In [1]:

```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd



from sklearn import metrics
from sklearn.metrics import classification_report,confusion_matrix
```

In [5]:

```
import warnings
warnings.filterwarnings(action="ignore")
%matplotlib inline
pd.set_option("display.max_rows", 1000)
pd.set_option("display.max_columns", 1000)
```

In [8]:

```
fires = pd.read_csv("/content/drive/MyDrive/forestfires.csv")     #reading the dataset
fires.head(15)     #show the first 15 instances of dataset
```

Out[8]:

	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
0	7	5	mar	fri	86.2	26.2	94.3	5.1	8.2	51	6.7	0.0	0.0
1	7	4	oct	tue	90.6	35.4	669.1	6.7	18.0	33	0.9	0.0	0.0
2	7	4	oct	sat	90.6	43.7	686.9	6.7	14.6	33	1.3	0.0	0.0
3	8	6	mar	fri	91.7	33.3	77.5	9.0	8.3	97	4.0	0.2	0.0
4	8	6	mar	sun	89.3	51.3	102.2	9.6	11.4	99	1.8	0.0	0.0
5	8	6	aug	sun	92.3	85.3	488.0	14.7	22.2	29	5.4	0.0	0.0
6	8	6	aug	mon	92.3	88.9	495.6	8.5	24.1	27	3.1	0.0	0.0
7	8	6	aug	mon	91.5	145.4	608.2	10.7	8.0	86	2.2	0.0	0.0
8	8	6	sep	tue	91.0	129.5	692.6	7.0	13.1	63	5.4	0.0	0.0
9	7	5	sep	sat	92.5	88.0	698.6	7.1	22.8	40	4.0	0.0	0.0
10	7	5	sep	sat	92.5	88.0	698.6	7.1	17.8	51	7.2	0.0	0.0
11	7	5	sep	sat	92.8	73.2	713.0	22.6	19.3	38	4.0	0.0	0.0
12	6	5	aug	fri	63.5	70.8	665.3	0.8	17.0	72	6.7	0.0	0.0
13	6	5	sep	mon	90.9	126.5	686.5	7.0	21.3	42	2.2	0.0	0.0
14	6	5	sep	wed	92.9	133.3	699.6	9.2	26.4	21	4.5	0.0	0.0

In [9]:

```
#show the last 10 instances of dataset
fires.tail(10)
```

Out[9]:

	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
507	2	4	aug	fri	91.0	166.9	752.6	7.1	25.9	41	3.6	0.0	0.00
508	1	2	aug	fri	91.0	166.9	752.6	7.1	25.9	41	3.6	0.0	0.00
509	5	4	aug	fri	91.0	166.9	752.6	7.1	21.1	71	7.6	1.4	2.17
510	6	5	aug	fri	91.0	166.9	752.6	7.1	18.2	62	5.4	0.0	0.43

	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
511	8	6	aug	sun	81.6	56.7	665.6	1.9	27.8	35	2.7	0.0	0.00
512	4	3	aug	sun	81.6	56.7	665.6	1.9	27.8	32	2.7	0.0	6.44
513	2	4	aug	sun	81.6	56.7	665.6	1.9	21.9	71	5.8	0.0	54.29
514	7	4	aug	sun	81.6	56.7	665.6	1.9	21.2	70	6.7	0.0	11.16
515	1	4	aug	sat	94.4	146.0	614.7	11.3	25.6	42	4.0	0.0	0.00
516	6	3	nov	tue	79.5	3.0	106.7	1.1	11.8	31	4.5	0.0	0.00

In [10]:

```
#generate descriptive statistics of each attribute
fires.describe().T
```

