untitled2

March 28, 2025

[1]:
```
#NAME: SHIVANI GADKARI
#ROLL NO.: 13342
```

[4]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import classification_report, accuracy_score
```

[5]:
```
df = pd.read_csv('diabetes.csv')
```

[6]:
```
df
```

[6]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	\
0	6	148	72	35	0	33.6	
1	1	85	66	29	0	26.6	
2	8	183	64	0	0	23.3	
3	1	89	66	23	94	28.1	
4	0	137	40	35	168	43.1	
..	…	…	…	…	…	…	
763	10	101	76	48	180	32.9	
764	2	122	70	27	0	36.8	
765	5	121	72	23	112	26.2	
766	1	126	60	0	0	30.1	
767	1	93	70	31	0	30.4	

	DiabetesPedigreeFunction	Age	Outcome
0	0.627	50	1
1	0.351	31	0
2	0.672	32	1
3	0.167	21	0
4	2.288	33	1
..	…	…	…
763	0.171	63	0
764	0.340	27	0
765	0.245	30	0
766	0.349	47	1

```
767                     0.315   23        0

[768 rows x 9 columns]
```

[7]:
```
df.shape
```

[7]:
```
(768, 9)
```

[8]:
```
df.head
```

[8]:
```
<bound method NDFrame.head of      Pregnancies  Glucose  BloodPressure
SkinThickness  Insulin   BMI  \
0              6      148             72             35        0  33.6
1              1       85             66             29        0  26.6
2              8      183             64              0        0  23.3
3              1       89             66             23       94  28.1
4              0      137             40             35      168  43.1
..           …      …              …              …       …   …
763           10      101             76             48      180  32.9
764            2      122             70             27        0  36.8
765            5      121             72             23      112  26.2
766            1      126             60              0        0  30.1
767            1       93             70             31        0  30.4

     DiabetesPedigreeFunction  Age  Outcome
0                       0.627   50        1
1                       0.351   31        0
2                       0.672   32        1
3                       0.167   21        0
4                       2.288   33        1
..                        …  …      …
763                     0.171   63        0
764                     0.340   27        0
765                     0.245   30        0
766                     0.349   47        1
767                     0.315   23        0

[768 rows x 9 columns]>
```

[10]:
```
df.isnull().sum()
```

[10]:
```
Pregnancies                 0
Glucose                     0
BloodPressure               0
SkinThickness               0
Insulin                     0
BMI                         0
```

```
DiabetesPedigreeFunction    0
Age                         0
Outcome                     0
dtype: int64
```

[12]:
```
X = df.drop(columns=['Outcome'])
y = df['Outcome']
```

[13]:
```
#Defining X_train,X_test ,etc
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
 ↪random_state=42)
```

[14]:
```
#Model is trained using GausianNB:
from sklearn.naive_bayes import GaussianNB
gaussian = GaussianNB()
gaussian.fit(X_train, y_train)
```

[14]:
```
GaussianNB()
```

[23]:
```
#Making predictions:
y_pred_all = model.predict(X_test_scaled)
probabilities_all = model.predict_proba(X_test_scaled)
```

[24]:
```
#Printing prediction and probablities:
print("\nPredictions for all test samples:")
for i, (prediction, prob) in enumerate(zip(y_pred_all, probabilities_all)):#␣
 ↪all predicticts probablites for each sample
    print(f"Sample {i+1}: Predicted Outcome = {prediction}, P(No) = {prob[0]:.
 ↪4f}, P(Yes) = {prob[1]:.4f}")
```

```
Predictions for all test samples:
Sample 1: Predicted Outcome = 0, P(No) = 0.7226, P(Yes) = 0.2774
Sample 2: Predicted Outcome = 0, P(No) = 0.8140, P(Yes) = 0.1860
Sample 3: Predicted Outcome = 0, P(No) = 0.8853, P(Yes) = 0.1147
Sample 4: Predicted Outcome = 0, P(No) = 0.8286, P(Yes) = 0.1714
Sample 5: Predicted Outcome = 0, P(No) = 0.5316, P(Yes) = 0.4684
Sample 6: Predicted Outcome = 0, P(No) = 0.5601, P(Yes) = 0.4399
Sample 7: Predicted Outcome = 0, P(No) = 0.9838, P(Yes) = 0.0162
Sample 8: Predicted Outcome = 1, P(No) = 0.3404, P(Yes) = 0.6596
Sample 9: Predicted Outcome = 1, P(No) = 0.4638, P(Yes) = 0.5362
Sample 10: Predicted Outcome = 1, P(No) = 0.2294, P(Yes) = 0.7706
Sample 11: Predicted Outcome = 0, P(No) = 0.7461, P(Yes) = 0.2539
Sample 12: Predicted Outcome = 1, P(No) = 0.1086, P(Yes) = 0.8914
Sample 13: Predicted Outcome = 0, P(No) = 0.6671, P(Yes) = 0.3329
Sample 14: Predicted Outcome = 0, P(No) = 0.6871, P(Yes) = 0.3129
```

```
Sample 15: Predicted Outcome = 0, P(No) = 0.9147, P(Yes) = 0.0853
Sample 16: Predicted Outcome = 0, P(No) = 0.6137, P(Yes) = 0.3863
Sample 17: Predicted Outcome = 0, P(No) = 0.8606, P(Yes) = 0.1394
Sample 18: Predicted Outcome = 0, P(No) = 0.9235, P(Yes) = 0.0765
Sample 19: Predicted Outcome = 1, P(No) = 0.1422, P(Yes) = 0.8578
Sample 20: Predicted Outcome = 1, P(No) = 0.4448, P(Yes) = 0.5552
Sample 21: Predicted Outcome = 0, P(No) = 0.7916, P(Yes) = 0.2084
Sample 22: Predicted Outcome = 0, P(No) = 0.9209, P(Yes) = 0.0791
Sample 23: Predicted Outcome = 1, P(No) = 0.4688, P(Yes) = 0.5312
Sample 24: Predicted Outcome = 0, P(No) = 0.9046, P(Yes) = 0.0954
Sample 25: Predicted Outcome = 1, P(No) = 0.4650, P(Yes) = 0.5350
Sample 26: Predicted Outcome = 1, P(No) = 0.1201, P(Yes) = 0.8799
Sample 27: Predicted Outcome = 0, P(No) = 0.8755, P(Yes) = 0.1245
Sample 28: Predicted Outcome = 0, P(No) = 0.9689, P(Yes) = 0.0311
Sample 29: Predicted Outcome = 0, P(No) = 0.7520, P(Yes) = 0.2480
Sample 30: Predicted Outcome = 0, P(No) = 0.8805, P(Yes) = 0.1195
Sample 31: Predicted Outcome = 1, P(No) = 0.0928, P(Yes) = 0.9072
Sample 32: Predicted Outcome = 1, P(No) = 0.1327, P(Yes) = 0.8673
Sample 33: Predicted Outcome = 1, P(No) = 0.2423, P(Yes) = 0.7577
Sample 34: Predicted Outcome = 1, P(No) = 0.1670, P(Yes) = 0.8330
Sample 35: Predicted Outcome = 1, P(No) = 0.3795, P(Yes) = 0.6205
Sample 36: Predicted Outcome = 1, P(No) = 0.3141, P(Yes) = 0.6859
Sample 37: Predicted Outcome = 1, P(No) = 0.0346, P(Yes) = 0.9654
Sample 38: Predicted Outcome = 0, P(No) = 0.7502, P(Yes) = 0.2498
Sample 39: Predicted Outcome = 1, P(No) = 0.4913, P(Yes) = 0.5087
Sample 40: Predicted Outcome = 1, P(No) = 0.2764, P(Yes) = 0.7236
Sample 41: Predicted Outcome = 0, P(No) = 0.9288, P(Yes) = 0.0712
Sample 42: Predicted Outcome = 1, P(No) = 0.4057, P(Yes) = 0.5943
Sample 43: Predicted Outcome = 1, P(No) = 0.4182, P(Yes) = 0.5818
Sample 44: Predicted Outcome = 0, P(No) = 0.6694, P(Yes) = 0.3306
Sample 45: Predicted Outcome = 0, P(No) = 0.9712, P(Yes) = 0.0288
Sample 46: Predicted Outcome = 1, P(No) = 0.4930, P(Yes) = 0.5070
Sample 47: Predicted Outcome = 1, P(No) = 0.3498, P(Yes) = 0.6502
Sample 48: Predicted Outcome = 0, P(No) = 0.7749, P(Yes) = 0.2251
Sample 49: Predicted Outcome = 0, P(No) = 0.6303, P(Yes) = 0.3697
Sample 50: Predicted Outcome = 1, P(No) = 0.0442, P(Yes) = 0.9558
Sample 51: Predicted Outcome = 0, P(No) = 0.9504, P(Yes) = 0.0496
Sample 52: Predicted Outcome = 1, P(No) = 0.3362, P(Yes) = 0.6638
Sample 53: Predicted Outcome = 1, P(No) = 0.1903, P(Yes) = 0.8097
Sample 54: Predicted Outcome = 0, P(No) = 0.7519, P(Yes) = 0.2481
Sample 55: Predicted Outcome = 0, P(No) = 0.9057, P(Yes) = 0.0943
Sample 56: Predicted Outcome = 0, P(No) = 0.9573, P(Yes) = 0.0427
Sample 57: Predicted Outcome = 1, P(No) = 0.2290, P(Yes) = 0.7710
Sample 58: Predicted Outcome = 0, P(No) = 0.9940, P(Yes) = 0.0060
Sample 59: Predicted Outcome = 0, P(No) = 0.5870, P(Yes) = 0.4130
Sample 60: Predicted Outcome = 1, P(No) = 0.2525, P(Yes) = 0.7475
Sample 61: Predicted Outcome = 1, P(No) = 0.2643, P(Yes) = 0.7357
Sample 62: Predicted Outcome = 0, P(No) = 0.6519, P(Yes) = 0.3481
```

```
Sample 63: Predicted Outcome = 0, P(No) = 0.7994, P(Yes) = 0.2006
Sample 64: Predicted Outcome = 0, P(No) = 0.7839, P(Yes) = 0.2161
Sample 65: Predicted Outcome = 0, P(No) = 0.9221, P(Yes) = 0.0779
Sample 66: Predicted Outcome = 1, P(No) = 0.3828, P(Yes) = 0.6172
Sample 67: Predicted Outcome = 0, P(No) = 0.9480, P(Yes) = 0.0520
Sample 68: Predicted Outcome = 1, P(No) = 0.2740, P(Yes) = 0.7260
Sample 69: Predicted Outcome = 0, P(No) = 0.9610, P(Yes) = 0.0390
Sample 70: Predicted Outcome = 1, P(No) = 0.2848, P(Yes) = 0.7152
Sample 71: Predicted Outcome = 1, P(No) = 0.3241, P(Yes) = 0.6759
Sample 72: Predicted Outcome = 0, P(No) = 0.9280, P(Yes) = 0.0720
Sample 73: Predicted Outcome = 0, P(No) = 0.8171, P(Yes) = 0.1829
Sample 74: Predicted Outcome = 0, P(No) = 0.8851, P(Yes) = 0.1149
Sample 75: Predicted Outcome = 0, P(No) = 0.9072, P(Yes) = 0.0928
Sample 76: Predicted Outcome = 1, P(No) = 0.4851, P(Yes) = 0.5149
Sample 77: Predicted Outcome = 0, P(No) = 0.8318, P(Yes) = 0.1682
Sample 78: Predicted Outcome = 0, P(No) = 0.8567, P(Yes) = 0.1433
Sample 79: Predicted Outcome = 0, P(No) = 0.8659, P(Yes) = 0.1341
Sample 80: Predicted Outcome = 0, P(No) = 0.7662, P(Yes) = 0.2338
Sample 81: Predicted Outcome = 1, P(No) = 0.3496, P(Yes) = 0.6504
Sample 82: Predicted Outcome = 0, P(No) = 0.8517, P(Yes) = 0.1483
Sample 83: Predicted Outcome = 0, P(No) = 0.9369, P(Yes) = 0.0631
Sample 84: Predicted Outcome = 0, P(No) = 0.6293, P(Yes) = 0.3707
Sample 85: Predicted Outcome = 0, P(No) = 0.7380, P(Yes) = 0.2620
Sample 86: Predicted Outcome = 1, P(No) = 0.1697, P(Yes) = 0.8303
Sample 87: Predicted Outcome = 1, P(No) = 0.1032, P(Yes) = 0.8968
Sample 88: Predicted Outcome = 0, P(No) = 0.6902, P(Yes) = 0.3098
Sample 89: Predicted Outcome = 0, P(No) = 0.8759, P(Yes) = 0.1241
Sample 90: Predicted Outcome = 0, P(No) = 0.9142, P(Yes) = 0.0858
Sample 91: Predicted Outcome = 0, P(No) = 0.9324, P(Yes) = 0.0676
Sample 92: Predicted Outcome = 0, P(No) = 0.7634, P(Yes) = 0.2366
Sample 93: Predicted Outcome = 0, P(No) = 0.9956, P(Yes) = 0.0044
Sample 94: Predicted Outcome = 1, P(No) = 0.4601, P(Yes) = 0.5399
Sample 95: Predicted Outcome = 1, P(No) = 0.4874, P(Yes) = 0.5126
Sample 96: Predicted Outcome = 1, P(No) = 0.3556, P(Yes) = 0.6444
Sample 97: Predicted Outcome = 0, P(No) = 0.6315, P(Yes) = 0.3685
Sample 98: Predicted Outcome = 0, P(No) = 0.8686, P(Yes) = 0.1314
Sample 99: Predicted Outcome = 1, P(No) = 0.3873, P(Yes) = 0.6127
Sample 100: Predicted Outcome = 0, P(No) = 0.9150, P(Yes) = 0.0850
Sample 101: Predicted Outcome = 1, P(No) = 0.2838, P(Yes) = 0.7162
Sample 102: Predicted Outcome = 0, P(No) = 0.9357, P(Yes) = 0.0643
Sample 103: Predicted Outcome = 1, P(No) = 0.2236, P(Yes) = 0.7764
Sample 104: Predicted Outcome = 0, P(No) = 0.5024, P(Yes) = 0.4976
Sample 105: Predicted Outcome = 1, P(No) = 0.3566, P(Yes) = 0.6434
Sample 106: Predicted Outcome = 0, P(No) = 0.7817, P(Yes) = 0.2183
Sample 107: Predicted Outcome = 0, P(No) = 0.7393, P(Yes) = 0.2607
Sample 108: Predicted Outcome = 1, P(No) = 0.2634, P(Yes) = 0.7366
Sample 109: Predicted Outcome = 0, P(No) = 0.8639, P(Yes) = 0.1361
Sample 110: Predicted Outcome = 1, P(No) = 0.4573, P(Yes) = 0.5427
```

```
Sample 111: Predicted Outcome = 0, P(No) = 0.8995, P(Yes) = 0.1005
Sample 112: Predicted Outcome = 0, P(No) = 0.6827, P(Yes) = 0.3173
Sample 113: Predicted Outcome = 0, P(No) = 0.9792, P(Yes) = 0.0208
Sample 114: Predicted Outcome = 1, P(No) = 0.2718, P(Yes) = 0.7282
Sample 115: Predicted Outcome = 0, P(No) = 0.8227, P(Yes) = 0.1773
Sample 116: Predicted Outcome = 0, P(No) = 0.6631, P(Yes) = 0.3369
Sample 117: Predicted Outcome = 1, P(No) = 0.2327, P(Yes) = 0.7673
Sample 118: Predicted Outcome = 0, P(No) = 0.7785, P(Yes) = 0.2215
Sample 119: Predicted Outcome = 0, P(No) = 0.9340, P(Yes) = 0.0660
Sample 120: Predicted Outcome = 1, P(No) = 0.4278, P(Yes) = 0.5722
Sample 121: Predicted Outcome = 0, P(No) = 0.9406, P(Yes) = 0.0594
Sample 122: Predicted Outcome = 0, P(No) = 0.7023, P(Yes) = 0.2977
Sample 123: Predicted Outcome = 0, P(No) = 0.7601, P(Yes) = 0.2399
Sample 124: Predicted Outcome = 0, P(No) = 0.9167, P(Yes) = 0.0833
Sample 125: Predicted Outcome = 0, P(No) = 0.7383, P(Yes) = 0.2617
Sample 126: Predicted Outcome = 0, P(No) = 0.5830, P(Yes) = 0.4170
Sample 127: Predicted Outcome = 0, P(No) = 0.9584, P(Yes) = 0.0416
Sample 128: Predicted Outcome = 1, P(No) = 0.1257, P(Yes) = 0.8743
Sample 129: Predicted Outcome = 1, P(No) = 0.0310, P(Yes) = 0.9690
Sample 130: Predicted Outcome = 1, P(No) = 0.2722, P(Yes) = 0.7278
Sample 131: Predicted Outcome = 1, P(No) = 0.2990, P(Yes) = 0.7010
Sample 132: Predicted Outcome = 1, P(No) = 0.1497, P(Yes) = 0.8503
Sample 133: Predicted Outcome = 0, P(No) = 0.9101, P(Yes) = 0.0899
Sample 134: Predicted Outcome = 0, P(No) = 0.5818, P(Yes) = 0.4182
Sample 135: Predicted Outcome = 1, P(No) = 0.1852, P(Yes) = 0.8148
Sample 136: Predicted Outcome = 0, P(No) = 0.8877, P(Yes) = 0.1123
Sample 137: Predicted Outcome = 0, P(No) = 0.8439, P(Yes) = 0.1561
Sample 138: Predicted Outcome = 1, P(No) = 0.1501, P(Yes) = 0.8499
Sample 139: Predicted Outcome = 1, P(No) = 0.2008, P(Yes) = 0.7992
Sample 140: Predicted Outcome = 0, P(No) = 0.9863, P(Yes) = 0.0137
Sample 141: Predicted Outcome = 0, P(No) = 0.9041, P(Yes) = 0.0959
Sample 142: Predicted Outcome = 0, P(No) = 0.9592, P(Yes) = 0.0408
Sample 143: Predicted Outcome = 0, P(No) = 0.7986, P(Yes) = 0.2014
Sample 144: Predicted Outcome = 0, P(No) = 0.6218, P(Yes) = 0.3782
Sample 145: Predicted Outcome = 0, P(No) = 0.8754, P(Yes) = 0.1246
Sample 146: Predicted Outcome = 0, P(No) = 0.7079, P(Yes) = 0.2921
Sample 147: Predicted Outcome = 0, P(No) = 0.8504, P(Yes) = 0.1496
Sample 148: Predicted Outcome = 0, P(No) = 0.9778, P(Yes) = 0.0222
Sample 149: Predicted Outcome = 0, P(No) = 0.5766, P(Yes) = 0.4234
Sample 150: Predicted Outcome = 1, P(No) = 0.2542, P(Yes) = 0.7458
Sample 151: Predicted Outcome = 0, P(No) = 0.8739, P(Yes) = 0.1261
Sample 152: Predicted Outcome = 0, P(No) = 0.5135, P(Yes) = 0.4865
Sample 153: Predicted Outcome = 0, P(No) = 0.7735, P(Yes) = 0.2265
Sample 154: Predicted Outcome = 0, P(No) = 0.7951, P(Yes) = 0.2049
```

[25]:
```
#Model Evaluation:
print("\nModel Evaluation:")
```

```
print("Accuracy:", accuracy_score(y_test, y_pred_all))
```

```
Model Evaluation:
Accuracy: 0.7532467532467533
```

[27]:
```
#Classification Report:
print("\nClassification Report:") # Genrates precision , recall,,f-1 score(mean of precison & recall) and
                                 # support(no of occurence of each sample in dataset)
print(classification_report(y_test, y_pred_all))
```

```
Classification Report:
              precision    recall  f1-score   support

           0       0.81      0.80      0.81        99
           1       0.65      0.67      0.66        55

    accuracy                           0.75       154
   macro avg       0.73      0.74      0.73       154
weighted avg       0.76      0.75      0.75       154
```

[28]:
```
#creating sample datapoints:
sample_data = pd.DataFrame({
    'Pregnancies': [1],
    'Glucose': [85],
    'BloodPressure': [66],
    'SkinThickness': [29],
    'Insulin': [0],
    'BMI': [26.6],
    'DiabetesPedigreeFunction': [0.351],
    'Age': [50]
})
```

[29]:
```
#Scalimng sample data:
sample_scaled = scaler.transform(sample_data)
```

[30]:
```
#rediction making for sample data:
sample_prediction = model.predict(sample_scaled)
```

[31]:
```
#Getting prediction probablity for it:
sample_probability = model.predict_proba(sample_scaled)
```

[32]:
```
#Printing Prediction & Probablity:
print("\nPrediction for the sample data:")
print(f"Predicted Outcome: {sample_prediction[0]}")
print(f"Probabilities- P(No): {sample_probability[0][0]:.4f}, P(Yes):␣
 ↪{sample_probability[0][1]:.4f}")
```

```
Prediction for the sample data:
Predicted Outcome: 0
Probabilities- P(No): 0.8856, P(Yes): 0.1144
```

[ ]: