

Python 3.11.5 | packaged by Anaconda, Inc. | (main, Sep 11 2023, 13:26:23) [MSC v.1916 64 bit (AMD64)]  
Type "copyright", "credits" or "license" for more information.

IPython 8.15.0 -- An enhanced Interactive Python.

Restarting kernel...

```
In [1]: 'C:/Users/rincy/Desktop/ML_IDS - RFC/Code/ML-IDS_RF wo PCA.py' = 'C:/Users/rincy/Desktop/ML_IDS  
- RFC/Code'
```

Number of data points: 125973

Column names and data types:

duration	int64
protocol	object
service	object
flag	object
src_bytes	int64
dst_bytes	int64
land	int64
wrong_fragment	int64
urgent	int64
hot	int64
num_failed_logins	int64
logged_in	int64
num_compromised	int64
root_shell	int64
su_attempted	int64
num_root	int64
num_file_creations	int64
num_shells	int64
num_access_files	int64
num_outbound_cmds	int64
is_host_login	int64
is_guest_login	int64
count	int64
srv_count	int64
serror_rate	float64
srv_serror_rate	float64
rerror_rate	float64
srv_rerror_rate	float64
same_srv_rate	float64
diff_srv_rate	float64
srv_diff_host_rate	float64
dst_host_count	int64
dst_host_srv_count	int64
dst_host_same_srv_rate	float64
dst_host_diff_srv_rate	float64
dst_host_same_src_port_rate	float64
dst_host_srv_diff_host_rate	float64
dst_host_serror_rate	float64
dst_host_srv_serror_rate	float64
dst_host_rerror_rate	float64
dst_host_srv_rerror_rate	float64
class	object
level	int64

dtype: object

First few rows of the dataset:

	duration	protocol	service	...	dst_host_srv_rerror_rate	class	level
0	0	tcp	ftp_data	...	0.00	normal	20
1	0	udp	other	...	0.00	normal	15
2	0	tcp	private	...	0.00	neptune	19
3	0	tcp	http	...	0.01	normal	21
4	0	tcp	http	...	0.00	normal	21

[5 rows x 43 columns]

Number of data points: 125973

Number of features used: 3

Data points:

	src_bytes	dst_bytes	protocol
0	491	0	1
1	146	0	2
2	0	0	1
3	232	8153	1
4	199	420	1

```

...      ...      ...      ...
125968      0      0      1
125969      105     145     2
125970     2231     384     1
125971      0      0      1
125972     151      0      1

```

[125973 rows x 3 columns]

Fitting 3 folds for each of 16 candidates, totalling 48 fits

Confusion Matrix:

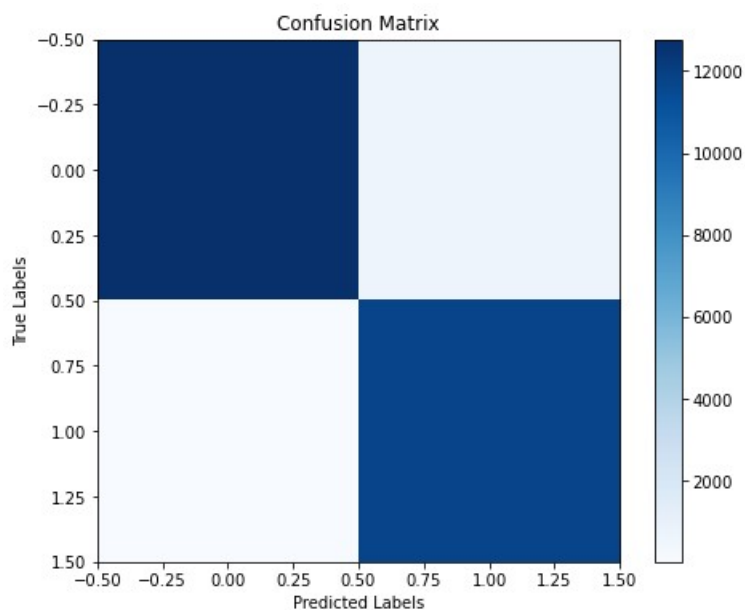
	Predicted 0	Predicted 1
Actual 0	12771	651
Actual 1	23	11750

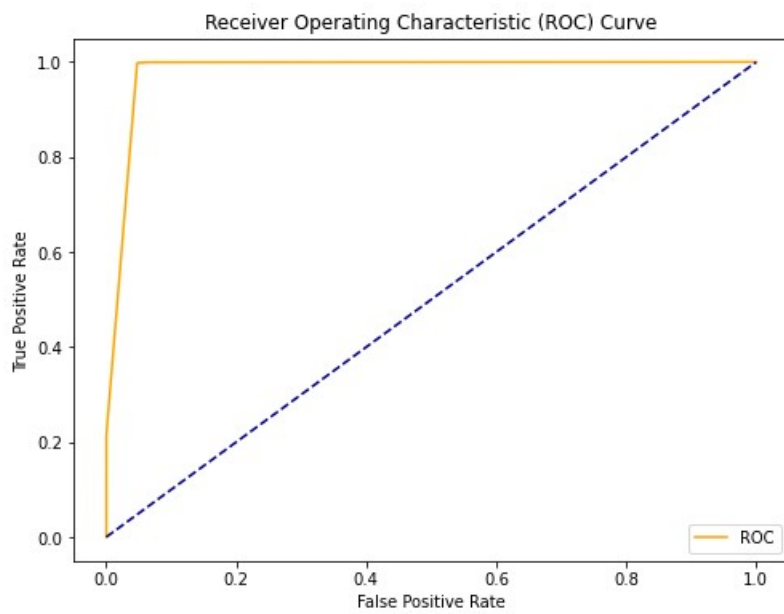
Classification Report:

	precision	recall	f1-score	support
0	1.00	0.95	0.97	13422
1	0.95	1.00	0.97	11773
accuracy			0.97	25195
macro avg	0.97	0.97	0.97	25195
weighted avg	0.97	0.97	0.97	25195

Best Hyperparameters: {'max\_depth': 20, 'min\_samples\_leaf': 1, 'min\_samples\_split': 5, 'n\_estimators': 100}

Test Accuracy (without PCA) with Best Hyperparameters: 0.9732486604485017





AUC Score: 0.9810941505952206

Average Training Complexity (seconds): 4.973517507314682

Average Running Complexity (seconds): 0.22885978718598685

In [2]: