Program Offered: M. Tech / Data Science¶

Course Title: Machine Learning Supervised Classification (MLSC)

Group Number: 16

Name of the Project Members

- Tania Khan
- Vinod A
- Sandeep Ram
- Emmanuel Pradeep Aloysius

1. Libraries Required
    a. import numpy as np
    b. import pandas as pd
    c. import random
    d. pd.options.display.max_rows = None
    e. import seaborn as sns
    f. from sklearn.model_selection import train_test_split
    g. from sklearn.linear_model import LogisticRegression
    h. from datetime import datetime
    i. import matplotlib.pyplot as plt
    j. from warnings import filterwarnings
    k. filterwarnings('ignore')
    l. from sklearn.utils import resample
    m. from sklearn.metrics import accuracy_score,precision_score, recall_score, f1_score, roc_auc_score, roc_curve, confusion_matrix, ConfusionMatrixDisplay
    n. from imblearn.over_sampling import SMOTE

2. Data Preparation
    2.1 Read the Data

```
data = pd.read_csv("Rainfall_prediction_data.csv")
```

    2.2 Check the Data Type

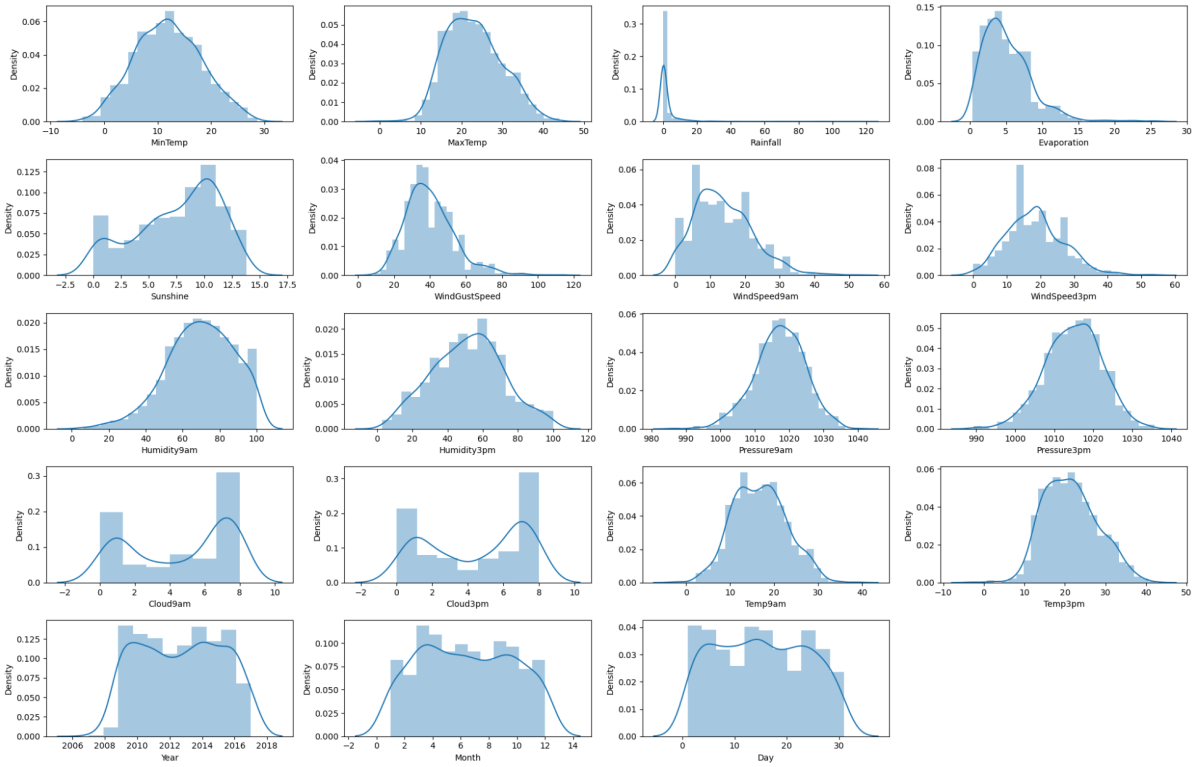
```
data.dtypes
```

    2.3 Remove Insignificant Variables, if its applicable.

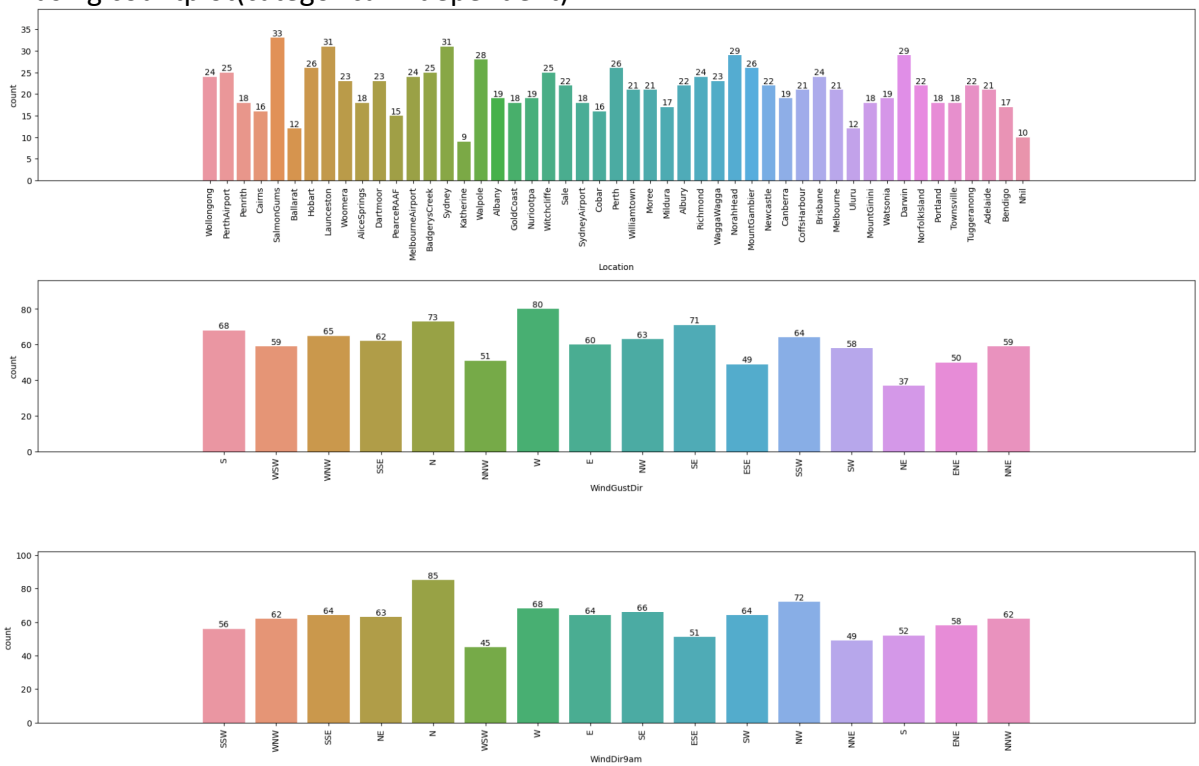
```
data.drop(['Date'],axis=1,inplace=True)
```

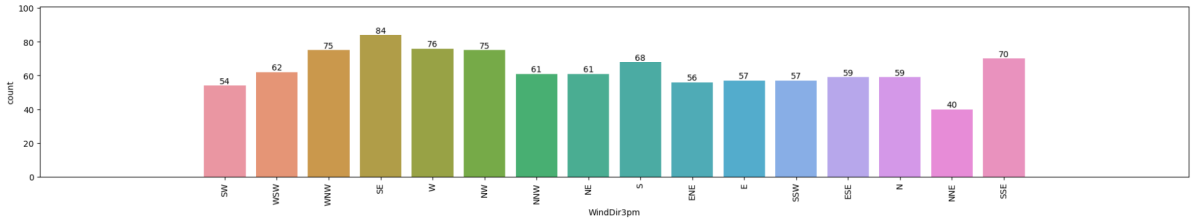
    2.4 Distribution of variable

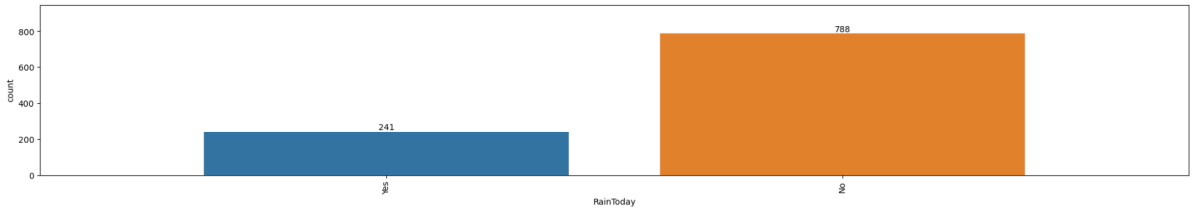
A. using DistPlot (numerical independent)



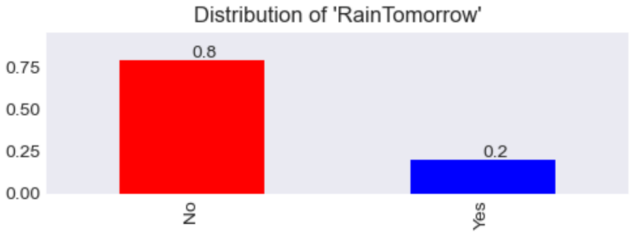
B. using countplot(categorical independent)





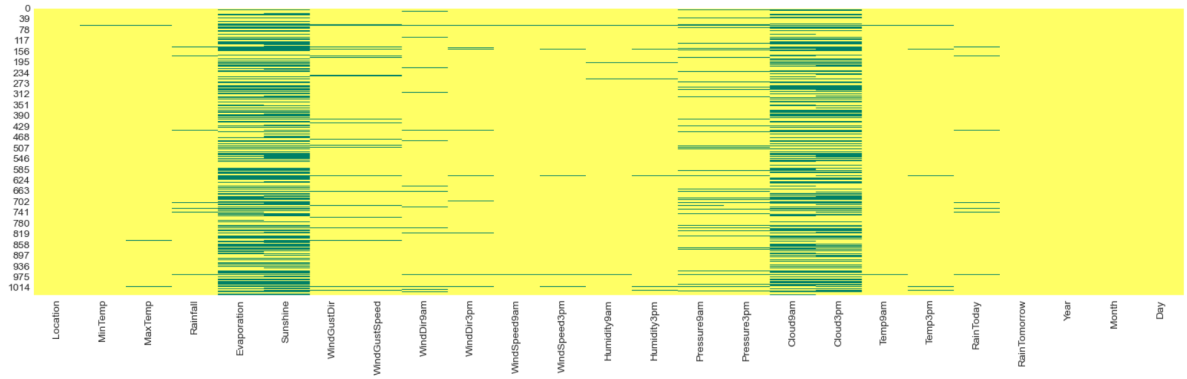


C. using histplot(Dependent)



2.5 Missing Value Treatment

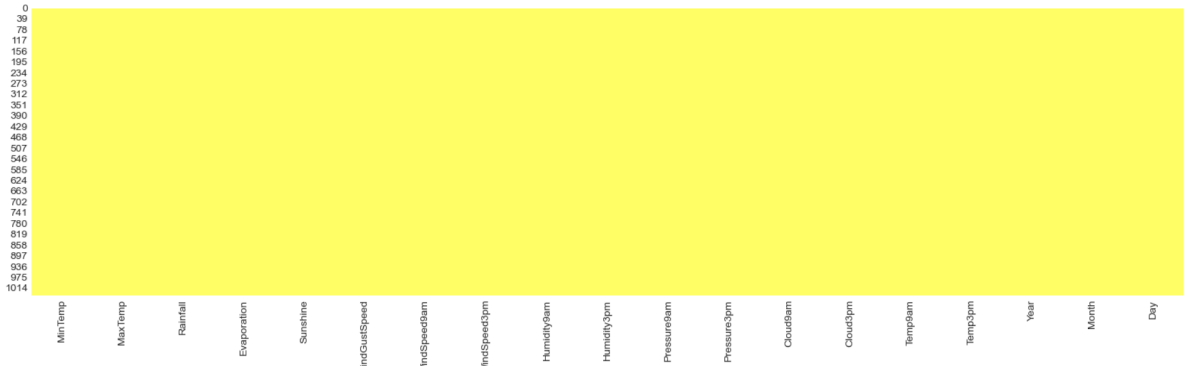
Before Treatment



a. Impute the Categorical Variables with mode

b. imputation of numerical variables using KNN Imputer

After Treatment

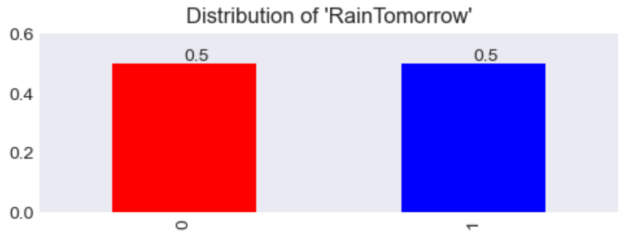


2.6 Dummy encoding the categorical Variables

a. Separating Dependent and Independent variables into separate datasets

b. **Encode the each categorical variable and create (n-1) dummy variables for n categories of the variable.**

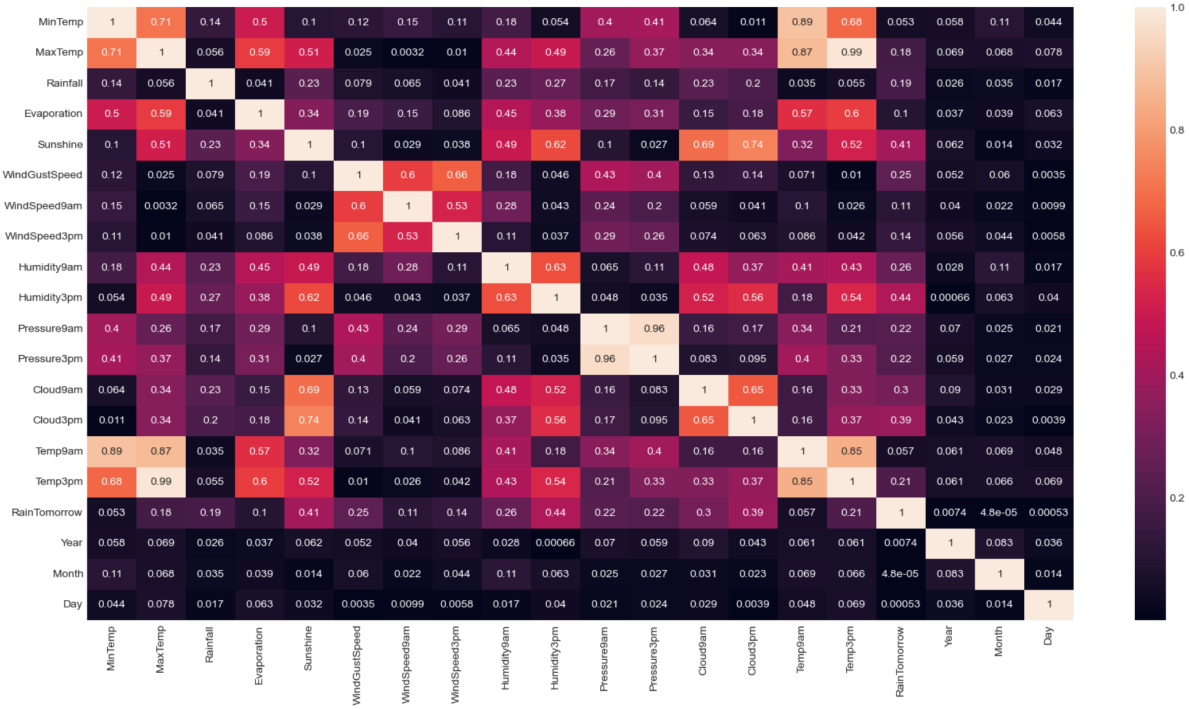
c. **Balancing of dependent variable**



2.**7 Scale the Data**

a. Using Standard scaler scaled the data

b. Checking the corelation between vaiables and removing the 'Day' from the dataset, as it is least corelated with dependent



2.8 Train-test split

a. the dataset is been splited with 80:20 ratio and with random_state=42

3. Logistic Regression

```
import statsmodels.api as sm
logreg = sm.Logit(np.array(ytrain),np.array(xtrain)).fit()
print(logreg.summary())
```

Warning: Maximum number of iterations has been exceeded.
Current function value: 0.348321
Iterations: 35

Logit Regression Results

==========
Dep. Variable: y No. Observations: 1326
Model: Logit Df Residuals: 1213
Method: MLE Df Model: 112
Date: Sun, 07 May 2023 Pseudo R-squ.: 0.4975
Time: 20:57:39 Log-Likelihood: -461.87
converged: False LL-Null: -919.11
Covariance Type: nonrobust LLR p-value: 4.821e-126
==========

	coef	std err	z	P>|z|	[0.025	0.975]
x1	0.4049	0.131	3.098	0.002	0.149	0.661
x2	-0.1507	0.356	-0.424	0.672	-0.848	0.546
x3	-0.3264	0.640	-0.510	0.610	-1.581	0.928
x4	-0.1603	0.105	-1.527	0.127	-0.366	0.045
x5	-0.0411	0.143	-0.287	0.774	-0.322	0.240
x6	-0.2941	0.194	-1.519	0.129	-0.674	0.085
x7	1.2655	0.190	6.674	0.000	0.894	1.637
x8	-0.2609	0.152	-1.713	0.087	-0.559	0.038
x9	0.0512	0.158	0.323	0.746	-0.259	0.361
x10	-0.0603	0.203	-0.297	0.767	-0.459	0.338
x11	2.2683	0.291	7.793	0.000	1.698	2.839
x12	1.5606	0.497	3.143	0.002	0.587	2.534
x13	-2.2727	0.505	-4.497	0.000	-3.263	-1.282
x14	-0.1775	0.169	-1.047	0.295	-0.510	0.155
x15	0.4780	0.173	2.767	0.006	0.139	0.817
x16	-0.9003	0.500	-1.801	0.072	-1.880	0.079
x17	1.1825	0.716	1.652	0.099	-0.221	2.586
x18	0.0463	0.099	0.468	0.640	-0.148	0.240
x19	0.0221	0.104	0.213	0.831	-0.181	0.225
x20	0.0004	0.000	1.320	0.187	-0.000	0.001
x21	0.6371	0.838	0.760	0.447	-1.006	2.280
x22	-1.2252	1.007	-1.216	0.224	-3.200	0.749
x23	2.6853	0.966	2.780	0.005	0.792	4.578
x24	1.0263	0.817	1.256	0.209	-0.576	2.628
x25	-0.5248	1.026	-0.512	0.609	-2.535	1.486
x26	-2.6556	1.156	-2.298	0.022	-4.921	-0.391
x27	1.8453	0.916	2.015	0.044	0.051	3.6
x28	2.4529	0.953	2.573	0.010	0.584	4.322
x29	-1.3607	1.137	-1.197	0.231	-3.589	0.867
x30	-26.7526	7.2e+05	-3.71e-05	1.000	-1.41e+06	1.41e+06
x31	-0.8837	0.968	-0.913	0.361	-2.781	1.013
x32	-0.1175	0.869	-0.135	0.892	-1.822	1.587
x33	-0.7692	0.922	-0.834	0.404	-2.576	1.038
x34	-1.1980	1.137	-1.054	0.292	-3.426	1.030

