the performance of the CNN model, I trained the models with different numbers of convolutional blocks and compared their validation accuracies, reduced the convolutional block to two, removed dropout and increased the batch size, there was no obvious changes as the model still trained well with an accuracy of 96%.

Question C

To check the effect of the learning rate on the performance of the CNN algorithm, I trained the CNN model with another learning rate of 0.0001, used SGD optimizer with the learning rate and trained the model to 10 epochs with a batch size of 24 then I observed it gave a less accuracy of 93% compared to other learning rates which means that the learning rate can have a significant impact on the performance of the CNN depending on how high or low it is.

Question D

Based on the hyper-parameter tuning performed on the model, there was no case of overfitting as the model performed well. To check for overfitting, I plotted the training accuracy against the validation accuracy. They both improved and converged to similarly high values, which suggests that there was no significant overfitting as the models generalized well.