Out[10]:

	count	mean	std	min	25%	50%	75%	max
X	517.0	4.669246	2.313778	1.0	3.0	4.00	7.00	9.00
Y	517.0	4.299807	1.229900	2.0	4.0	4.00	5.00	9.00
FFMC	517.0	90.644681	5.520111	18.7	90.2	91.60	92.90	96.20
DMC	517.0	110.872340	64.046482	1.1	68.6	108.30	142.40	291.30
DC	517.0	547.940039	248.066192	7.9	437.7	664.20	713.90	860.60
ISI	517.0	9.021663	4.559477	0.0	6.5	8.40	10.80	56.10
temp	517.0	18.889168	5.806625	2.2	15.5	19.30	22.80	33.30
RH	517.0	44.288201	16.317469	15.0	33.0	42.00	53.00	100.00
wind	517.0	4.017602	1.791653	0.4	2.7	4.00	4.90	9.40
rain	517.0	0.021663	0.295959	0.0	0.0	0.00	0.00	6.40
area	517.0	12.847292	63.655818	0.0	0.0	0.52	6.57	1090.84

In [11]:

```
#given area of land burnt, but we have to predict if there is fire or not so changing val
ues of area to 0 and 1 only
#here 0 represet there is not fire and 1 represent fire, changing all values of area whic
h are greater than 0 to 1
fires['area'].values[fires['area'].values > 0] = 1

#renaming the area attribute to output for clear understanding
fires = fires.rename(columns={'area': 'output'})
fires.head(5)
```

Out[11]:

	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	output
0	7	5	mar	fri	86.2	26.2	94.3	5.1	8.2	51	6.7	0.0	0.0
1	7	4	oct	tue	90.6	35.4	669.1	6.7	18.0	33	0.9	0.0	0.0
2	7	4	oct	sat	90.6	43.7	686.9	6.7	14.6	33	1.3	0.0	0.0
3	8	6	mar	fri	91.7	33.3	77.5	9.0	8.3	97	4.0	0.2	0.0
4	8	6	mar	sun	89.3	51.3	102.2	9.6	11.4	99	1.8	0.0	0.0

In [12]:

```
#Compute pairwise correlation of columns
fires.corr()
```

Out[12]:

	X	Y	FFMC	DMC	DC	ISI	temp	RH	wind	rain	output
X	1.000000	0.539548	0.021039	0.048384	0.085916	0.006210	0.051258	0.085223	0.018798	0.065387	0.062491

	X	Y	FFMC	DMC	DC	ISI	temp	RH	wind	rain	output
X	1.000000	0.539548	-0.021039	-0.048384	-0.085916	0.006210	-0.051258	0.085223	0.018798	0.065387	0.062491
Y	0.539548	1.000000	0.046308	0.007782	0.101178	0.024488	0.024103	0.062221	0.020341	0.033234	0.056892
FFMC	-0.021039	-0.046308	1.000000	0.382619	0.330512	0.531805	0.431532	-0.300995	-0.028485	0.056702	0.073823
DMC	-0.048384	0.007782	0.382619	1.000000	0.682192	0.305128	0.469594	0.073795	-0.105342	0.074790	0.062672
DC	-0.085916	-0.101178	0.330512	0.682192	1.000000	0.229154	0.496208	-0.039192	-0.203466	0.035861	0.096724
ISI	0.006210	-0.024488	0.531805	0.305128	0.229154	1.000000	0.394287	-0.132517	0.106826	0.067668	0.035663
temp	-0.051258	-0.024103	0.431532	0.469594	0.496208	0.394287	1.000000	-0.527390	-0.227116	0.069491	0.076047
RH	0.085223	0.062221	-0.300995	0.073795	-0.039192	-0.132517	-0.527390	1.000000	0.069410	0.099751	-0.035587
wind	0.018798	-0.020341	-0.028485	-0.105342	-0.203466	0.106826	-0.227116	0.069410	1.000000	0.061119	0.055702
rain	0.065387	0.033234	0.056702	0.074790	0.035861	0.067668	0.069491	0.099751	0.061119	1.000000	0.025550
output	0.062491	0.056892	0.073823	0.062672	0.096724	0.035663	0.076047	-0.035587	0.055702	0.025550	1.000000

In [ ]: