

STC EXPERIMENT 6

```
[1]: import numpy as np
import scipy as sp
import matplotlib.pyplot as plt
import pandas as pd
    matplotlib inline
from sklearn.metrics import confusion_matrix

[2]: from sklearn import svm

[3]: df1=pd.read_csv("/home/spit/Downloads/traindata_10percent.csv")

[4]: df1
```

							0.1	0.2	0.3	0.4	...	9.1	1.00.1	\
0	0	tcp	http	SF	181	5450	0	0	0	0	...	19	1.0	
1	0	tcp	http	SF	235	1337	0	0	0	0	...	29	1.0	
2	0	tcp	http	SF	219	1337	0	0	0	0	...	39	1.0	
3	0	tcp	http	SF	217	2032	0	0	0	0	...	49	1.0	
4	0	tcp	http	SF	217	2032	0	0	0	0	...	59	1.0	
5	0	tcp	http	SF	212	1940	0	0	0	0	...	69	1.0	
6	0	tcp	http	SF	159	4087	0	0	0	0	...	79	1.0	
7	0	tcp	http	SF	210	151	0	0	0	0	...	89	1.0	
8	0	tcp	http	SF	212	786	0	0	0	1	...	99	1.0	
9	0	tcp	http	SF	210	624	0	0	0	0	...	109	1.0	
10	0	tcp	http	SF	177	1985	0	0	0	0	...	119	1.0	
11	0	tcp	http	SF	222	773	0	0	0	0	...	129	1.0	
12	0	tcp	http	SF	256	1169	0	0	0	0	...	139	1.0	
13	0	tcp	http	SF	241	259	0	0	0	0	...	149	1.0	
14	0	tcp	http	SF	260	1837	0	0	0	0	...	159	1.0	
15	0	tcp	http	SF	241	261	0	0	0	0	...	169	1.0	
16	0	tcp	http	SF	257	818	0	0	0	0	...	179	1.0	
17	0	tcp	http	SF	233	255	0	0	0	0	...	189	1.0	
18	0	tcp	http	SF	233	504	0	0	0	0	...	199	1.0	
19	0	tcp	http	SF	256	1273	0	0	0	0	...	209	1.0	
20	0	tcp	http	SF	234	255	0	0	0	0	...	219	1.0	
21	0	tcp	http	SF	241	259	0	0	0	0	...	229	1.0	
22	0	tcp	http	SF	239	968	0	0	0	0	...	239	1.0	
23	0	tcp	http	SF	245	1919	0	0	0	0	...	249	1.0	
24	0	tcp	http	SF	248	2129	0	0	0	0	...	255	1.0	
25	0	tcp	http	SF	354	1752	0	0	0	0	...	255	1.0	
26	0	tcp	http	SF	193	3991	0	0	0	0	...	255	1.0	
27	0	tcp	http	SF	214	14959	0	0	0	0	...	255	1.0	
28	0	tcp	http	SF	212	1309	0	0	0	0	...	255	1.0	
29	0	tcp	http	SF	215	3670	0	0	0	0	...	255	1.0	
...	

493990	0	tcp	http	SF	296	617	0	0	0	0	...	255	1.0
493991	0	tcp	http	SF	294	29288	0	0	0	0	...	255	1.0
493992	0	tcp	http	SF	285	34557	0	0	0	0	...	255	1.0
493993	0	tcp	http	SF	316	3665	0	0	0	0	...	255	1.0
493994	0	tcp	http	SF	335	10440	0	0	0	0	...	255	1.0
493995	0	tcp	http	SF	284	10592	0	0	0	0	...	255	1.0
493996	0	tcp	http	SF	242	7066	0	0	0	0	...	255	1.0
493997	0	tcp	http	SF	223	3707	0	0	0	0	...	255	1.0
493998	0	tcp	http	SF	204	1731	0	0	0	0	...	255	1.0
493999	0	tcp	http	SF	148	1122	0	0	0	0	...	255	1.0
494000	0	tcp	http	S0	0	0	0	0	0	0	...	255	1.0
494001	0	tcp	http	SF	215	2649	0	0	0	0	...	255	1.0
494002	0	tcp	http	SF	341	326	0	0	0	0	...	255	1.0
494003	0	tcp	http	SF	341	1943	0	0	0	0	...	255	1.0
494004	0	tcp	http	SF	341	1663	0	0	0	0	...	255	1.0
494005	0	tcp	http	SF	235	501	0	0	0	0	...	255	1.0
494006	0	tcp	http	SF	320	13828	0	0	0	0	...	255	1.0
494007	0	tcp	http	SF	319	1435	0	0	0	0	...	255	1.0
494008	0	tcp	http	SF	335	3435	0	0	0	0	...	255	1.0
494009	0	tcp	http	SF	291	236	0	0	0	0	...	255	1.0
494010	0	tcp	http	SF	308	662	0	0	0	0	...	255	1.0
494011	0	tcp	http	SF	291	1862	0	0	0	0	...	255	1.0
494012	0	tcp	http	SF	289	244	0	0	0	0	...	255	1.0
494013	0	tcp	http	SF	306	662	0	0	0	0	...	255	1.0
494014	0	tcp	http	SF	289	1862	0	0	0	0	...	255	1.0
494015	0	tcp	http	SF	310	1881	0	0	0	0	...	255	1.0
494016	0	tcp	http	SF	282	2286	0	0	0	0	...	255	1.0
494017	0	tcp	http	SF	203	1200	0	0	0	0	...	255	1.0
494018	0	tcp	http	SF	291	1200	0	0	0	0	...	255	1.0
494019	0	tcp	http	SF	219	1234	0	0	0	0	...	255	1.0

	0.00.6	0.11.1	0.00.7	0.00.8	0.00.9	0.00.10	0.00.11	normal.
0	0.0	0.05	0.00	0.00	0.00	0.00	0.0	normal.
1	0.0	0.03	0.00	0.00	0.00	0.00	0.0	normal.
2	0.0	0.03	0.00	0.00	0.00	0.00	0.0	normal.
3	0.0	0.02	0.00	0.00	0.00	0.00	0.0	normal.
4	0.0	0.02	0.00	0.00	0.00	0.00	0.0	normal.
5	0.0	1.00	0.04	0.00	0.00	0.00	0.0	normal.
6	0.0	0.09	0.04	0.00	0.00	0.00	0.0	normal.
7	0.0	0.12	0.04	0.00	0.00	0.00	0.0	normal.
8	0.0	0.12	0.05	0.00	0.00	0.00	0.0	normal.
9	0.0	0.06	0.05	0.00	0.00	0.00	0.0	normal.
10	0.0	0.04	0.04	0.00	0.00	0.00	0.0	normal.
11	0.0	0.03	0.04	0.00	0.00	0.00	0.0	normal.
12	0.0	0.25	0.04	0.00	0.00	0.00	0.0	normal.
13	0.0	0.07	0.04	0.00	0.00	0.00	0.0	normal.
14	0.0	0.04	0.04	0.00	0.00	0.00	0.0	normal.
15	0.0	0.03	0.04	0.00	0.00	0.00	0.0	normal.
16	0.0	0.02	0.03	0.00	0.00	0.00	0.0	normal.

17	0.0	0.02	0.03	0.00	0.00	0.00	0.0	normal.
18	0.0	0.02	0.03	0.00	0.00	0.00	0.0	normal.
19	0.0	0.01	0.03	0.00	0.00	0.00	0.0	normal.
20	0.0	0.01	0.03	0.00	0.00	0.00	0.0	normal.
21	0.0	0.01	0.03	0.00	0.00	0.00	0.0	normal.
22	0.0	0.33	0.03	0.00	0.00	0.00	0.0	normal.
23	0.0	0.08	0.03	0.00	0.00	0.00	0.0	normal.
24	0.0	0.04	0.03	0.00	0.00	0.00	0.0	normal.
25	0.0	0.20	0.04	0.00	0.00	0.00	0.0	normal.
26	0.0	1.00	0.05	0.00	0.00	0.00	0.0	normal.
27	0.0	0.09	0.05	0.00	0.00	0.00	0.0	normal.
28	0.0	0.05	0.05	0.00	0.00	0.00	0.0	normal.
29	0.0	0.03	0.05	0.00	0.00	0.00	0.0	normal.
...
493990	0.0	0.04	0.10	0.00	0.00	0.00	0.0	normal.
493991	0.0	0.03	0.10	0.00	0.00	0.00	0.0	normal.
493992	0.0	0.14	0.10	0.00	0.00	0.00	0.0	normal.
493993	0.0	0.06	0.09	0.00	0.00	0.00	0.0	normal.
493994	0.0	0.04	0.08	0.00	0.00	0.00	0.0	normal.
493995	0.0	0.03	0.07	0.00	0.00	0.00	0.0	normal.
493996	0.0	0.02	0.05	0.00	0.00	0.00	0.0	normal.
493997	0.0	0.11	0.05	0.00	0.00	0.00	0.0	normal.
493998	0.0	0.05	0.04	0.00	0.00	0.00	0.0	normal.
493999	0.0	0.33	0.04	0.00	0.00	0.00	0.0	normal.
494000	0.0	0.15	0.04	0.08	0.00	0.08	0.0	normal.
494001	0.0	0.04	0.04	0.04	0.00	0.04	0.0	normal.
494002	0.0	1.00	0.05	0.00	0.01	0.00	0.0	normal.
494003	0.0	0.09	0.05	0.00	0.01	0.00	0.0	normal.
494004	0.0	0.05	0.05	0.00	0.01	0.00	0.0	normal.
494005	0.0	0.50	0.05	0.00	0.01	0.00	0.0	normal.
494006	0.0	0.10	0.05	0.00	0.01	0.00	0.0	normal.
494007	0.0	0.17	0.07	0.00	0.01	0.00	0.0	normal.
494008	0.0	0.06	0.07	0.00	0.01	0.00	0.0	normal.
494009	0.0	0.04	0.06	0.00	0.01	0.00	0.0	normal.
494010	0.0	0.03	0.06	0.00	0.01	0.00	0.0	normal.
494011	0.0	0.02	0.05	0.00	0.01	0.00	0.0	normal.
494012	0.0	0.02	0.05	0.00	0.01	0.00	0.0	normal.
494013	0.0	0.02	0.05	0.00	0.01	0.00	0.0	normal.
494014	0.0	0.01	0.05	0.00	0.01	0.00	0.0	normal.
494015	0.0	0.01	0.05	0.00	0.01	0.00	0.0	normal.
494016	0.0	0.17	0.05	0.00	0.01	0.00	0.0	normal.
494017	0.0	0.06	0.05	0.06	0.01	0.00	0.0	normal.
494018	0.0	0.04	0.05	0.04	0.01	0.00	0.0	normal.
494019	0.0	0.17	0.05	0.00	0.01	0.00	0.0	normal.

[494020 rows x 42 columns]

```
[5]: df2=pd.read_csv("/home/spit/Downloads/testdata_10percent.csv.zip")
[6]: df2
```

[6]	0	udp	private	SF	105	146	0.1	0.2	0.3	0.4	...	241	0.95	\
:														
0	14	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
1	0	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
2	0	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
3	0	udp	private	SF	105	0	0	0	0	0	...	241	0.95	
4	0	udp	private	SF	105	0	0	0	0	0	...	241	0.95	
5	0	udp	private	SF	105	145	0	0	0	0	...	241	0.95	
6	0	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
7	0	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
8	0	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
9	0	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
10	5	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
11	0	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
12	0	udp	private	SF	105	0	0	0	0	0	...	241	0.95	
13	0	udp	private	SF	105	0	0	0	0	0	...	241	0.95	
14	14	udp	private	SF	105	146	0	0	0	0	...	241	0.95	
15	0	udp	private	SF	105	0	0	0	0	0	...	242	0.95	
16	5	udp	private	SF	105	146	0	0	0	0	...	243	0.95	
17	0	icmp	urp_i	SF	182	0	0	0	0	0	...	10	0.04	
18	0	icmp	urp_i	SF	182	0	0	0	0	0	...	11	0.04	
19	0	icmp	urp_i	SF	182	0	0	0	0	0	...	12	0.05	
20	0	icmp	urp_i	SF	182	0	0	0	0	0	...	13	0.05	
21	0	icmp	urp_i	SF	182	0	0	0	0	0	...	14	0.05	
22	0	icmp	urp_i	SF	182	0	0	0	0	0	...	15	0.06	
23	0	udp	private	SF	105	146	0	0	0	0	...	238	0.93	
24	0	udp	private	SF	105	146	0	0	0	0	...	238	0.93	
25	0	udp	private	SF	105	146	0	0	0	0	...	238	0.93	
26	0	udp	private	SF	105	146	0	0	0	0	...	238	0.93	
27	0	udp	private	SF	105	146	0	0	0	0	...	238	0.93	
28	0	icmp	eco_i	SF	30	0	0	0	0	0	...	3	0.01	
29	0	udp	private	SF	105	146	0	0	0	0	...	237	0.93	
...
399969	0	tcp	http	SF	283	466	0	0	0	0	...	255	1.00	
399970	0	tcp	http	SF	289	1096	0	0	0	0	...	255	1.00	
399971	0	tcp	http	SF	289	1862	0	0	0	0	...	255	1.00	
399972	0	tcp	http	SF	203	242	0	0	0	0	...	255	1.00	
399973	0	tcp	http	S1	196	0	0	0	0	0	...	255	1.00	
399974	0	tcp	http	SF	143	17463	0	0	0	0	...	255	1.00	
399975	0	tcp	http	SF	202	4017	0	0	0	0	...	255	1.00	
399976	0	tcp	http	SF	0	234	0	0	0	0	...	255	1.00	
399977	0	tcp	http	SF	203	1200	0	0	0	0	...	255	1.00	
399978	0	tcp	http	SF	198	2169	0	0	0	0	...	255	1.00	
399979	0	tcp	http	SF	197	466	0	0	0	0	...	255	1.00	
399980	0	tcp	http	SF	203	1862	0	0	0	0	...	255	1.00	
399981	0	tcp	http	SF	203	1096	0	0	0	0	...	255	1.00	
399982	0	tcp	http	SF	284	2286	0	0	0	0	...	255	1.00	
399983	0	tcp	http	SF	290	4017	0	0	0	0	...	255	1.00	

399984	0	tcp	http	SF	291	234	0	0	0	0	...	255	1.00
399985	0	tcp	http	SF	291	242	0	0	0	0	...	255	1.00
399986	0	tcp	http	SF	231	17463	0	0	0	0	...	255	1.00
399987	0	tcp	http	SF	291	1200	0	0	0	0	...	255	1.00
399988	0	tcp	http	SF	285	466	0	0	0	0	...	255	1.00
399989	0	tcp	http	SF	286	2169	0	0	0	0	...	255	1.00
399990	0	tcp	http	SF	291	1096	0	0	0	0	...	255	1.00
399991	0	tcp	http	SF	291	1862	0	0	0	0	...	255	1.00
399992	0	tcp	http	SF	159	15808	0	0	0	0	...	255	1.00
399993	0	tcp	http	SF	219	244	0	0	0	0	...	255	1.00
399994	0	tcp	http	SF	212	2288	0	0	0	0	...	255	1.00
399995	0	tcp	http	SF	219	236	0	0	0	0	...	255	1.00
399996	0	tcp	http	SF	218	3610	0	0	0	0	...	255	1.00
399997	0	tcp	http	SF	219	1234	0	0	0	0	...	255	1.00
399998	0	tcp	http	SF	219	1098	0	0	0	0	...	255	1.00

