Configuration 1

Parameters: Em_sz: 50 Latent_sz: 200 Dropout_rate: 0.3 Filters: 64 Batch_Size: 20 Learning_rate: 0.001

```
Model: "sequential"
```

| Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 42, 50)	468,100
dropout (Dropout)	(None, 42, 50)	0
conv1d (Conv1D)	(None, 38, 64)	16,064
max_pooling1d (MaxPooling1D)	(None, 9, 64)	0
lstm (LSTM)	(None, 200)	212,000
flatten (Flatten)	(None, 200)	0
dense (Dense)	(None, 3)	603

Training and validation acc
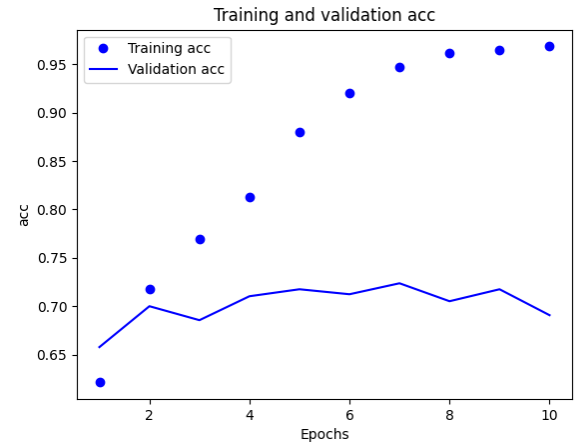
```
Total params: 696,767 (2.66 MB)
Trainable params: 696,767 (2.66 MB)
Non-trainable params: 0 (0.00 B)
Epoch 1/10
194/194 ──────────────── 7s 22ms/step – accuracy: 0.5868 – loss: 0.9373 – val_accuracy: 0.6577 – val_loss: 0.7661
Epoch 2/10
194/194 ──────────────── 3s 18ms/step – accuracy: 0.7055 – loss: 0.6998 – val_accuracy: 0.7000 – val_loss: 0.7176
Epoch 3/10
194/194 ──────────────── 4s 18ms/step – accuracy: 0.7776 – loss: 0.5304 – val_accuracy: 0.6856 – val_loss: 0.7555
Epoch 4/10
194/194 ──────────────── 4s 19ms/step – accuracy: 0.8167 – loss: 0.4335 – val_accuracy: 0.7103 – val_loss: 0.7978
Epoch 5/10
194/194 ──────────────── 4s 22ms/step – accuracy: 0.8784 – loss: 0.3032 – val_accuracy: 0.7175 – val_loss: 0.8478
Epoch 6/10
194/194 ──────────────── 3s 17ms/step – accuracy: 0.9305 – loss: 0.1886 – val_accuracy: 0.7124 – val_loss: 1.0045
Epoch 7/10
194/194 ──────────────── 3s 17ms/step – accuracy: 0.9548 – loss: 0.1384 – val_accuracy: 0.7237 – val_loss: 1.1058
Epoch 8/10
194/194 ──────────────── 4s 18ms/step – accuracy: 0.9684 – loss: 0.1040 – val_accuracy: 0.7052 – val_loss: 1.2102
Epoch 9/10
194/194 ──────────────── 3s 16ms/step – accuracy: 0.9696 – loss: 0.1014 – val_accuracy: 0.7175 – val_loss: 1.4375
Epoch 10/10
194/194 ──────────────── 3s 16ms/step – accuracy: 0.9669 – loss: 0.0902 – val_accuracy: 0.6907 – val_loss: 1.3109
31/31 ──────────────── 0s 6ms/step – accuracy: 0.6856 – loss: 1.3312
1.3108787536621094
0.6907216310501099
31/31 ──────────────── 1s 12ms/step
p =  [[3.9010658e-05 9.9933499e-01 6.2597112e-04]
 [1.5836004e-05 9.9863333e-01 1.3508486e-03]
 [2.2608885e-03 1.7927801e-03 9.9594635e-01]
 ...
 [1.0360047e-01 7.5830537e-01 1.3809419e-01]
 [1.3500453e-02 3.8128307e-03 9.8268676e-01]
 [1.0477717e-03 4.6584815e-01 5.3310406e-01]]
1/1 ──────────────── 0s 34ms/step
pred  2
y_test[7:8]  [[False  True False]]
prediction   positive
```

This configuration has around 700,000 parameters. The trained model achieves 70% accuracy on a test set, which is good enough performance. However, the validation loss for this configuration is high. This indicates that while the model fits the training data well, it may not generalize as effectively to unseen data. In a specific test case, the model predicted the status as positive, though this prediction doesn't necessarily reflect its overall generalization ability. The higher validation loss suggests there may be overfitting, where the model memorizes the training data but struggles to adapt to new data.

Configuration 2

Parameters: Em_sz: 30 Latent_sz: 100 Dropout_rate: 0.3 Filters: 32 Batch_Size: 30 Learning_rate: 0.01

```
Model: "sequential"
_____
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 42, 30)            280,860

 dropout (Dropout)           (None, 42, 30)            0

 conv1d (Conv1D)             (None, 38, 32)            4,832

 max_pooling1d (MaxPooling1D) (None, 9, 32)            0

 lstm (LSTM)                 (None, 100)               53,200

 flatten (Flatten)           (None, 100)               0

 dense (Dense)               (None, 3)                 303

=================================================================
Total params: 339,195 (1.29 MB)
Trainable params: 339,195 (1.29 MB)
Non-trainable params: 0 (0.00 B)
```


Training and validation acc
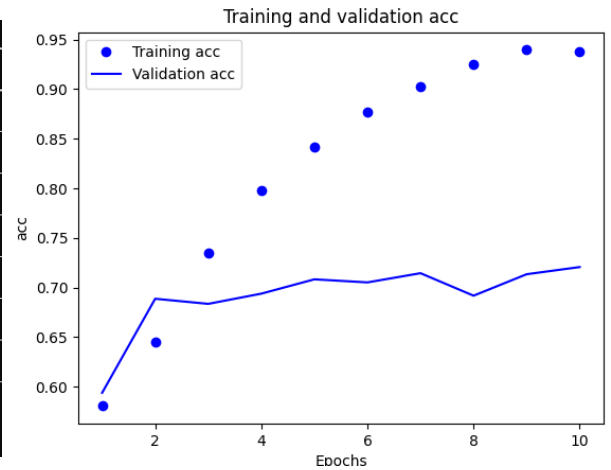
```
Epoch 1/10
130/130 ──────────────── 4s 14ms/step - accuracy: 0.5508 - loss: 0.9580 - val_accuracy: 0.5938 - val_loss: 0.8477
Epoch 2/10
130/130 ──────────────── 1s 10ms/step - accuracy: 0.6251 - loss: 0.8050 - val_accuracy: 0.6887 - val_loss: 0.7490
Epoch 3/10
130/130 ──────────────── 1s 10ms/step - accuracy: 0.7409 - loss: 0.6063 - val_accuracy: 0.6835 - val_loss: 0.7864
Epoch 4/10
130/130 ──────────────── 1s 10ms/step - accuracy: 0.8136 - loss: 0.4693 - val_accuracy: 0.6938 - val_loss: 0.8242
Epoch 5/10
130/130 ──────────────── 1s 10ms/step - accuracy: 0.8483 - loss: 0.3938 - val_accuracy: 0.7082 - val_loss: 0.8492
Epoch 6/10
130/130 ──────────────── 1s 10ms/step - accuracy: 0.8759 - loss: 0.3275 - val_accuracy: 0.7052 - val_loss: 0.9288
Epoch 7/10
130/130 ──────────────── 1s 10ms/step - accuracy: 0.9125 - loss: 0.2504 - val_accuracy: 0.7144 - val_loss: 1.0332
Epoch 8/10
130/130 ──────────────── 1s 10ms/step - accuracy: 0.9192 - loss: 0.2284 - val_accuracy: 0.6918 - val_loss: 1.0844
Epoch 9/10
130/130 ──────────────── 1s 10ms/step - accuracy: 0.9343 - loss: 0.1929 - val_accuracy: 0.7134 - val_loss: 1.0897
Epoch 10/10
130/130 ──────────────── 1s 11ms/step - accuracy: 0.9395 - loss: 0.1723 - val_accuracy: 0.7206 - val_loss: 1.0333
31/31 ──────────────── 0s 4ms/step - accuracy: 0.7179 - loss: 1.0650
1.0332735776901245
0.7206185460090637
31/31 ──────────────── 1s 11ms/step
p =  [[2.1900518e-05 9.9799711e-01 1.9809257e-03]
 [1.0039723e-05 9.9882919e-01 1.1607175e-03]
 [9.7659032e-04 7.6068356e-03 9.9141663e-01]

 ...

 [5.8772750e-02 8.3579302e-01 1.0543420e-01]
 [2.8098035e-01 2.8292355e-01 4.3609613e-01]
 [2.1711414e-05 9.9806434e-01 1.9139099e-03]]
1/1 ──────────────── 0s 32ms/step
pred  0
y_test[7:8]  [[False  True False]]
prediction  negative
```

This configuration has slightly better accuracy on the test set, with a 72% accuracy compared to configuration 1. The validation loss is nearly identical to configuration 1 which suggests that the model performs almost similar on the unseen data. There are around 400,000 trainable parameters in this configuration, quite less in comparison to the first configuration. Adding few epochs could potentially achieve the same validation loss as configuration 1.

Configuration 3

Parameters: Em_sz: 70 Latent_sz: 300 Dropout_rate: 0.5 Filters: 128 Batch_Size: 50 Learning_rate: 0.0005

```
Model: "sequential"
┌─────────────────────────────┬──────────────────────┬──────────────┐
│ Layer (type)                │ Output Shape         │      Param # │
├─────────────────────────────┼──────────────────────┼──────────────┤
│ embedding (Embedding)       │ (None, 42, 70)       │      655,340 │
│ dropout (Dropout)           │ (None, 42, 70)       │            0 │
│ conv1d (Conv1D)             │ (None, 38, 128)      │       44,928 │
│ max_pooling1d (MaxPooling1D)│ (None, 9, 128)       │            0 │
│ lstm (LSTM)                 │ (None, 300)          │      514,800 │
│ flatten (Flatten)           │ (None, 300)          │            0 │
│ dense (Dense)               │ (None, 3)            │          903 │
└─────────────────────────────┴──────────────────────┴──────────────┘
Total params: 1,215,971 (4.64 MB)
Trainable params: 1,215,971 (4.64 MB)
Non-trainable params: 0 (0.00 B)
```


Training and validation acc
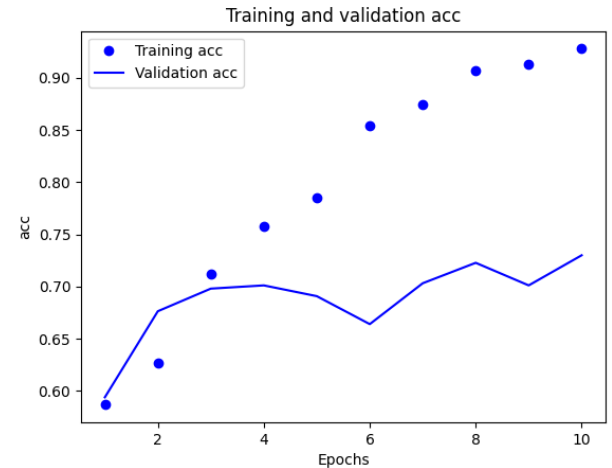
```
Epoch 1/10
78/78 ──────────────── 6s 42ms/step - accuracy: 0.5554 - loss: 0.9713 - val_accuracy: 0.5938 - val_loss: 0.9344
Epoch 2/10
78/78 ──────────────── 3s 36ms/step - accuracy: 0.6149 - loss: 0.8809 - val_accuracy: 0.6763 - val_loss: 0.8054
Epoch 3/10
78/78 ──────────────── 4s 48ms/step - accuracy: 0.7051 - loss: 0.7188 - val_accuracy: 0.6979 - val_loss: 0.7139
Epoch 4/10
78/78 ──────────────── 4s 48ms/step - accuracy: 0.7556 - loss: 0.5789 - val_accuracy: 0.7010 - val_loss: 0.7197
Epoch 5/10
78/78 ──────────────── 4s 48ms/step - accuracy: 0.7968 - loss: 0.4833 - val_accuracy: 0.6907 - val_loss: 0.7284
Epoch 6/10
78/78 ──────────────── 4s 47ms/step - accuracy: 0.8558 - loss: 0.3884 - val_accuracy: 0.6639 - val_loss: 0.8085
Epoch 7/10
78/78 ──────────────── 4s 46ms/step - accuracy: 0.8879 - loss: 0.3045 - val_accuracy: 0.7031 - val_loss: 0.7769
Epoch 8/10
78/78 ──────────────── 4s 46ms/step - accuracy: 0.9034 - loss: 0.2591 - val_accuracy: 0.7227 - val_loss: 0.8078
Epoch 9/10
78/78 ──────────────── 4s 47ms/step - accuracy: 0.9124 - loss: 0.2234 - val_accuracy: 0.7010 - val_loss: 0.8513
Epoch 10/10
78/78 ──────────────── 4s 47ms/step - accuracy: 0.9328 - loss: 0.1934 - val_accuracy: 0.7299 - val_loss: 0.8315
31/31 ──────────────── 0s 10ms/step - accuracy: 0.7189 - loss: 0.8420
0.8314988017082214
0.7298969030380249
31/31 ──────────────── 1s 18ms/step
p =  [[2.0870168e-05 9.9909186e-01 8.8730396e-04]
 [1.2035190e-06 9.9925989e-01 7.3890219e-04]
 [7.6636323e-04 2.9588489e-02 9.6964520e-01]
 ...
 [5.8105517e-02 8.7830400e-01 6.3590460e-02]
 [4.6681482e-01 1.3392235e-01 3.9926276e-01]
 [1.1143826e-04 9.8658341e-01 1.3305144e-02]]
1/1 ──────────────── 0s 34ms/step
pred  0
y_test[7:8]  [[False  True False]]
prediction   negative
```

This configuration has the most accuracy score compared to all there configuration even if the difference with the second is small. The accuracy score is ~73%. In this configuration the learning rate is 0.0005 and has the largest trainable parameters. The training loss is almost similar to the second configuration. However, the first one has the lowest training loss but higher validation loss which could be reflection of overfitting issue. When compared to accuracy, and loss, the third configuration has the highest accuracy as well as small gap in training and validation loss making it the better option for the generalization. The model learns the underlying patterns while lowest overfitting issue. The learning rate in this configuration is lowest which ensures slow and more stable convergence.

In comparison, configuration 3 has the best generalization than configuration 1 and 2.

The above results are when I ran the script in my windows environment. In order to run the script in linux: Following environment must be created:

conda create -n sp3 python=3.9.0 matplotlib nltk=3.9.1 numpy=1.19.5 pandas=1.3.4 scikit-learn=1.5.2 seaborn tensorflow=2.4.1 pydot notebook pytorch torchaudio tensorflow-gpu keras