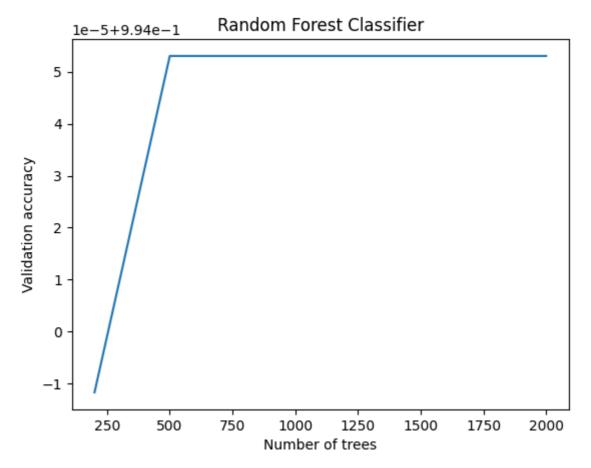
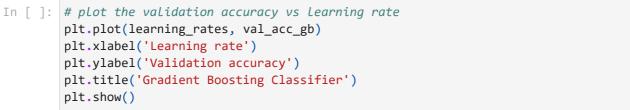
```
In [ ]: # imports
        import os
        import pandas as pd
        import numpy as np
        # Load data
        train = pd.read_csv('../data/processed/train_data_processed.csv')
        test = pd.read_csv('../data/processed/test_data_processed.csv')
        val = pd.read_csv('../data/processed/val_data_processed.csv')
In [ ]: # more feature engineering
        # use encoder to encode OCCURRED_ON_DATE column
        from sklearn.preprocessing import LabelEncoder
        le = LabelEncoder()
        train['OCCURRED_ON_DATE'] = le.fit_transform(train['OCCURRED_ON_DATE'])
        test['OCCURRED_ON_DATE'] = le.transform(test['OCCURRED_ON_DATE'])
        val['OCCURRED_ON_DATE'] = le.transform(val['OCCURRED_ON_DATE'])
In [ ]: # save Le
        import joblib
        joblib.dump(le, '../models/datetime_encoder.pkl')
Out[ ]: ['../models/datetime_encoder.pkl']
In [ ]: #drop _id column
        test = test.drop('_id', axis=1)
        val = val.drop('_id', axis=1)
In [ ]: # define the target variable
        y_train = train['Severe_crimes']
        y_test = test['Severe_crimes']
        y_val = val['Severe_crimes']
        # define the features
        X_train = train.drop(['Severe_crimes'], axis=1)
        X_test = test.drop(['Severe_crimes'], axis=1)
        X_val = val.drop(['Severe_crimes'], axis=1)
In [ ]: # test on 9 different models
        # first: random forest classifier
        from sklearn.ensemble import RandomForestClassifier
        # define the number of trees
        n_estimators = [200, 500, 1000, 1500, 2000]
        val_acc_rf = []
        # fit the model with the different number of trees
        for n in n estimators:
            rf = RandomForestClassifier(n_estimators=n, random_state=42)
            rf.fit(X_train, y_train)
            print(f'Number of trees: {n}')
            print(f'Train accuracy: {rf.score(X_train, y_train)}')
            print(f'Test accuracy: {rf.score(X_test, y_test)}')
            print(f'Validation accuracy: {rf.score(X_val, y_val)}')
```

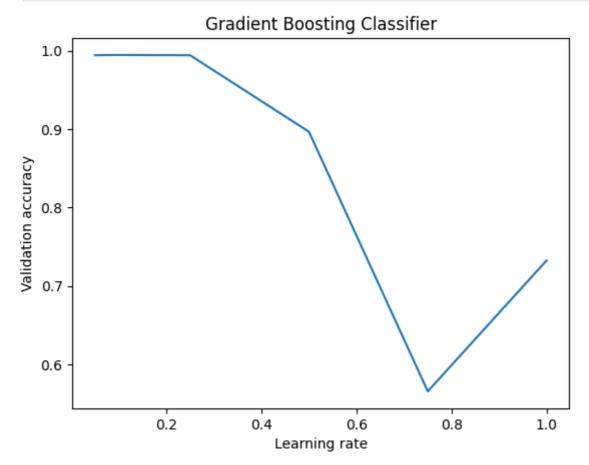
```
val_acc_rf.append(rf.score(X_val, y_val))
           print('----')
      Number of trees: 200
      Train accuracy: 0.9998541139856059
      Test accuracy: 0.9952540480178671
      Validation accuracy: 0.9939883645765999
      Number of trees: 500
      Train accuracy: 0.9998541139856059
      Test accuracy: 0.9955332216638749
      Validation accuracy: 0.9940530058177117
      -----
      Number of trees: 1000
      Train accuracy: 0.9998541139856059
      Test accuracy: 0.9952540480178671
      Validation accuracy: 0.9940530058177117
      -----
      Number of trees: 1500
      Train accuracy: 0.9998541139856059
      Test accuracy: 0.9955332216638749
      Validation accuracy: 0.9940530058177117
      Number of trees: 2000
      Train accuracy: 0.9998541139856059
      Test accuracy: 0.9955332216638749
      Validation accuracy: 0.9940530058177117
In [ ]: # draw the validation accuracy vs number of trees
       import matplotlib.pyplot as plt
       plt.plot(n_estimators, val_acc_rf)
       plt.xlabel('Number of trees')
       plt.ylabel('Validation accuracy')
       plt.title('Random Forest Classifier')
       plt.show()
       print(val_acc_rf)
```