	0.01	0.00.6	0.00.7	0.00.8	0.00.9	0.00.10	0.00.11	normal.
0	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
1	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
2	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
3	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
4	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
5	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
6	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
7	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
8	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
9	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
10	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
11	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
12	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
13	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
14	0.01	0.02	0.00	0.00	0.00	0.0	0.0	normal.
15	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
16	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
17	0.01	0.05	0.00	0.00	0.00	0.0	0.0	normal.
18	0.01	0.05	0.00	0.00	0.00	0.0	0.0	normal.
19	0.01	0.05	0.00	0.00	0.00	0.0	0.0	normal.
20	0.01	0.06	0.00	0.00	0.00	0.0	0.0	normal.
21	0.01	0.06	0.00	0.00	0.00	0.0	0.0	normal.
22	0.01	0.07	0.00	0.00	0.00	0.0	0.0	normal.
23	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
24	0.01	0.01	0.00	0.00	0.00	0.0	0.0	normal.
25	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
26	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
27	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
28	0.01	0.07	0.00	0.00	0.00	0.0	0.0	normal.
29	0.01	0.00	0.00	0.00	0.00	0.0	0.0	normal.
...
399969	0.00	0.12	0.05	0.00	0.01	0.0	0.0	normal.

399970	0.00	0.11	0.05	0.00	0.01	0.0	0.0	normal.
399971	0.00	0.10	0.05	0.00	0.01	0.0	0.0	normal.
399972	0.00	0.09	0.05	0.00	0.01	0.0	0.0	normal.
399973	0.00	0.08	0.05	0.08	0.01	0.0	0.0	normal.
399974	0.00	0.08	0.05	0.08	0.01	0.0	0.0	normal.
399975	0.00	0.07	0.05	0.07	0.01	0.0	0.0	normal.
399976	0.00	0.07	0.05	0.07	0.01	0.0	0.0	normal.
399977	0.00	0.06	0.05	0.06	0.01	0.0	0.0	normal.
399978	0.00	0.06	0.05	0.06	0.01	0.0	0.0	normal.
399979	0.00	0.06	0.05	0.06	0.01	0.0	0.0	normal.
399980	0.00	0.05	0.05	0.05	0.01	0.0	0.0	normal.
399981	0.00	0.05	0.05	0.05	0.01	0.0	0.0	normal.
399982	0.00	0.05	0.05	0.05	0.01	0.0	0.0	normal.
399983	0.00	0.05	0.05	0.05	0.01	0.0	0.0	normal.
399984	0.00	0.04	0.05	0.04	0.01	0.0	0.0	normal.
399985	0.00	0.04	0.05	0.04	0.01	0.0	0.0	normal.
399986	0.00	0.04	0.05	0.04	0.01	0.0	0.0	normal.
399987	0.00	0.04	0.05	0.04	0.01	0.0	0.0	normal.
399988	0.00	0.04	0.05	0.04	0.01	0.0	0.0	normal.
399989	0.00	0.04	0.05	0.04	0.01	0.0	0.0	normal.
399990	0.00	0.03	0.05	0.03	0.01	0.0	0.0	normal.
399991	0.00	0.03	0.05	0.03	0.01	0.0	0.0	normal.
399992	0.00	1.00	0.05	0.00	0.01	0.0	0.0	normal.
399993	0.00	0.50	0.05	0.00	0.01	0.0	0.0	normal.
399994	0.00	0.33	0.05	0.00	0.01	0.0	0.0	normal.
399995	0.00	0.25	0.05	0.00	0.01	0.0	0.0	normal.
399996	0.00	0.20	0.05	0.00	0.01	0.0	0.0	normal.
399997	0.00	0.17	0.05	0.00	0.01	0.0	0.0	normal.
399998	0.00	0.14	0.05	0.00	0.01	0.0	0.0	normal.

