Report

Domestic vs Wild Cat Classification

Approach

- Load dataset
- Preprocess and split the dataset
- Build and tune a feed forward fully connected neural network model
- feed forward :- Sequential model and fully connected :- dense layers
- choose the right optimizer
- Analyze the loss and accuracy for both the training and validation datasets
- Test the model built on test dataset
- Perform calculations
- Change the optimizer and other parameters(learning rate, momentum, batch size) and perform the above steps again.

Preprocessing

The dataset is split into train and test in the ratio of 80:20

Further train data is also split into train and validation in the ratio of 80:20

Model

Feed forward model is built using keras.models.Sequential( ) and then made fully connected using keras.layers.Dense

Activation Function used: 'softmax' in output layer and 'relu' in all other layers

Dropout rate : 0.9

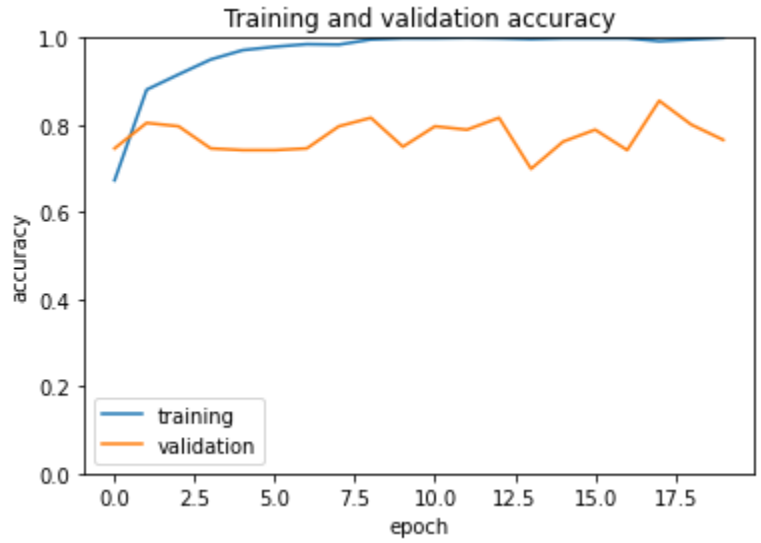
Optimizers Used

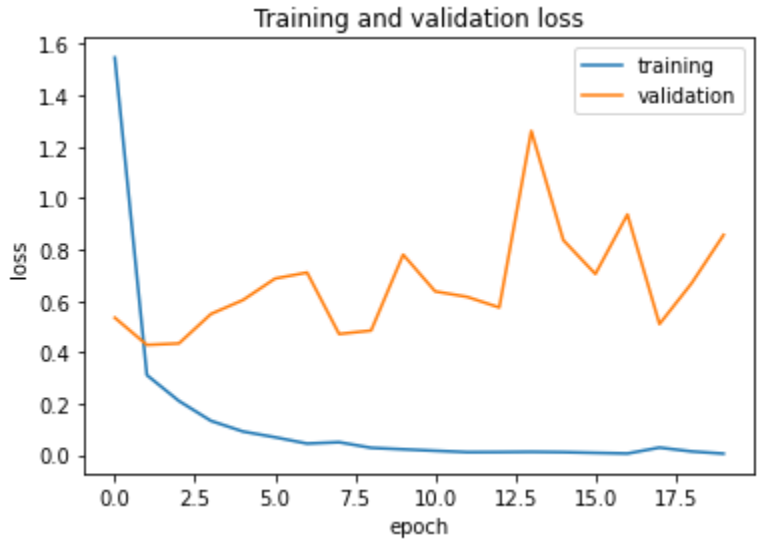
- Adam
- SGD with batch_size = 256
- SGD(learning rate = 0.01, momentum = 0.9) with batch_size = 16
- SGD(learning rate = 0.01, momentum = 0.5) with batch_size = 16
- SGD(learning rate = 0.01) with batch_size = 16
- RMSprop

- Adagrad
- Adadelta

Plots and Evaluation Metrics

- Adam





Overall Categorical Accuracy: 88.75%

Confusion Matrix

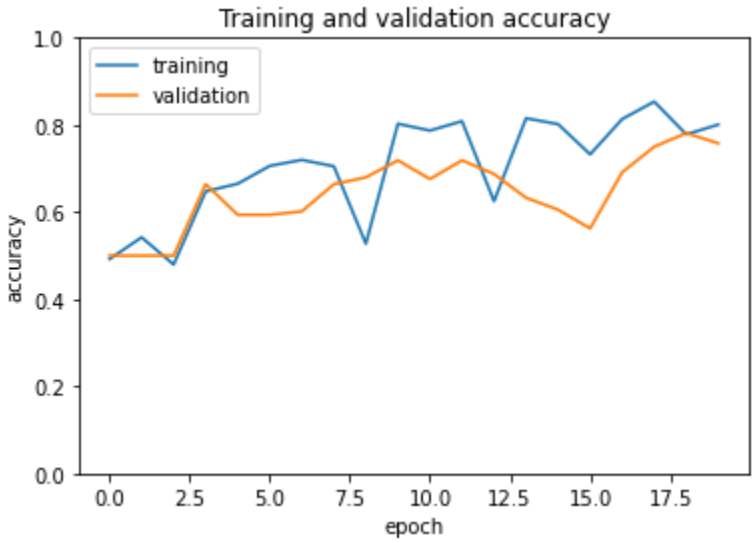
[[138 22]

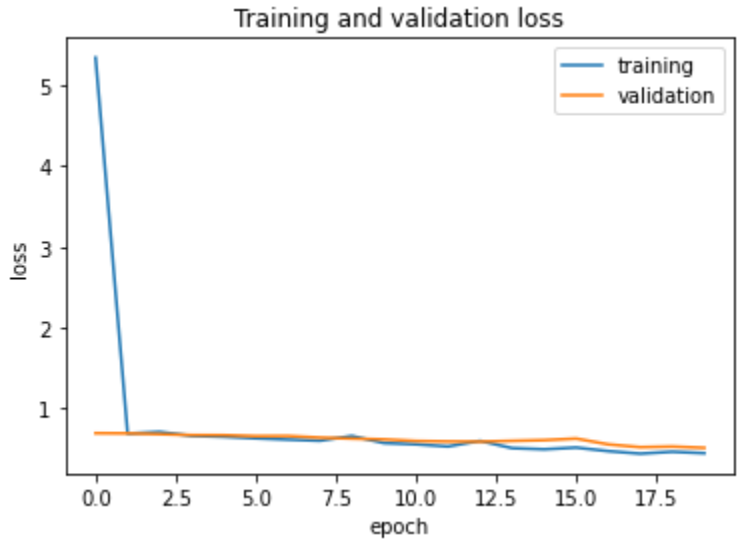
[14 146]]

AUC: 0.888

	Precision	Recall	F-Score	Support	Specificity
big_cats	0.869048	0.9125	0.890244	160.0	0.862500
cats	0.907895	0.8625	0.884615	160.0	0.869048

- SGD (lr = 0.01) with batch_size = 256





Overall Categorical Accuracy: 83.75%

Confusion Matrix

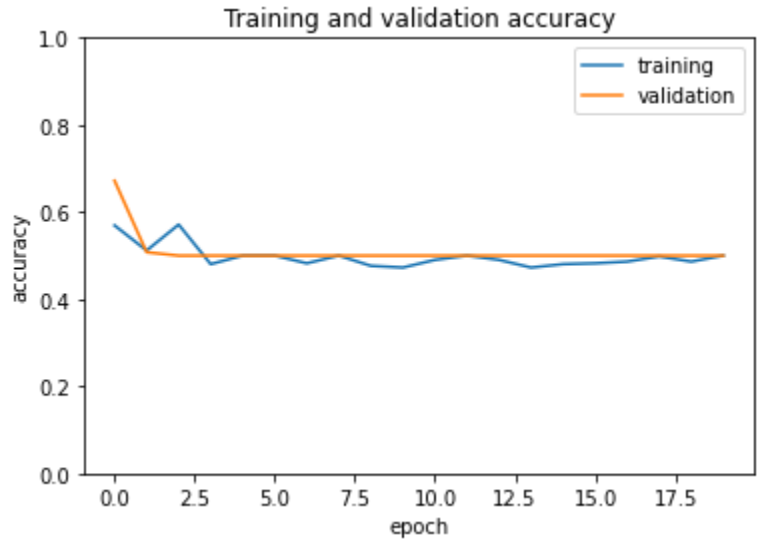
[[118 42]

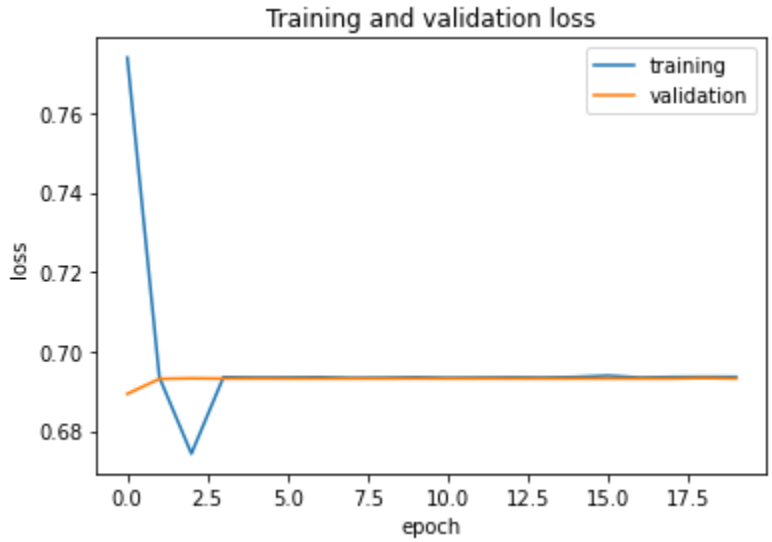
[10 150]]

AUC: 0.838

	Precision	Recall	F-Score	Support	Specificity
big_cats	0.781250	0.9375	0.852273	160.0	0.73750
cats	0.921875	0.7375	0.819444	160.0	0.78125

- SGD(lr = 0.01, momentum = 0.9) with batch_size = 16





Overall Categorical Accuracy: 60.94%

Confusion Matrix

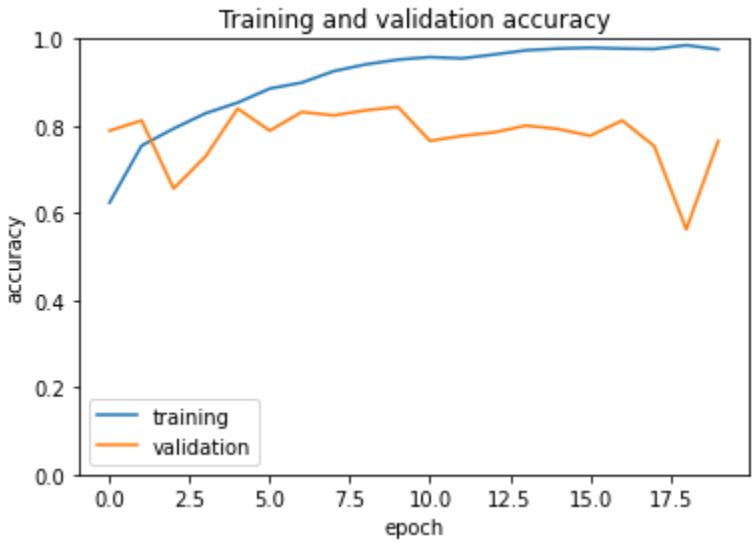
[[125 35]

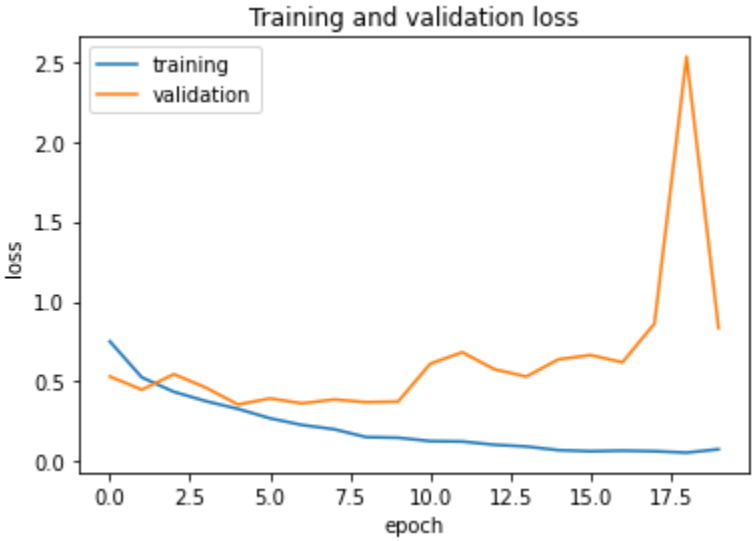
[90 70]]

AUC: 0.609

	Precision	Recall	F-Score	Support	Specificity
cats	0.581395	0.78125	0.666667	160.0	0.666667
big_cats	0.666667	0.43750	0.528302	160.0	0.781250

- SGD(lr=0.01, momentum = 0.5) with batch_size = 16





Overall Categorical Accuracy: 84.06%

Confusion Matrix

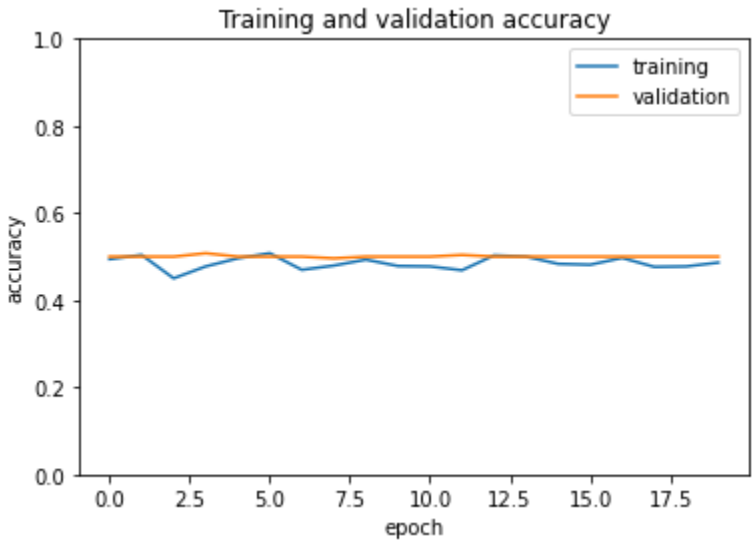
[[128 32]

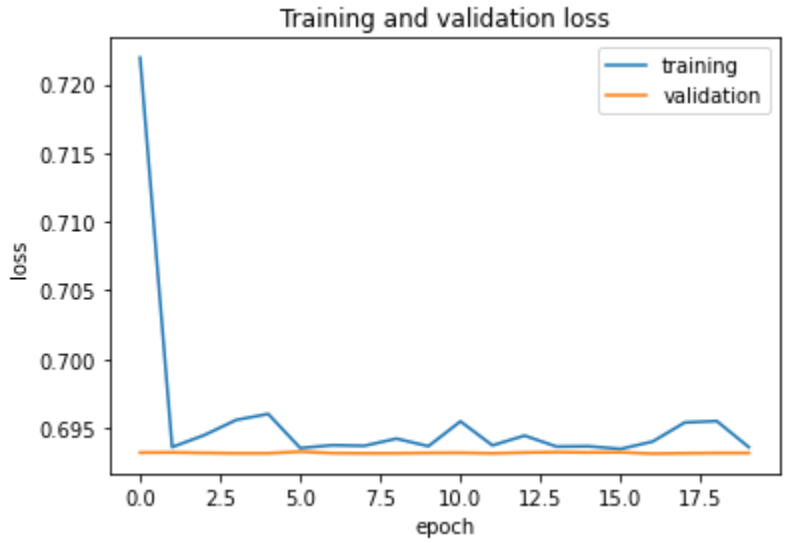
[19 141]]

AUC: 0.841

	Precision	Recall	F-Score	Support	Specificity
big_cats	0.815029	0.88125	0.846847	160.0	0.800000
cats	0.870748	0.80000	0.833876	160.0	0.815029

- SGD(lr = 0.01) with batch_size = 1





Overall Categorical Accuracy: 50.00%

Confusion Matrix

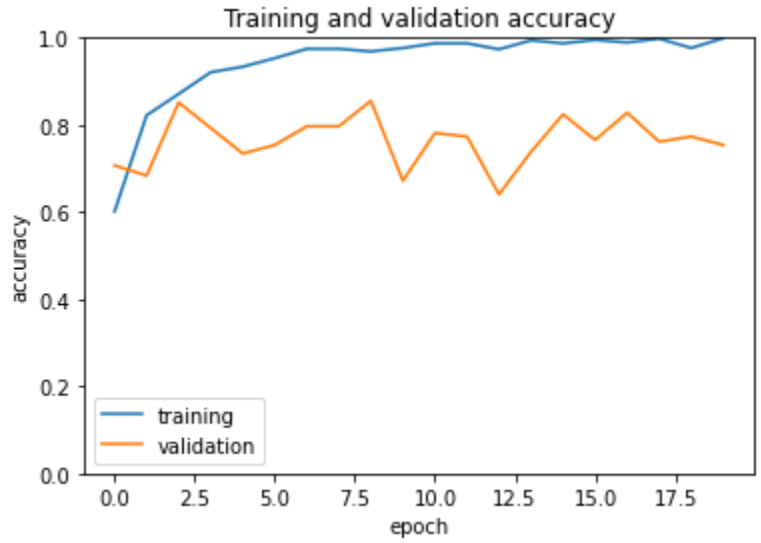
[[160 0]

[160 0]]

AUC: 0.500

	Precision	Recall	F-Score	Support	Specificity
cats	0.5	1.0	0.666667	160.0	NaN
big_cats	0.0	0.0	0.000000	160.0	1.0

- RMSprop



Overall Categorical Accuracy: 88.44%

Confusion Matrix

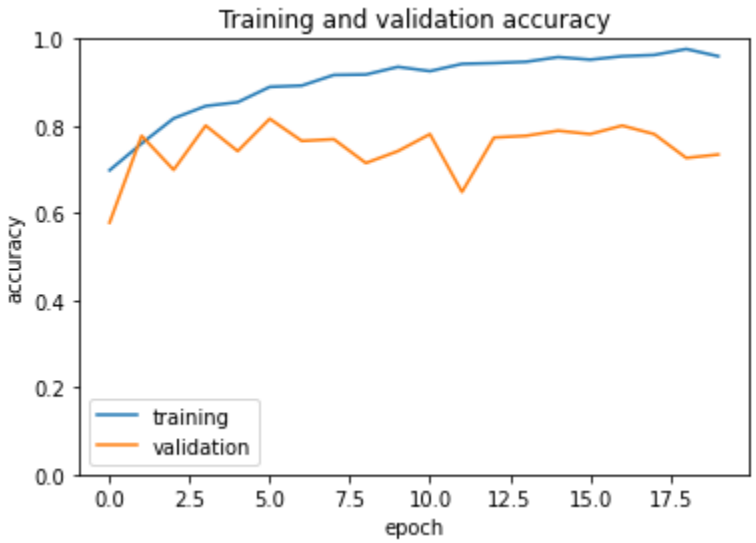
[[133 27]

[10 150]]

AUC: 0.884

	Precision	Recall	F-Score	Support	Specificity
big_cats	0.847458	0.93750	0.890208	160.0	0.831250
cats	0.930070	0.83125	0.877888	160.0	0.847458

- Adagrad



Overall Categorical Accuracy: 90.00%

Confusion Matrix

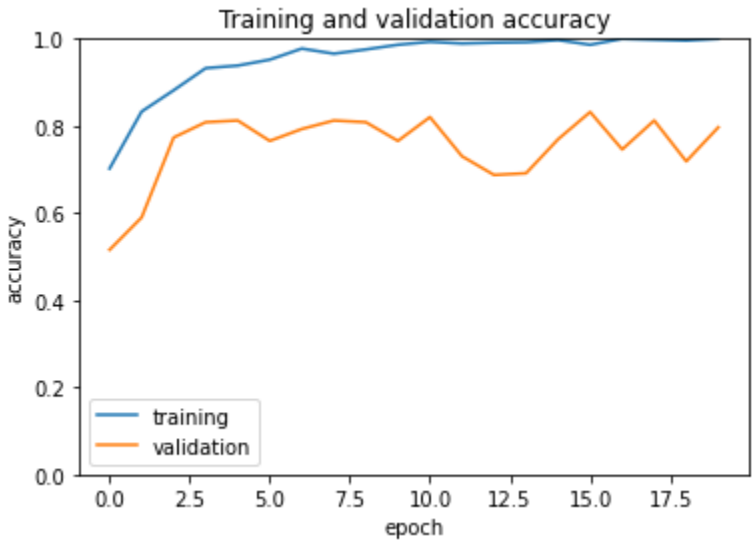
[[143 17]

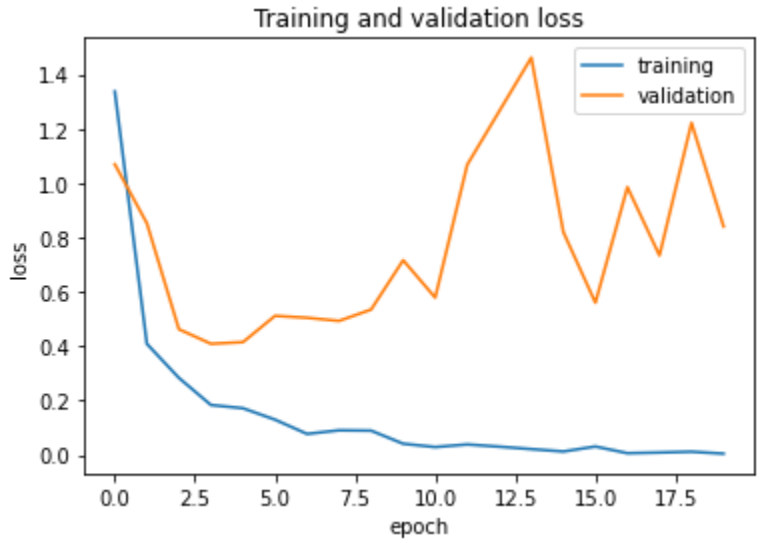
[15 145]]

AUC: 0.900

	Precision	Recall	F-Score	Support	Specificity
big_cats	0.895062	0.90625	0.900621	160.0	0.893750
cats	0.905063	0.89375	0.899371	160.0	0.895062

- Adadelta





Overall Categorical Accuracy: 90.94%

Confusion Matrix

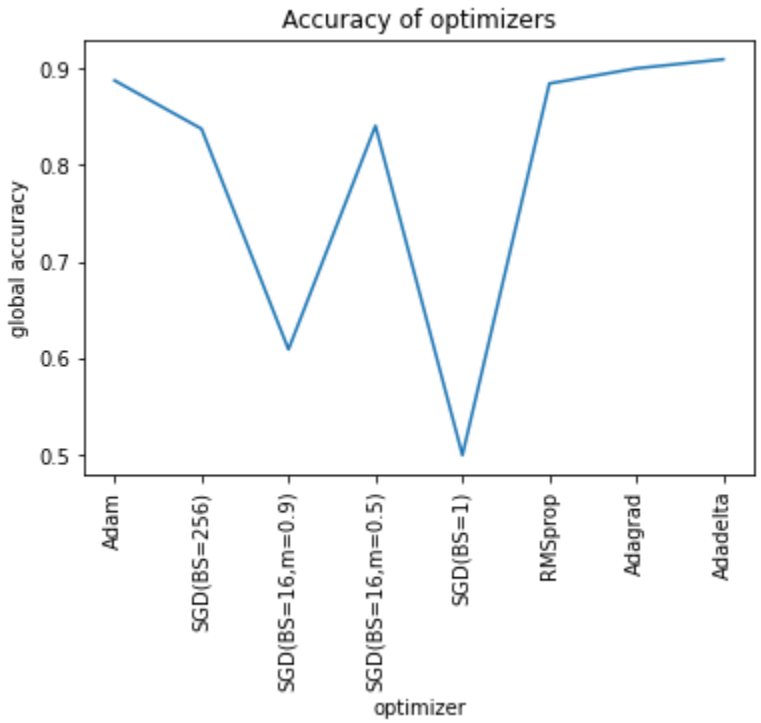
[[144 16]

[13 147]]

AUC: 0.909

	Precision	Recall	F-Score	Support	Specificity
big_cats	0.901840	0.91875	0.910217	160.0	0.90000
cats	0.917197	0.90000	0.908517	160.0	0.90184

Accuracy Results



Observations

- As observed with SGD optimizer, increase in momentum decreases the accuracy.
- As observed, SGD with batch size = 1 resulted in the least accuracy among all optimizers used and Adadelta showed maximum accuracy.