[0.9939883645765999, 0.9940530058177117, 0.9940530058177117, 0.9940530058177117, 0.9940530058177117]

```
Learning rate: 0.05
Accuracy score (training): 0.994
Accuracy score (validation): 0.994
Learning rate: 0.1
Accuracy score (training): 0.995
Accuracy score (validation): 0.994
-----
Learning rate: 0.25
Accuracy score (training): 0.995
Accuracy score (validation): 0.994
Learning rate: 0.5
Accuracy score (training): 0.901
Accuracy score (validation): 0.897
-----
Learning rate: 0.75
Accuracy score (training): 0.568
Accuracy score (validation): 0.566
Learning rate: 1
Accuracy score (training): 0.730
Accuracy score (validation): 0.733
```





```
In [ ]: # model 3: Cnn
        from keras.models import Sequential
        from keras.layers import Dense, Dropout
        from keras import regularizers
        from sklearn.metrics import f1_score
        # test different dropout rates
        dropout_rates = [0.1, 0.2, 0.3, 0.4, 0.5]
        f1_scores = []
        val_acc_scores = []
        for rate in dropout_rates:
            model = Sequential()
            model.add(Dense(64, input_dim=7, activation='relu'))
            model.add(Dropout(rate))
            model.add(Dense(32, activation='relu'))
            model.add(Dropout(rate))
            model.add(Dense(1, activation='sigmoid'))
            # compile the model
            model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accura
            # fit the model
            history = model.fit(X_train, y_train, epochs=50, batch_size=64, validation_d
            # use the model on validation data and evaluate
            y_pred = model.predict(X_val)
            y_pred = (y_pred > 0.5)
            # f1 score
            f1 = f1_score(y_val, y_pred)
            f1_scores.append(f1)
            # validation accuracy
            val_acc = accuracy_score(y_val, y_pred)
            val_acc_scores.append(val_acc)
        print(f1_scores)
```

WARNING:tensorflow:From c:\Users\wangd\.conda\envs\BCAIML\Lib\site-packages\keras\src\losses.py:2976: The name tf.losses.sparse_softmax_cross_entropy is deprecate d. Please use tf.compat.v1.losses.sparse_softmax_cross_entropy instead.

WARNING:tensorflow:From c:\Users\wangd\.conda\envs\BCAIML\Lib\site-packages\keras \src\backend.py:873: The name tf.get_default_graph is deprecated. Please use tf.c ompat.v1.get_default_graph instead.

WARNING:tensorflow:From c:\Users\wangd\.conda\envs\BCAIML\Lib\site-packages\keras\src\optimizers__init__.py:309: The name tf.train.Optimizer is deprecated. Pleas e use tf.compat.v1.train.Optimizer instead.

Epoch 1/50

WARNING:tensorflow:From c:\Users\wangd\.conda\envs\BCAIML\Lib\site-packages\keras\src\utils\tf_utils.py:492: The name tf.ragged.RaggedTensorValue is deprecated. P lease use tf.compat.v1.ragged.RaggedTensorValue instead.

WARNING:tensorflow:From c:\Users\wangd\.conda\envs\BCAIML\Lib\site-packages\keras\src\engine\base_layer_utils.py:384: The name tf.executing_eagerly_outside_functions is deprecated. Please use tf.compat.v1.executing_eagerly_outside_functions in stead.

```
0.9153 - val_loss: 0.1361 - val_accuracy: 0.9366
Epoch 2/50
0.9326 - val_loss: 0.1357 - val_accuracy: 0.9366
Epoch 3/50
0.9357 - val_loss: 0.1256 - val_accuracy: 0.9418
Epoch 4/50
0.9401 - val_loss: 0.0971 - val_accuracy: 0.9529
Epoch 5/50
0.9557 - val_loss: 0.0911 - val_accuracy: 0.9679
Epoch 6/50
0.9646 - val_loss: 0.0846 - val_accuracy: 0.9715
Epoch 7/50
0.9666 - val loss: 0.0854 - val accuracy: 0.9725
Epoch 8/50
0.9686 - val_loss: 0.0806 - val_accuracy: 0.9727
Epoch 9/50
0.9705 - val loss: 0.0797 - val accuracy: 0.9753
Epoch 10/50
0.9717 - val_loss: 0.0751 - val_accuracy: 0.9785
Epoch 11/50
0.9731 - val_loss: 0.0766 - val_accuracy: 0.9754
Epoch 12/50
0.9745 - val_loss: 0.0752 - val_accuracy: 0.9772
Epoch 13/50
0.9757 - val_loss: 0.0743 - val_accuracy: 0.9780
```

```
Epoch 14/50
0.9763 - val_loss: 0.0713 - val_accuracy: 0.9791
Epoch 15/50
0.9771 - val_loss: 0.0688 - val_accuracy: 0.9786
Epoch 16/50
0.9787 - val_loss: 0.0652 - val_accuracy: 0.9793
Epoch 17/50
0.9787 - val loss: 0.0607 - val accuracy: 0.9862
Epoch 18/50
0.9786 - val_loss: 0.0643 - val_accuracy: 0.9814
Epoch 19/50
0.9792 - val_loss: 0.0557 - val_accuracy: 0.9861
Epoch 20/50
0.9797 - val_loss: 0.0591 - val_accuracy: 0.9837
Epoch 21/50
0.9805 - val_loss: 0.0543 - val_accuracy: 0.9880
Epoch 22/50
0.9806 - val_loss: 0.0709 - val_accuracy: 0.9818
Epoch 23/50
0.9809 - val loss: 0.0552 - val accuracy: 0.9873
Epoch 24/50
0.9816 - val_loss: 0.0604 - val_accuracy: 0.9876
Epoch 25/50
0.9816 - val_loss: 0.0553 - val_accuracy: 0.9869
Epoch 26/50
0.9821 - val_loss: 0.0552 - val_accuracy: 0.9868
Epoch 27/50
0.9827 - val loss: 0.0655 - val accuracy: 0.9805
Epoch 28/50
0.9835 - val_loss: 0.0550 - val_accuracy: 0.9888
Epoch 29/50
0.9835 - val loss: 0.0511 - val accuracy: 0.9856
Epoch 30/50
0.9849 - val_loss: 0.0467 - val_accuracy: 0.9896
Epoch 31/50
0.9849 - val_loss: 0.0454 - val_accuracy: 0.9893
Epoch 32/50
0.9853 - val_loss: 0.0523 - val_accuracy: 0.9872
Epoch 33/50
0.9854 - val_loss: 0.0458 - val_accuracy: 0.9904
```

```
Epoch 34/50
0.9853 - val_loss: 0.0495 - val_accuracy: 0.9893
Epoch 35/50
0.9856 - val_loss: 0.0519 - val_accuracy: 0.9843
Epoch 36/50
0.9830 - val_loss: 0.0657 - val_accuracy: 0.9835
Epoch 37/50
0.9806 - val loss: 0.0587 - val accuracy: 0.9805
Epoch 38/50
0.9808 - val_loss: 0.0570 - val_accuracy: 0.9811
Epoch 39/50
0.9820 - val_loss: 0.0532 - val_accuracy: 0.9867
Epoch 40/50
0.9843 - val_loss: 0.0469 - val_accuracy: 0.9910
Epoch 41/50
0.9845 - val_loss: 0.0734 - val_accuracy: 0.9791
Epoch 42/50
0.9851 - val_loss: 0.0472 - val_accuracy: 0.9906
Epoch 43/50
0.9858 - val loss: 0.0549 - val accuracy: 0.9893
Epoch 44/50
0.9864 - val_loss: 0.0439 - val_accuracy: 0.9904
Epoch 45/50
0.9864 - val_loss: 0.0423 - val_accuracy: 0.9928
Epoch 46/50
0.9873 - val_loss: 0.0420 - val_accuracy: 0.9917
Epoch 47/50
0.9882 - val loss: 0.0427 - val accuracy: 0.9912
Epoch 48/50
0.9868 - val_loss: 0.0610 - val_accuracy: 0.9793
Epoch 49/50
964/964 [============] - 1s 1ms/step - loss: 0.0572 - accuracy:
0.9862 - val loss: 0.0462 - val accuracy: 0.9870
Epoch 50/50
0.9869 - val_loss: 0.0433 - val_accuracy: 0.9928
484/484 [========= ] - 0s 756us/step
Epoch 1/50
0.9122 - val_loss: 0.1780 - val_accuracy: 0.9365
Epoch 2/50
0.9325 - val_loss: 0.1220 - val_accuracy: 0.9426
Epoch 3/50
```

```
0.9434 - val_loss: 0.1002 - val_accuracy: 0.9560
Epoch 4/50
0.9516 - val_loss: 0.0952 - val_accuracy: 0.9668
Epoch 5/50
964/964 [============] - 1s 1ms/step - loss: 0.1153 - accuracy:
0.9578 - val_loss: 0.0922 - val_accuracy: 0.9685
Epoch 6/50
0.9615 - val_loss: 0.0870 - val_accuracy: 0.9721
Epoch 7/50
0.9651 - val_loss: 0.0859 - val_accuracy: 0.9705
Epoch 8/50
0.9679 - val_loss: 0.0861 - val_accuracy: 0.9750
Epoch 9/50
0.9692 - val loss: 0.0837 - val accuracy: 0.9743
Epoch 10/50
964/964 [============] - 1s 2ms/step - loss: 0.0943 - accuracy:
0.9699 - val_loss: 0.0818 - val_accuracy: 0.9739
Epoch 11/50
0.9713 - val_loss: 0.0932 - val_accuracy: 0.9720
Epoch 12/50
0.9717 - val_loss: 0.0828 - val_accuracy: 0.9766
Epoch 13/50
0.9730 - val_loss: 0.0860 - val_accuracy: 0.9750
Epoch 14/50
0.9732 - val_loss: 0.0794 - val_accuracy: 0.9773
Epoch 15/50
0.9745 - val loss: 0.0770 - val accuracy: 0.9783
Epoch 16/50
0.9742 - val_loss: 0.0732 - val_accuracy: 0.9787
Epoch 17/50
0.9746 - val_loss: 0.0731 - val_accuracy: 0.9794
Epoch 18/50
0.9759 - val_loss: 0.0722 - val_accuracy: 0.9776
Epoch 19/50
0.9754 - val_loss: 0.0747 - val_accuracy: 0.9808
0.9761 - val_loss: 0.0716 - val_accuracy: 0.9794
Epoch 21/50
0.9766 - val loss: 0.0679 - val accuracy: 0.9793
Epoch 22/50
0.9767 - val_loss: 0.0721 - val_accuracy: 0.9783
Epoch 23/50
```

```
0.9775 - val_loss: 0.0700 - val_accuracy: 0.9792
Epoch 24/50
0.9778 - val_loss: 0.0631 - val_accuracy: 0.9830
Epoch 25/50
964/964 [============] - 1s 1ms/step - loss: 0.0751 - accuracy:
0.9790 - val_loss: 0.0787 - val_accuracy: 0.9816
Epoch 26/50
0.9793 - val_loss: 0.0751 - val_accuracy: 0.9802
Epoch 27/50
0.9796 - val_loss: 0.0592 - val_accuracy: 0.9835
Epoch 28/50
0.9799 - val_loss: 0.0550 - val_accuracy: 0.9884
Epoch 29/50
0.9821 - val loss: 0.0692 - val accuracy: 0.9804
Epoch 30/50
0.9812 - val_loss: 0.0568 - val_accuracy: 0.9858
Epoch 31/50
0.9817 - val_loss: 0.0611 - val_accuracy: 0.9807
Epoch 32/50
0.9801 - val_loss: 0.0581 - val_accuracy: 0.9813
Epoch 33/50
0.9811 - val_loss: 0.0603 - val_accuracy: 0.9858
Epoch 34/50
0.9810 - val_loss: 0.0593 - val_accuracy: 0.9877
Epoch 35/50
0.9820 - val loss: 0.0614 - val accuracy: 0.9832
Epoch 36/50
0.9784 - val_loss: 0.0578 - val_accuracy: 0.9885
Epoch 37/50
0.9794 - val_loss: 0.0575 - val_accuracy: 0.9846
Epoch 38/50
0.9797 - val_loss: 0.0590 - val_accuracy: 0.9833
Epoch 39/50
0.9802 - val_loss: 0.0534 - val_accuracy: 0.9915
0.9812 - val_loss: 0.0543 - val_accuracy: 0.9920
Epoch 41/50
0.9815 - val loss: 0.0635 - val accuracy: 0.9910
Epoch 42/50
0.9811 - val_loss: 0.0637 - val_accuracy: 0.9800
Epoch 43/50
```

```
0.9815 - val_loss: 0.0577 - val_accuracy: 0.9867
Epoch 44/50
0.9814 - val_loss: 0.0538 - val_accuracy: 0.9813
Epoch 45/50
964/964 [============] - 1s 1ms/step - loss: 0.0671 - accuracy:
0.9815 - val_loss: 0.0611 - val_accuracy: 0.9798
Epoch 46/50
0.9824 - val_loss: 0.0612 - val_accuracy: 0.9792
Epoch 47/50
0.9800 - val_loss: 0.0540 - val_accuracy: 0.9850
Epoch 48/50
0.9790 - val_loss: 0.0515 - val_accuracy: 0.9923
Epoch 49/50
0.9821 - val loss: 0.0508 - val accuracy: 0.9913
Epoch 50/50
964/964 [=============] - 1s 1ms/step - loss: 0.0615 - accuracy:
0.9826 - val_loss: 0.0495 - val_accuracy: 0.9897