[399999 rows x 42 columns]

[7]:

```

    'protocol_type',
    'service',
    'flag',
    'src_bytes',
    'dst_bytes',
    'land',
    'wrong_fragment',
    'urgent',
    'hot',
[8] 'num_failed_logins',
:  'logged_in',
    'num_compromised',
    'root_shell',
    'su_attempted',
    'num_root',
    'num_file_creations',
    'num_shells',
    'num_access_files',
    'num_outbound_cmds',
    'is_host_login',
    'is_guest_login',
    'count',
    'srv_count',
    'serror_rate',
    'srv_serror_rate',
    'rerror_rate',
    'srv_rerror_rate',
    'same_srv_rate',
    'diff_srv_rate',
    'srv_diff_host_rate',
    'dst_host_count',
    'dst_host_srv_count',
    'dst_host_same_srv_rate',
    'dst_host_diff_srv_rate',
    'dst_host_same_src_port_rate',
    'dst_host_srv_diff_host_rate',
    'dst_host_serror_rate',
    'dst_host_srv_serror_rate',
    'dst_host_rerror_rate',
    'dst_host_srv_rerror_rate',
    'outcome'
]

df2.columns = [
    'duration',
    'protocol_type',
    'service',

```

```

'flag',
'src_bytes',
'dst_bytes',
'land',
'wrong_fragment',
'urgent',
'hot',
'num_failed_logins',
'logged_in',
'num_compromised',
'root_shell',
'su_attempted',
'num_root',
'num_file_creations',
'num_shells',
'num_access_files',
'num_outbound_cmds',
'is_host_login',
'is_guest_login',
'count',
'srv_count',
'serror_rate',
'srv_serror_rate',
'rerror_rate',
'srv_rerror_rate',
'same_srv_rate',
'diff_srv_rate',
'srv_diff_host_rate',
'dst_host_count',
'dst_host_srv_count',
'dst_host_same_srv_rate',
'dst_host_diff_srv_rate',
'dst_host_same_src_port_rate',
'dst_host_srv_diff_host_rate',
'dst_host_serror_rate',
'dst_host_srv_serror_rate',
'dst_host_rerror_rate',
'dst_host_srv_rerror_rate',
'outcome'
]

```

[9]: df1.dtypes

```

[9]: duration          int64
    protocol_type      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
outcome             object
dtype: object

```

```

[10]: duration          int64
protocol_type          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
outcome             object
dtype: object

```

```
[12]: (494020, 38)      xtrain = df1.select_dtypes(exclude="object")
```

```
[13]:                  df1.shape
```

```
[13]: (494020, 42)
```

```
[14]:                  xtest = df2.select_dtypes(exclude="object")
```

```
[15]:                  xtest.shape
```

[15]: (399999, 38)

[16]: df2.shape

[16]: (399999, 42)

[17]: ytest=df2.outcome

[18]: ytest.value_counts()

[18]: neptune. 210910
normal. 149207
smurf. 36529
portsweep. 1804
ipsweep. 1314
teardrop. 100
back. 100
pod. 22
buffer_overflow. 9
land. 3
satan. 1
Name: outcome, dtype: int64

[19]: ytest.empty

[19]: False

[20]: ytrain=df1.outcome

[21]: ytrain.value_counts()

[21]: smurf. 280790
neptune. 107201
normal. 97277
back. 2203
satan. 1589
ipsweep. 1247
portsweep. 1040
warezclient. 1020
teardrop. 979
pod. 264
nmap. 231
guess_passwd. 53
buffer_overflow. 30
land. 21
warezmaster. 20
imap. 12
rootkit. 10
loadmodule. 9
ftp_write. 8
multihop. 7
phf. 4

[22]:

```
perl spy. 2
. Name: outcome, dtype: int64
```

```
[22]: False ytrain.empty
```

```
[23]: from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
```

```
[25]: model_LDA = LinearDiscriminantAnalysis()
```

```
[43]: #Train the model import time import sys t0=time.time()
      model_LDA.fit(xtrain,ytrain) tt=time.time() - t0 print(tt)
```

```
2.8985960483551025
```

```
/home/spit/anaconda3/lib/python3.7/site-
packages/sklearn/discriminant_analysis.py:388: UserWarning: Variables are
collinear.
  warnings.warn("Variables are collinear.")
```

```
[29]: #Prediction import time import sys t0=time.time()
      score = model_LDA.score(xtest,ytest) tt=time.time() - t0
      print(tt) print(score*100)
```

```
0.35146045684814453
99.1672479181198
```

```
[32]: from sklearn.metrics import classification_report t0=time.time()
      ypred = model_LDA.predict(xtest) tt=time.time()-t0
      print(tt) cr=classification_report(ypred,ytest) print(cr)
```

```
0.12423944473266602
```

```
precision    recall  f1-score   support
```

back.	1.00	0.31	0.47	325
buffer_overflow.	0.78	0.64	0.70	11
ipsweep.	0.98	0.76	0.86	1689
land.	0.67	0.07	0.12	30
neptune.	1.00	1.00	1.00	210775
normal.	0.98	1.00	0.99	146616
pod.	0.91	0.95	0.93	21
portsweep.	1.00	0.70	0.82	2562
satan.	1.00	0.00	0.00	1347
smurf.	0.99	0.99	0.99	36524
teardrop.	0.99	1.00	0.99	99
micro avg	0.99	0.99	0.99	399999
macro avg	0.94	0.67	0.72	399999
weighted avg	0.99	0.99	0.99	399999

```
[33]: #Create confusion Matrix
      from pandas import crosstab
      crosstab(ytest,ypred)
```

```
[34]:
```

col_0	back.	buffer_overflow.	ipsweep.	land.	neptune.	normal.	\
outcome							
back.	100	0	0	0	0	0	
buffer_overflow.	0	7	0	0	2	0	
ipsweep.	0	0	1290	0	1	3	
land.	0	0	0	2	0	0	
neptune.	0	0	0	14	210731	2	
normal.	225	4	399	14	41	146348	
pod.	0	0	0	0	0	0	
portsweep.	0	0	0	0	0	9	
satan.	0	0	0	0	0	0	
smurf.	0	0	0	0	0	254	
teardrop.	0	0	0	0	0	0	

col_0	pod.	portsweep.	satan.	smurf.	teardrop.
outcome					
back.	0	0	0	0	0
buffer_overflow.	0	0	0	0	0
ipsweep.	0	0	20	0	0
land.	0	0	1	0	0
neptune.	0	163	0	0	0
normal.	0	604	1323	249	0
pod.	20	0	2	0	0
portsweep.	0	1795	0	0	0
satan.	0	0	1	0	0
smurf.	0	0	0	36275	0
teardrop.	1	0	0	0	99

```
[35]: ytest.value_counts()
```

```
[35]: neptune.      210910
      normal.     149207
      smurf.      36529
      portsweep.  1804
      ipsweep.    1314
      teardrop.   100
      back.       100
      pod.        22
      buffer_overflow. 9
      land.       3
      satan.      1
      Name: outcome, dtype: int64
```

```
[36]: 36275/36529
```

```
[36]: 0.9930466204933067
```

```
[38]: #Save the model
      from sklearn.externals import joblib
```

```
[39]: joblib.dump(model_LDA,'LDA_trainedModel')
      ['LDA_trainedModel']
```

```
[39]: #Check the model
      mymodel = joblib.load('LDA_trainedModel')
```

```
[41]: mymodel.score(xtest,ytest)
```

```
[42]: 0.9916724791811979
```

```
[42]:
[ ]:
```