x35	-1.5167	0.903	-1.679	0.093	-3.287	0.253
x36	-0.0628	1.111	-0.057	0.955	-2.240	2.114
x37	-1.0554	0.859	-1.229	0.219	-2.739	0.628
x38	-0.5653	0.939	-0.602	0.547	-2.406	1.275
x39	0.3300	0.914	0.361	0.718	-1.461	2.121
x40	2.2425	0.876	2.559	0.010	0.525	3.960
x41	0.4876	0.937	0.521	0.603	-1.348	2.324
x42	-1.3408	0.907	-1.479	0.139	-3.118	0.436
x43	-0.2029	0.970	-0.209	0.834	-2.103	1.697
x44	-1.1687	1.044	-1.119	0.263	-3.216	0.878
x45	-1.3001	1.600	-0.813	0.416	-4.436	1.836
x46	-0.4915	0.869	-0.566	0.572	-2.194	1.211
x47	1.3011	0.801	1.624	0.104	-0.269	2.872
x48	1.5506	0.840	1.847	0.065	-0.095	3.196
x49	-1.8323	1.992	-0.920	0.358	-5.738	2.073
x50	1.2107	0.920	1.316	0.188	-0.593	3.014
x51	2.5072	0.800	3.136	0.002	0.940	4.074
x52	-0.9130	1.109	-0.823	0.410	-3.087	1.261
x53	-0.7286	0.924	-0.788	0.430	-2.540	1.083
x54	-0.0768	0.862	-0.089	0.929	-1.766	1.613
x55	-0.4695	0.890	-0.527	0.598	-2.214	1.275
x56	0.6405	0.844	0.759	0.448	-1.013	2.294
x57	0.5933	0.833	0.713	0.476	-1.039	2.225
x58	0.9895	0.956	1.035	0.301	-0.884	2.863
x59	-4.1224	2.133	-1.933	0.053	-8.302	0.057
x60	-1.3403	0.786	-1.705	0.088	-2.881	0.200
x61	0.9224	1.766	0.522	0.601	-2.539	4.383
x62	1.7723	0.802	2.209	0.027	0.200	3.345
x63	0.5159	0.799	0.646	0.518	-1.050	2.082
x64	0.1935	0.813	0.238	0.812	-1.400	1.787
x65	2.3950	0.918	2.608	0.009	0.595	4.195
x66	0.7526	0.806	0.934	0.351	-0.827	2.333
x67	-0.7743	0.913	-0.848	0.396	-2.564	1.015
x68	-0.6037	1.153	-0.523	0.601	-2.864	1.657
x69	-0.8581	0.679	-1.264	0.206	-2.189	0.472
x70	-0.2147	0.656	-0.327	0.743	-1.500	1.071
x71	0.1970	0.666	0.296	0.767	-1.108	1.502
x72	-0.4639	0.743	-0.624	0.533	-1.921	0.993
x73	-0.9122	0.663	-1.376	0.169	-2.212	0.387
x74	-0.2825	0.713	-0.396	0.692	-1.680	1.115
x75	-0.6084	0.743	-0.819	0.413	-2.065	0.848
x76	0.7848	0.640	1.227	0.220	-0.469	2.038
x77	0.7655	0.619	1.237	0.216	-0.448	1.979
x78	0.5312	0.683	0.778	0.436	-0.807	1.869
x79	-1.6236	0.777	-2.089	0.037	-3.147	-0.100
x80	-0.0190	0.691	-0.027	0.978	-1.373	1.335
x81	0.0744	0.628	0.118	0.906	-1.157	1.306

x82	0.4503	0.671	0.672	0.502	-0.864	1.765
x83	-0.8790	0.747	-1.177	0.239	-2.342	0.584
x84	0.2239	0.665	0.337	0.736	-1.079	1.527
x85	-0.9208	0.720	-1.279	0.201	-2.331	0.490
x86	0.3361	0.593	0.567	0.571	-0.826	1.498
x87	0.4091	0.633	0.646	0.518	-0.832	1.650
x88	-0.0009	0.693	-0.001	0.999	-1.359	1.357
x89	-0.0734	0.691	-0.106	0.915	-1.428	1.281
x90	0.5566	0.663	0.840	0.401	-0.743	1.856
x91	1.5638	0.685	2.282	0.023	0.220	2.907
x92	1.1755	0.623	1.886	0.059	-0.046	2.397
x93	0.1661	0.664	0.250	0.802	-1.135	1.467
x94	0.2549	0.701	0.363	0.716	-1.120	1.630
x95	0.4092	0.694	0.590	0.555	-0.951	1.769
x96	0.5999	0.681	0.880	0.379	-0.736	1.935
x97	0.2968	0.655	0.453	0.650	-0.987	1.581
x98	-0.2081	0.803	-0.259	0.795	-1.782	1.366
x99	0.0889	0.576	0.154	0.877	-1.039	1.217
x100	-0.3004	0.560	-0.537	0.592	-1.398	0.797
x101	-0.7291	0.662	-1.102	0.271	-2.026	0.568
x102	-0.0648	0.604	-0.107	0.915	-1.248	1.118
x103	-0.7074	0.706	-1.001	0.317	-2.092	0.677
x104	0.2237	0.676	0.331	0.741	-1.101	1.549
x105	0.3858	0.618	0.624	0.532	-0.825	1.597
x106	-1.2011	0.634	-1.893	0.058	-2.444	0.042
x107	-1.8833	0.575	-3.275	0.001	-3.010	-0.756
x108	-0.6979	0.601	-1.162	0.245	-1.875	0.480
x109	-0.2618	0.640	-0.409	0.683	-1.517	0.993
x110	-1.5697	0.676	-2.321	0.020	-2.895	-0.244
x111	-1.0074	0.641	-1.572	0.116	-2.263	0.248
x112	-2.1776	0.646	-3.372	0.001	-3.443	-0.912
x113	-0.9763	0.677	-1.443	0.149	-2.303	0.350

==========

Interpret the odds for each variable

1. RainToday = 0.4,it implies that the odds of getting rain tomorrow in creases by a factor of 0.4 due to one unit increase inRainToday, keepin g other variables constant

2. MinTemp = -0.15,it implies that the odds of getting rain tomorrow in creases by a factor of -0.15 due to one unit increase inMinTemp, keepin g other variables constant

3. MaxTemp = -0.33,it implies that the odds of getting rain tomorrow in creases by a factor of -0.33 due to one unit increase inMaxTemp, keepin g other variables constant

4. Rainfall = -0.16,it implies that the odds of getting rain tomorrow i ncreases by a factor of -0.16 due to one unit increase inRainfall, keep ing other variables constant

5. Evaporation = -0.04,it implies that the odds of getting rain tomorro w increases by a factor of -0.04 due to one unit increase inEvaporation , keeping other variables constant

6. Sunshine = -0.29,it implies that the odds of getting rain tomorrow i ncreases by a factor of -0.29 due to one unit increase inSunshine, keep ing other variables constant

7. WindGustSpeed = 1.27,it implies that the odds of getting rain tomorr ow increases by a factor of 1.27 due to one unit increase inWindGustSpe ed, keeping other variables constant

8. WindSpeed9am = -0.26,it implies that the odds of getting rain tomorr ow increases by a factor of -0.26 due to one unit increase inWindSpeed9 am, keeping other variables constant

9. WindSpeed3pm = 0.05,it implies that the odds of getting rain tomorro w increases by a factor of 0.05 due to one unit increase inWindSpeed3pm , keeping other variables constant

10. Humidity9am = -0.06,it implies that the odds of getting rain tomorr ow increases by a factor of -0.06 due to one unit increase inHumidity9a m, keeping other variables constant

11. Humidity3pm = 2.27,it implies that the odds of getting rain tomorro w increases by a factor of 2.27 due to one unit increase inHumidity3pm, keeping other variables constant

12. Pressure9am = 1.56,it implies that the odds of getting rain tomorro w increases by a factor of 1.56 due to one unit increase inPressure9am, keeping other variables constant

13. Pressure3pm = -2.27,it implies that the odds of getting rain tomorr ow increases by a factor of -2.27 due to one unit increase inPressure3p m, keeping other variables constant

14. Cloud9am = -0.18,it implies that the odds of getting rain tomorrow increases by a factor of -0.18 due to one unit increase inCloud9am, kee ping other variables constant

15. Cloud3pm = 0.48,it implies that the odds of getting rain tomorrow i ncreases by a factor of 0.48 due to one unit increase inCloud3pm, keepi ng other variables constant

16. Temp9am = -0.9,it implies that the odds of getting rain tomorrow in creases by a factor of -0.9 due to one unit increase inTemp9am, keeping other variables constant

17. Temp3pm = 1.18,it implies that the odds of getting rain tomorrow in creases by a factor of 1.18 due to one unit increase inTemp3pm, keeping other variables constant

18. Year = 0.05,it implies that the odds of getting rain tomorrow increases by a factor of 0.05 due to one unit increase inYear, keeping other variables constant

19. Month = 0.02,it implies that the odds of getting rain tomorrow increases by a factor of 0.02 due to one unit increase inMonth, keeping other variables constant

20. index = 0.0,it implies that the odds of getting rain tomorrow increases by a factor of 0.0 due to one unit increase inindex, keeping other variables constant

21. Location_Albany = 0.64,it implies that the odds of getting rain tomorrow increases by a factor of 0.64 due to one unit increase inLocation_Albany, keeping other variables constant

22. Location_Albury = -1.23,it implies that the odds of getting rain tomorrow increases by a factor of -1.23 due to one unit increase inLocation_Albury, keeping other variables constant

23. Location_AliceSprings = 2.69,it implies that the odds of getting rain tomorrow increases by a factor of 2.69 due to one unit increase inLocation_AliceSprings, keeping other variables constant

24. Location_BadgerysCreek = 1.03,it implies that the odds of getting rain tomorrow increases by a factor of 1.03 due to one unit increase inLocation_BadgerysCreek, keeping other variables constant

25. Location_Ballarat = -0.52,it implies that the odds of getting rain tomorrow increases by a factor of -0.52 due to one unit increase inLocation_Ballarat, keeping other variables constant

26. Location_Bendigo = -2.66,it implies that the odds of getting rain tomorrow increases by a factor of -2.66 due to one unit increase inLocation_Bendigo, keeping other variables constant

27. Location_Brisbane = 1.85,it implies that the odds of getting rain tomorrow increases by a factor of 1.85 due to one unit increase inLocation_Brisbane, keeping other variables constant

28. Location_Cairns = 2.45,it implies that the odds of getting rain tomorrow increases by a factor of 2.45 due to one unit increase inLocation_Cairns, keeping other variables constant

29. Location_Canberra = -1.36,it implies that the odds of getting rain tomorrow increases by a factor of -1.36 due to one unit increase inLocation_Canberra, keeping other variables constant

30. Location_Cobar = -26.75,it implies that the odds of getting rain tomorrow increases by a factor of -26.75 due to one unit increase inLocation_Cobar, keeping other variables constant

31. Location_CoffsHarbour = -0.88,it implies that the odds of getting rain tomorrow increases by a factor of -0.88 due to one unit increase in Location_CoffsHarbour, keeping other variables constant

32. Location_Dartmoor = -0.12,it implies that the odds of getting rain tomorrow increases by a factor of -0.12 due to one unit increase inLocation_Dartmoor, keeping other variables constant

33. Location_Darwin = -0.77,it implies that the odds of getting rain tomorrow increases by a factor of -0.77 due to one unit increase inLocation_Darwin, keeping other variables constant

34. Location_GoldCoast = -1.2,it implies that the odds of getting rain tomorrow increases by a factor of -1.2 due to one unit increase inLocation_GoldCoast, keeping other variables constant

35. Location_Hobart = -1.52,it implies that the odds of getting rain tomorrow increases by a factor of -1.52 due to one unit increase inLocation_Hobart, keeping other variables constant

36. Location_Katherine = -0.06,it implies that the odds of getting rain tomorrow increases by a factor of -0.06 due to one unit increase inLocation_Katherine, keeping other variables constant

37. Location_Launceston = -1.06,it implies that the odds of getting rain tomorrow increases by a factor of -1.06 due to one unit increase inLocation_Launceston, keeping other variables constant

38. Location_Melbourne = -0.57,it implies that the odds of getting rain tomorrow increases by a factor of -0.57 due to one unit increase inLocation_Melbourne, keeping other variables constant

39. Location_MelbourneAirport = 0.33,it implies that the odds of getting rain tomorrow increases by a factor of 0.33 due to one unit increase inLocation_MelbourneAirport, keeping other variables constant

40. Location_Mildura = 2.24,it implies that the odds of getting rain tomorrow increases by a factor of 2.24 due to one unit increase inLocation_Mildura, keeping other variables constant

41. Location_Moree = 0.49,it implies that the odds of getting rain tomorrow increases by a factor of 0.49 due to one unit increase inLocation_Moree, keeping other variables constant

42. Location_MountGambier = -1.34,it implies that the odds of getting rain tomorrow increases by a factor of -1.34 due to one unit increase in Location_MountGambier, keeping other variables constant

43. Location_MountGinini = -0.2,it implies that the odds of getting rain tomorrow increases by a factor of -0.2 due to one unit increase inLocation_MountGinini, keeping other variables constant

44. Location_Newcastle = -1.17,it implies that the odds of getting rain tomorrow increases by a factor of -1.17 due to one unit increase inLocation_Newcastle, keeping other variables constant

45. Location_Nhil = -1.3,it implies that the odds of getting rain tomorrow increases by a factor of -1.3 due to one unit increase inLocation_Nhil, keeping other variables constant

46. Location_NorahHead = -0.49,it implies that the odds of getting rain tomorrow increases by a factor of -0.49 due to one unit increase inLocation_NorahHead, keeping other variables constant

47. Location_NorfolkIsland = 1.3,it implies that the odds of getting rain tomorrow increases by a factor of 1.3 due to one unit increase inLocation_NorfolkIsland, keeping other variables constant

48. Location_Nuriootpa = 1.55,it implies that the odds of getting rain tomorrow increases by a factor of 1.55 due to one unit increase inLocation_Nuriootpa, keeping other variables constant

49. Location_PearceRAAF = -1.83,it implies that the odds of getting rain tomorrow increases by a factor of -1.83 due to one unit increase inLocation_PearceRAAF, keeping other variables constant

50. Location_Penrith = 1.21,it implies that the odds of getting rain tomorrow increases by a factor of 1.21 due to one unit increase inLocation_Penrith, keeping other variables constant

51. Location_Perth = 2.51,it implies that the odds of getting rain tomorrow increases by a factor of 2.51 due to one unit increase inLocation_Perth, keeping other variables constant

52. Location_PerthAirport = -0.91,it implies that the odds of getting rain tomorrow increases by a factor of -0.91 due to one unit increase in Location_PerthAirport, keeping other variables constant

53. Location_Portland = -0.73,it implies that the odds of getting rain tomorrow increases by a factor of -0.73 due to one unit increase inLocation_Portland, keeping other variables constant

54. Location_Richmond = -0.08,it implies that the odds of getting rain tomorrow increases by a factor of -0.08 due to one unit increase inLocation_Richmond, keeping other variables constant

55. Location_Sale = -0.47,it implies that the odds of getting rain tomorrow increases by a factor of -0.47 due to one unit increase inLocation_Sale, keeping other variables constant

56. Location_SalmonGums = 0.64,it implies that the odds of getting rain tomorrow increases by a factor of 0.64 due to one unit increase inLocation_SalmonGums, keeping other variables constant

57. Location_Sydney = 0.59,it implies that the odds of getting rain tomorrow increases by a factor of 0.59 due to one unit increase inLocation_Sydney, keeping other variables constant

58. Location_SydneyAirport = 0.99,it implies that the odds of getting rain tomorrow increases by a factor of 0.99 due to one unit increase inLocation_SydneyAirport, keeping other variables constant

59. Location_Townsville = -4.12,it implies that the odds of getting rain tomorrow increases by a factor of -4.12 due to one unit increase inLocation_Townsville, keeping other variables constant

60. Location_Tuggeranong = -1.34,it implies that the odds of getting ra in tomorrow increases by a factor of -1.34 due to one unit increase inL ocation_Tuggeranong, keeping other variables constant

61. Location_Uluru = 0.92,it implies that the odds of getting rain tomo rrow increases by a factor of 0.92 due to one unit increase inLocation_ Uluru, keeping other variables constant

62. Location_WaggaWagga = 1.77,it implies that the odds of getting rain tomorrow increases by a factor of 1.77 due to one unit increase inLocat ion_WaggaWagga, keeping other variables constant

63. Location_Walpole = 0.52,it implies that the odds of getting rain to morrow increases by a factor of 0.52 due to one unit increase inLocatio n_Walpole, keeping other variables constant

64. Location_Watsonia = 0.19,it implies that the odds of getting rain t omorrow increases by a factor of 0.19 due to one unit increase inLocati on_Watsonia, keeping other variables constant

65. Location_Williamtown = 2.39,it implies that the odds of getting rai n tomorrow increases by a factor of 2.39 due to one unit increase inLoc ation_Williamtown, keeping other variables constant

66. Location_Witchcliffe = 0.75,it implies that the odds of getting rai n tomorrow increases by a factor of 0.75 due to one unit increase inLoc ation_Witchcliffe, keeping other variables constant

67. Location_Wollongong = -0.77,it implies that the odds of getting rai n tomorrow increases by a factor of -0.77 due to one unit increase inLo cation_Wollongong, keeping other variables constant

68. Location_Woomera = -0.6,it implies that the odds of getting rain to morrow increases by a factor of -0.6 due to one unit increase inLocatio n_Woomera, keeping other variables constant

69. WindGustDir_ENE = -0.86,it implies that the odds of getting rain to morrow increases by a factor of -0.86 due to one unit increase inWindGu stDir_ENE, keeping other variables constant

70. WindGustDir_ESE = -0.21,it implies that the odds of getting rain to morrow increases by a factor of -0.21 due to one unit increase inWindGu stDir_ESE, keeping other variables constant

71. WindGustDir_N = 0.2,it implies that the odds of getting rain tomorr ow increases by a factor of 0.2 due to one unit increase inWindGustDir_ N, keeping other variables constant

72. WindGustDir_NE = -0.46,it implies that the odds of getting rain tom orrow increases by a factor of -0.46 due to one unit increase inWindGus tDir_NE, keeping other variables constant

73. WindGustDir_NNE = -0.91,it implies that the odds of getting rain to morrow increases by a factor of -0.91 due to one unit increase inWindGu stDir_NNE, keeping other variables constant

74. WindGustDir_NNW = -0.28,it implies that the odds of getting rain to morrow increases by a factor of -0.28 due to one unit increase inWindGustDir_NNW, keeping other variables constant

75. WindGustDir_NW = -0.61,it implies that the odds of getting rain tomorrow increases by a factor of -0.61 due to one unit increase inWindGustDir_NW, keeping other variables constant

76. WindGustDir_S = 0.78,it implies that the odds of getting rain tomorrow increases by a factor of 0.78 due to one unit increase inWindGustDir_S, keeping other variables constant

77. WindGustDir_SE = 0.77,it implies that the odds of getting rain tomorrow increases by a factor of 0.77 due to one unit increase inWindGustDir_SE, keeping other variables constant

78. WindGustDir_SSE = 0.53,it implies that the odds of getting rain tomorrow increases by a factor of 0.53 due to one unit increase inWindGustDir_SSE, keeping other variables constant

79. WindGustDir_SSW = -1.62,it implies that the odds of getting rain to morrow increases by a factor of -1.62 due to one unit increase inWindGustDir_SSW, keeping other variables constant

80. WindGustDir_SW = -0.02,it implies that the odds of getting rain tomorrow increases by a factor of -0.02 due to one unit increase inWindGustDir_SW, keeping other variables constant

81. WindGustDir_W = 0.07,it implies that the odds of getting rain tomorrow increases by a factor of 0.07 due to one unit increase inWindGustDir_W, keeping other variables constant

82. WindGustDir_WNW = 0.45,it implies that the odds of getting rain tomorrow increases by a factor of 0.45 due to one unit increase inWindGustDir_WNW, keeping other variables constant

83. WindGustDir_WSW = -0.88,it implies that the odds of getting rain to morrow increases by a factor of -0.88 due to one unit increase inWindGustDir_WSW, keeping other variables constant

84. WindDir9am_ENE = 0.22,it implies that the odds of getting rain tomorrow increases by a factor of 0.22 due to one unit increase inWindDir9am_ENE, keeping other variables constant

85. WindDir9am_ESE = -0.92,it implies that the odds of getting rain tomorrow increases by a factor of -0.92 due to one unit increase inWindDir9am_ESE, keeping other variables constant

86. WindDir9am_N = 0.34,it implies that the odds of getting rain tomorrow increases by a factor of 0.34 due to one unit increase inWindDir9am_N, keeping other variables constant

87. WindDir9am_NE = 0.41,it implies that the odds of getting rain tomorrow increases by a factor of 0.41 due to one unit increase inWindDir9am_NE, keeping other variables constant

88. WindDir9am_NNE = -0.0,it implies that the odds of getting rain tomo rrow increases by a factor of -0.0 due to one unit increase inWindDir9a m_NNE, keeping other variables constant

89. WindDir9am_NNW = -0.07,it implies that the odds of getting rain tom orrow increases by a factor of -0.07 due to one unit increase inWindDir 9am_NNW, keeping other variables constant

90. WindDir9am_NW = 0.56,it implies that the odds of getting rain tomor row increases by a factor of 0.56 due to one unit increase inWindDir9am _NW, keeping other variables constant

91. WindDir9am_S = 1.56,it implies that the odds of getting rain tomorr ow increases by a factor of 1.56 due to one unit increase inWindDir9am_ S, keeping other variables constant

92. WindDir9am_SE = 1.18,it implies that the odds of getting rain tomor row increases by a factor of 1.18 due to one unit increase inWindDir9am _SE, keeping other variables constant

93. WindDir9am_SSE = 0.17,it implies that the odds of getting rain tomo rrow increases by a factor of 0.17 due to one unit increase inWindDir9a m_SSE, keeping other variables constant

94. WindDir9am_SSW = 0.25,it implies that the odds of getting rain tomo rrow increases by a factor of 0.25 due to one unit increase inWindDir9a m_SSW, keeping other variables constant

95. WindDir9am_SW = 0.41,it implies that the odds of getting rain tomor row increases by a factor of 0.41 due to one unit increase inWindDir9am _SW, keeping other variables constant

96. WindDir9am_W = 0.6,it implies that the odds of getting rain tomorro w increases by a factor of 0.6 due to one unit increase inWindDir9am_W, keeping other variables constant

97. WindDir9am_WNW = 0.3,it implies that the odds of getting rain tomor row increases by a factor of 0.3 due to one unit increase inWindDir9am_ WNW, keeping other variables constant

98. WindDir9am_WSW = -0.21,it implies that the odds of getting rain tom orrow increases by a factor of -0.21 due to one unit increase inWindDir 9am_WSW, keeping other variables constant

99. WindDir3pm_ENE = 0.09,it implies that the odds of getting rain tomo rrow increases by a factor of 0.09 due to one unit increase inWindDir3p m_ENE, keeping other variables constant

100. WindDir3pm_ESE = -0.3,it implies that the odds of getting rain tom orrow increases by a factor of -0.3 due to one unit increase inWindDir3 pm_ESE, keeping other variables constant

101. WindDir3pm_N = -0.73,it implies that the odds of getting rain tomo rrow increases by a factor of -0.73 due to one unit increase inWindDir3 pm_N, keeping other variables constant

102. WindDir3pm_NE = -0.06,it implies that the odds of getting rain tomorrow increases by a factor of -0.06 due to one unit increase inWindDir3pm_NE, keeping other variables constant

103. WindDir3pm_NNE = -0.71,it implies that the odds of getting rain tomorrow increases by a factor of -0.71 due to one unit increase inWindDir3pm_NNE, keeping other variables constant

104. WindDir3pm_NNW = 0.22,it implies that the odds of getting rain tomorrow increases by a factor of 0.22 due to one unit increase inWindDir3pm_NNW, keeping other variables constant

105. WindDir3pm_NW = 0.39,it implies that the odds of getting rain tomorrow increases by a factor of 0.39 due to one unit increase inWindDir3pm_NW, keeping other variables constant

106. WindDir3pm_S = -1.2,it implies that the odds of getting rain tomorrow increases by a factor of -1.2 due to one unit increase inWindDir3pm_S, keeping other variables constant

107. WindDir3pm_SE = -1.88,it implies that the odds of getting rain tomorrow increases by a factor of -1.88 due to one unit increase inWindDir3pm_SE, keeping other variables constant

108. WindDir3pm_SSE = -0.7,it implies that the odds of getting rain tomorrow increases by a factor of -0.7 due to one unit increase inWindDir3pm_SSE, keeping other variables constant

109. WindDir3pm_SSW = -0.26,it implies that the odds of getting rain tomorrow increases by a factor of -0.26 due to one unit increase inWindDir3pm_SSW, keeping other variables constant

110. WindDir3pm_SW = -1.57,it implies that the odds of getting rain tomorrow increases by a factor of -1.57 due to one unit increase inWindDir3pm_SW, keeping other variables constant

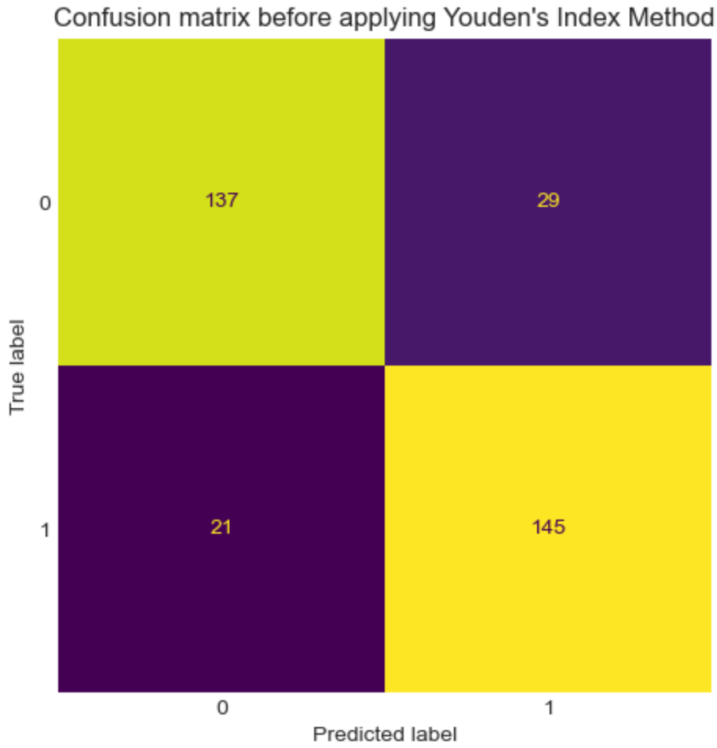
111. WindDir3pm_W = -1.01,it implies that the odds of getting rain tomorrow increases by a factor of -1.01 due to one unit increase inWindDir3pm_W, keeping other variables constant

112. WindDir3pm_WNW = -2.18,it implies that the odds of getting rain tomorrow increases by a factor of -2.18 due to one unit increase inWindDir3pm_WNW, keeping other variables constant

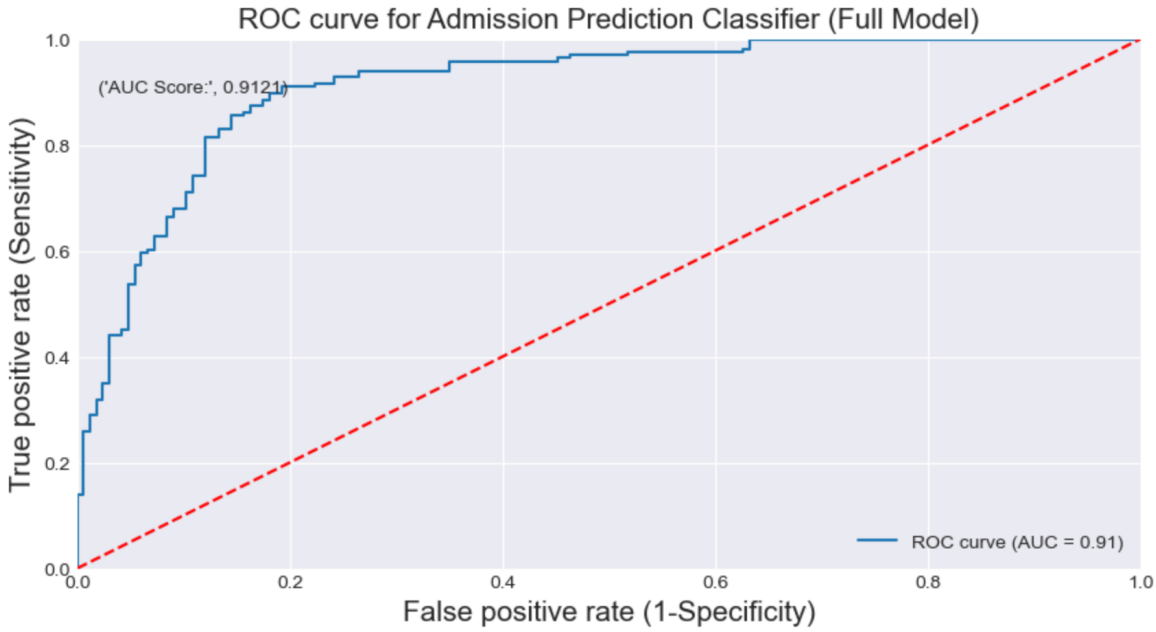
113. WindDir3pm_WSW = -0.98,it implies that the odds of getting rain tomorrow increases by a factor of -0.98 due to one unit increase inWindDir3pm_WSW, keeping other variables constant

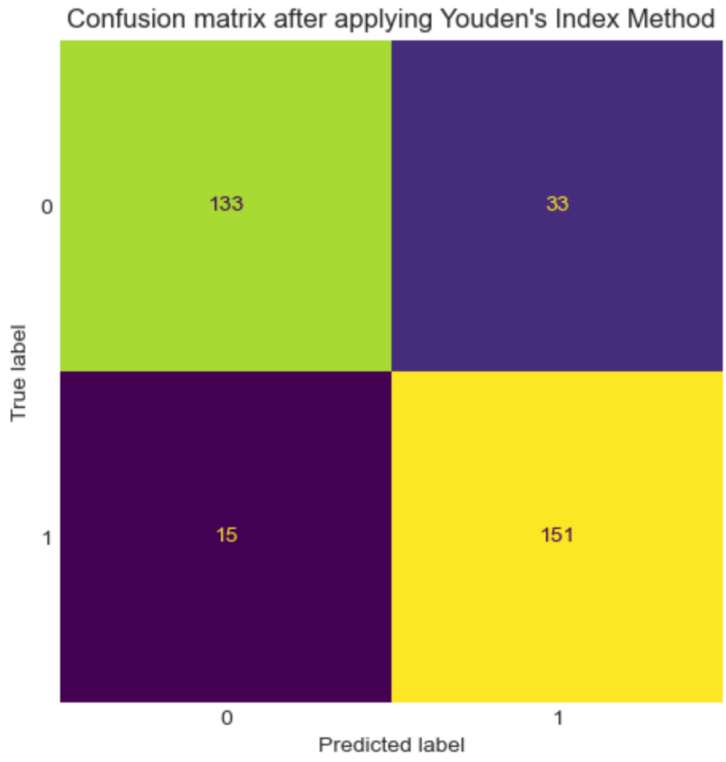
Compute various performance metrics.

Precision: 0.83
Recall: 0.87
Specificity: 0.83
F1_Score: 0.85
Accuracy: 0.85
Kappa Score: 0.7
AUC: 0.91



ROC_AUC score





3.1 Identify the Best Cut-off Value

Interpretation: From the above output, write your inference.

- The model achieved an overall accuracy of 0.85, which means that it correctly predicted the class labels of 85% of the instances in the dataset.
- Looking at the precision and recall values, we can see that the model performed similarly for both classes.
- For class 0, the precision is 0.89 and the recall is 0.80, meaning that 89% of the instances classified as 0 were actually 0, and the model correctly identified 80% of all instances that were actually 0.
- For class 1, the precision is 0.82 and the recall is 0.90, meaning that 82% of the instances classified as 1 were actually 1, and the model correctly identified 90% of all instances that were actually 1.
- The f1-score gives an overall measure of the model's performance.The f1-score for class 0 is 0.84, and for class 1 is 0.86, indicating that the model has a similar overall performance for both classes

Using the threshold value of 0.44 calculated by Youden's Index method, Class 1 F1_Score of the model increased from 82% to 86% and Class 0 F1_Score of the model increased from 82% to 84%