484/484 [========= ] - 0s 764us/step
Epoch 1/50
0.8976 - val_loss: 0.1919 - val_accuracy: 0.9365
Epoch 2/50
0.9307 - val_loss: 0.1679 - val_accuracy: 0.9365
Epoch 3/50
0.9319 - val_loss: 0.1639 - val_accuracy: 0.9365
Epoch 4/50
0.9321 - val_loss: 0.1322 - val_accuracy: 0.9365
Epoch 5/50
0.9324 - val_loss: 0.1176 - val_accuracy: 0.9365
Epoch 6/50
0.9324 - val loss: 0.1088 - val accuracy: 0.9365
Epoch 7/50
0.9324 - val_loss: 0.1025 - val_accuracy: 0.9365
Epoch 8/50
964/964 [============] - 1s 1ms/step - loss: 0.1350 - accuracy:
0.9325 - val loss: 0.1061 - val accuracy: 0.9365
Epoch 9/50
0.9324 - val_loss: 0.1031 - val_accuracy: 0.9365
Epoch 10/50
0.9324 - val loss: 0.1015 - val accuracy: 0.9365
Epoch 11/50
0.9323 - val loss: 0.0982 - val accuracy: 0.9365
Epoch 12/50
0.9325 - val_loss: 0.0984 - val_accuracy: 0.9365
Epoch 13/50
```

```
0.9329 - val_loss: 0.0958 - val_accuracy: 0.9365
Epoch 14/50
0.9413 - val_loss: 0.0938 - val_accuracy: 0.9623
0.9588 - val_loss: 0.0931 - val_accuracy: 0.9722
Epoch 16/50
0.9622 - val_loss: 0.0879 - val_accuracy: 0.9699
Epoch 17/50
0.9651 - val_loss: 0.0899 - val_accuracy: 0.9725
Epoch 18/50
0.9653 - val_loss: 0.0837 - val_accuracy: 0.9684
Epoch 19/50
0.9662 - val_loss: 0.0845 - val_accuracy: 0.9716
Epoch 20/50
0.9677 - val_loss: 0.0860 - val_accuracy: 0.9715
Epoch 21/50
0.9677 - val_loss: 0.0857 - val_accuracy: 0.9723
Epoch 22/50
0.9674 - val_loss: 0.0823 - val_accuracy: 0.9710
Epoch 23/50
0.9697 - val_loss: 0.0876 - val_accuracy: 0.9703
Epoch 24/50
0.9708 - val loss: 0.0812 - val accuracy: 0.9727
Epoch 25/50
0.9714 - val_loss: 0.0780 - val_accuracy: 0.9745
Epoch 26/50
0.9724 - val loss: 0.0722 - val accuracy: 0.9770
Epoch 27/50
0.9722 - val_loss: 0.0698 - val_accuracy: 0.9781
Epoch 28/50
964/964 [============] - 1s 1ms/step - loss: 0.0854 - accuracy:
0.9724 - val loss: 0.0687 - val accuracy: 0.9796
Epoch 29/50
0.9751 - val_loss: 0.0707 - val_accuracy: 0.9794
Epoch 30/50
0.9756 - val loss: 0.0634 - val accuracy: 0.9825
Epoch 31/50
0.9757 - val loss: 0.0667 - val accuracy: 0.9785
Epoch 32/50
0.9762 - val_loss: 0.0653 - val_accuracy: 0.9828
Epoch 33/50
```

```
0.9775 - val_loss: 0.0623 - val_accuracy: 0.9800
Epoch 34/50
0.9782 - val_loss: 0.0588 - val_accuracy: 0.9835
0.9789 - val_loss: 0.0581 - val_accuracy: 0.9861
Epoch 36/50
0.9789 - val_loss: 0.0536 - val_accuracy: 0.9915
Epoch 37/50
0.9756 - val_loss: 0.0632 - val_accuracy: 0.9807
Epoch 38/50
0.9801 - val_loss: 0.0627 - val_accuracy: 0.9807
Epoch 39/50
0.9773 - val_loss: 0.0604 - val_accuracy: 0.9820
Epoch 40/50
964/964 [=============] - 1s 1ms/step - loss: 0.0718 - accuracy:
0.9796 - val_loss: 0.0539 - val_accuracy: 0.9897
Epoch 41/50
0.9791 - val_loss: 0.0524 - val_accuracy: 0.9907
Epoch 42/50
0.9790 - val_loss: 0.0737 - val_accuracy: 0.9720
Epoch 43/50
0.9818 - val_loss: 0.0561 - val_accuracy: 0.9853
Epoch 44/50
0.9819 - val loss: 0.0509 - val accuracy: 0.9911
Epoch 45/50
0.9811 - val_loss: 0.0621 - val_accuracy: 0.9807
Epoch 46/50
0.9804 - val loss: 0.0500 - val accuracy: 0.9915
Epoch 47/50
0.9795 - val_loss: 0.0573 - val_accuracy: 0.9798
Epoch 48/50
0.9811 - val loss: 0.0556 - val accuracy: 0.9826
Epoch 49/50
0.9816 - val_loss: 0.0522 - val_accuracy: 0.9853
Epoch 50/50
0.9801 - val loss: 0.0548 - val accuracy: 0.9909
484/484 [========= ] - 0s 767us/step
Epoch 1/50
0.8949 - val_loss: 0.1783 - val_accuracy: 0.9365
Epoch 2/50
0.9296 - val loss: 0.1743 - val accuracy: 0.9365
```

```
Epoch 3/50
0.9315 - val_loss: 0.1703 - val_accuracy: 0.9365
Epoch 4/50
0.9322 - val_loss: 0.1424 - val_accuracy: 0.9365
Epoch 5/50
0.9322 - val_loss: 0.1299 - val_accuracy: 0.9365
Epoch 6/50
0.9323 - val loss: 0.1129 - val accuracy: 0.9365
Epoch 7/50
0.9324 - val_loss: 0.1064 - val_accuracy: 0.9365
Epoch 8/50
0.9331 - val_loss: 0.1014 - val_accuracy: 0.9423
Epoch 9/50
0.9441 - val_loss: 0.1016 - val_accuracy: 0.9621
Epoch 10/50
0.9493 - val_loss: 0.0956 - val_accuracy: 0.9668
Epoch 11/50
0.9503 - val_loss: 0.0910 - val_accuracy: 0.9681
Epoch 12/50
964/964 [=============] - 1s 1ms/step - loss: 0.1168 - accuracy:
0.9517 - val loss: 0.0938 - val accuracy: 0.9663
Epoch 13/50
0.9575 - val_loss: 0.0899 - val_accuracy: 0.9666
Epoch 14/50
0.9611 - val_loss: 0.0877 - val_accuracy: 0.9698
Epoch 15/50
0.9627 - val_loss: 0.0854 - val_accuracy: 0.9696
Epoch 16/50
0.9645 - val loss: 0.0851 - val accuracy: 0.9721
Epoch 17/50
0.9663 - val_loss: 0.0882 - val_accuracy: 0.9726
Epoch 18/50
0.9662 - val loss: 0.0865 - val accuracy: 0.9690
Epoch 19/50
0.9669 - val_loss: 0.0833 - val_accuracy: 0.9750
Epoch 20/50
0.9679 - val_loss: 0.0809 - val_accuracy: 0.9747
Epoch 21/50
0.9695 - val_loss: 0.0754 - val_accuracy: 0.9760
Epoch 22/50
0.9694 - val_loss: 0.0831 - val_accuracy: 0.9721
```

```
Epoch 23/50
0.9709 - val_loss: 0.0771 - val_accuracy: 0.9747
Epoch 24/50
0.9719 - val_loss: 0.0769 - val_accuracy: 0.9772
Epoch 25/50
0.9728 - val_loss: 0.0711 - val_accuracy: 0.9774
Epoch 26/50
0.9735 - val loss: 0.0739 - val accuracy: 0.9809
Epoch 27/50
0.9740 - val_loss: 0.0640 - val_accuracy: 0.9801
Epoch 28/50
0.9755 - val_loss: 0.0722 - val_accuracy: 0.9778
Epoch 29/50
0.9753 - val_loss: 0.0600 - val_accuracy: 0.9844
Epoch 30/50
0.9751 - val_loss: 0.0605 - val_accuracy: 0.9853
Epoch 31/50
0.9772 - val_loss: 0.0724 - val_accuracy: 0.9773
Epoch 32/50
0.9776 - val loss: 0.0674 - val accuracy: 0.9789
Epoch 33/50
0.9764 - val_loss: 0.0623 - val_accuracy: 0.9867
Epoch 34/50
0.9772 - val_loss: 0.0629 - val_accuracy: 0.9786
Epoch 35/50
0.9776 - val_loss: 0.0670 - val_accuracy: 0.9791
Epoch 36/50
0.9771 - val loss: 0.0594 - val accuracy: 0.9826
Epoch 37/50
0.9779 - val_loss: 0.0606 - val_accuracy: 0.9797
Epoch 38/50
0.9776 - val loss: 0.0551 - val accuracy: 0.9855
Epoch 39/50
0.9783 - val_loss: 0.0583 - val_accuracy: 0.9887
Epoch 40/50
0.9789 - val_loss: 0.0597 - val_accuracy: 0.9902
Epoch 41/50
0.9773 - val_loss: 0.0605 - val_accuracy: 0.9791
Epoch 42/50
0.9776 - val_loss: 0.0594 - val_accuracy: 0.9844
```

```
Epoch 43/50
0.9794 - val_loss: 0.0590 - val_accuracy: 0.9802
Epoch 44/50
0.9802 - val_loss: 0.0526 - val_accuracy: 0.9884
Epoch 45/50
0.9788 - val_loss: 0.0689 - val_accuracy: 0.9789
Epoch 46/50
0.9787 - val loss: 0.0638 - val accuracy: 0.9788
Epoch 47/50
964/964 [=============] - 1s 2ms/step - loss: 0.0738 - accuracy:
0.9773 - val_loss: 0.0668 - val_accuracy: 0.9785
Epoch 48/50
0.9791 - val_loss: 0.0565 - val_accuracy: 0.9871
Epoch 49/50
0.9788 - val_loss: 0.0638 - val_accuracy: 0.9793
Epoch 50/50
0.9801 - val loss: 0.0575 - val accuracy: 0.9798
484/484 [=========] - 0s 914us/step
Epoch 1/50
0.8940 - val_loss: 0.1822 - val_accuracy: 0.9365
Epoch 2/50
0.9307 - val_loss: 0.1797 - val_accuracy: 0.9365
Epoch 3/50
0.9320 - val_loss: 0.1749 - val_accuracy: 0.9365
Epoch 4/50
0.9323 - val loss: 0.1538 - val accuracy: 0.9365
Epoch 5/50
0.9323 - val_loss: 0.1346 - val_accuracy: 0.9365
Epoch 6/50
0.9324 - val_loss: 0.1196 - val_accuracy: 0.9365
Epoch 7/50
0.9326 - val_loss: 0.1145 - val_accuracy: 0.9365
Epoch 8/50
0.9334 - val_loss: 0.1135 - val_accuracy: 0.9409
0.9362 - val_loss: 0.0969 - val_accuracy: 0.9591
Epoch 10/50
0.9536 - val loss: 0.0958 - val accuracy: 0.9696
Epoch 11/50
0.9608 - val_loss: 0.1038 - val_accuracy: 0.9590
Epoch 12/50
```

```
0.9617 - val_loss: 0.0989 - val_accuracy: 0.9648
Epoch 13/50
0.9620 - val_loss: 0.0909 - val_accuracy: 0.9685
Epoch 14/50
964/964 [============] - 1s 1ms/step - loss: 0.1052 - accuracy:
0.9652 - val_loss: 0.0902 - val_accuracy: 0.9661
Epoch 15/50
0.9647 - val_loss: 0.0917 - val_accuracy: 0.9722
Epoch 16/50
0.9683 - val_loss: 0.0840 - val_accuracy: 0.9769
Epoch 17/50
0.9693 - val_loss: 0.0839 - val_accuracy: 0.9774
Epoch 18/50
0.9696 - val loss: 0.0838 - val accuracy: 0.9754
Epoch 19/50
964/964 [============] - 1s 1ms/step - loss: 0.0959 - accuracy:
0.9703 - val_loss: 0.0902 - val_accuracy: 0.9721
Epoch 20/50
0.9712 - val_loss: 0.0794 - val_accuracy: 0.9785
Epoch 21/50
0.9700 - val_loss: 0.0781 - val_accuracy: 0.9765
Epoch 22/50
0.9714 - val_loss: 0.0711 - val_accuracy: 0.9785
Epoch 23/50
0.9716 - val_loss: 0.0811 - val_accuracy: 0.9800
Epoch 24/50
0.9718 - val loss: 0.0727 - val accuracy: 0.9785
Epoch 25/50
0.9726 - val_loss: 0.0741 - val_accuracy: 0.9863
Epoch 26/50
0.9714 - val_loss: 0.0728 - val_accuracy: 0.9768
Epoch 27/50
0.9731 - val_loss: 0.0711 - val_accuracy: 0.9782
Epoch 28/50
0.9727 - val_loss: 0.0697 - val_accuracy: 0.9777
0.9734 - val_loss: 0.0712 - val_accuracy: 0.9848
Epoch 30/50
0.9726 - val loss: 0.0638 - val accuracy: 0.9838
Epoch 31/50
0.9740 - val_loss: 0.0637 - val_accuracy: 0.9825
Epoch 32/50
```

```
0.9742 - val_loss: 0.0803 - val_accuracy: 0.9761
Epoch 33/50
0.9722 - val_loss: 0.0713 - val_accuracy: 0.9793
Epoch 34/50
964/964 [============] - 1s 1ms/step - loss: 0.0848 - accuracy:
0.9728 - val_loss: 0.0654 - val_accuracy: 0.9807
Epoch 35/50
0.9725 - val_loss: 0.0666 - val_accuracy: 0.9789
Epoch 36/50
0.9734 - val_loss: 0.0676 - val_accuracy: 0.9789
Epoch 37/50
0.9738 - val_loss: 0.0760 - val_accuracy: 0.9793
Epoch 38/50
0.9731 - val loss: 0.0665 - val accuracy: 0.9830
Epoch 39/50
964/964 [============] - 1s 1ms/step - loss: 0.0846 - accuracy:
0.9736 - val_loss: 0.0635 - val_accuracy: 0.9891
Epoch 40/50
0.9727 - val_loss: 0.0659 - val_accuracy: 0.9807
Epoch 41/50
0.9732 - val_loss: 0.0609 - val_accuracy: 0.9872
Epoch 42/50
0.9744 - val_loss: 0.0759 - val_accuracy: 0.9789
Epoch 43/50
0.9735 - val_loss: 0.0597 - val_accuracy: 0.9830
Epoch 44/50
0.9741 - val loss: 0.0723 - val accuracy: 0.9838
Epoch 45/50
0.9731 - val_loss: 0.0666 - val_accuracy: 0.9803
Epoch 46/50
0.9731 - val_loss: 0.0581 - val_accuracy: 0.9829
Epoch 47/50
0.9736 - val_loss: 0.0653 - val_accuracy: 0.9879
Epoch 48/50
0.9743 - val_loss: 0.0638 - val_accuracy: 0.9790
0.9743 - val_loss: 0.0705 - val_accuracy: 0.9770
Epoch 50/50
0.9740 - val loss: 0.0696 - val accuracy: 0.9787
484/484 [======== ] - 0s 720us/step
[0.9401617250673855, 0.9167533818938606, 0.9236599891716297, 0.848690591658584,
0.8411934552454284]
```

```
In [ ]: # draw the dropout rates vs f1 scores
import matplotlib.pyplot as plt
plt.plot(dropout_rates, f1_scores)
plt.xlabel('Dropout Rate')
plt.ylabel('F1 Score')
plt.title('Dropout Rate vs F1 Score')
plt.show()
```



0.88

0.86

0.84

0.10

Dropout Rate vs F1 Score

```
In [ ]: # draw the dropout rates vs validation accuracy
    plt.plot(dropout_rates, val_acc_scores)
    plt.xlabel('Dropout Rate')
    plt.ylabel('Validation Accuracy')
    plt.title('Dropout Rate vs Validation Accuracy')
    plt.show()
```

0.25

0.30

Dropout Rate

0.35

0.40

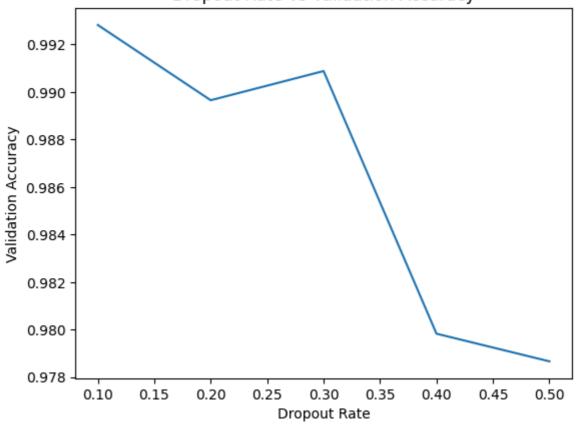
0.45

0.50

0.20

0.15

Dropout Rate vs Validation Accuracy



```
In []: #

In []: # model 4: XGBoost
    from xgboost import XGBClassifier

# test different Learning rates
learning_rates = [0.01, 0.05, 0.1, 0.2, 0.3]
val_acc_scores_xgb = []

for rate in learning_rates:
    xgb = XGBClassifier(learning_rate=rate)
    xgb.fit(X_train, y_train)
    print(f'Learning rate: {rate}')
    print(f'Train accuracy: {xgb.score(X_train, y_train)}')
    print(f'Test accuracy: {xgb.score(X_test, y_test)}')
    print(f'Validation accuracy: {xgb.score(X_val, y_val)}')
    val_acc_scores_xgb.append(xgb.score(X_val, y_val))
    print('------')
```

Learning rate: 0.01

Train accuracy: 0.9940673020813071 Test accuracy: 0.9949748743718593 Validation accuracy: 0.9942469295410472

Learning rate: 0.05

Train accuracy: 0.9946022174674188

Test accuracy: 0.9960915689558906

Validation accuracy: 0.9945054945054945

Learning rate: 0.1

Train accuracy: 0.994893989496207
Test accuracy: 0.9955332216638749
Validation accuracy: 0.9942469295410472

Learning rate: 0.2

Train accuracy: 0.9958503533683459
Test accuracy: 0.9955332216638749
Validation accuracy: 0.9939237233354881

Learning rate: 0.3

Train accuracy: 0.9966284121117811 Test accuracy: 0.9944165270798436 Validation accuracy: 0.9934712346477053

```
In [ ]: # draw the Learning rates vs validation accuracy
plt.plot(learning_rates, val_acc_scores_xgb)
plt.xlabel('Learning Rate')
plt.ylabel('Validation Accuracy')
plt.title('XGBoost Classifier')
plt.show()
```

