Breastcancer_prediction

August 1, 2025

[3]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LogisticRegression
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score
```

[31]:
```
cancer_df=pd.read_csv("Breast_cancer_dataset.csv")
cancer_df
```

[31]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	\
0	842302	M	17.99	10.38	122.80	1001.0	
1	842517	M	20.57	17.77	132.90	1326.0	
2	84300903	M	19.69	21.25	130.00	1203.0	
3	84348301	M	11.42	20.38	77.58	386.1	
4	84358402	M	20.29	14.34	135.10	1297.0	
..	…	…	…	…	…	…	
564	926424	M	21.56	22.39	142.00	1479.0	
565	926682	M	20.13	28.25	131.20	1261.0	
566	926954	M	16.60	28.08	108.30	858.1	
567	927241	M	20.60	29.33	140.10	1265.0	
568	92751	B	7.76	24.54	47.92	181.0	

	smoothness_mean	compactness_mean	concavity_mean	concave points_mean	\
0	0.11840	0.27760	0.30010	0.14710	
1	0.08474	0.07864	0.08690	0.07017	
2	0.10960	0.15990	0.19740	0.12790	
3	0.14250	0.28390	0.24140	0.10520	
4	0.10030	0.13280	0.19800	0.10430	
..	…	…	…	…	
564	0.11100	0.11590	0.24390	0.13890	
565	0.09780	0.10340	0.14400	0.09791	
566	0.08455	0.10230	0.09251	0.05302	
567	0.11780	0.27700	0.35140	0.15200	
568	0.05263	0.04362	0.00000	0.00000	

```
     …  radius_worst  texture_worst  perimeter_worst  area_worst  \
0    …        25.380          17.33           184.60      2019.0
1    …        24.990          23.41           158.80      1956.0
2    …        23.570          25.53           152.50      1709.0
3    …        14.910          26.50            98.87       567.7
4    …        22.540          16.67           152.20      1575.0
..   …           …              …                …          …
564  …        25.450          26.40           166.10      2027.0
565  …        23.690          38.25           155.00      1731.0
566  …        18.980          34.12           126.70      1124.0
567  …        25.740          39.42           184.60      1821.0
568  …         9.456          30.37            59.16       268.6

     smoothness_worst  compactness_worst  concavity_worst  \
0             0.16220            0.66560           0.7119
1             0.12380            0.18660           0.2416
2             0.14440            0.42450           0.4504
3             0.20980            0.86630           0.6869
4             0.13740            0.20500           0.4000
..                ...                ...              ...
564           0.14100            0.21130           0.4107
565           0.11660            0.19220           0.3215
566           0.11390            0.30940           0.3403
567           0.16500            0.86810           0.9387
568           0.08996            0.06444           0.0000

     concave points_worst  symmetry_worst  fractal_dimension_worst
0                  0.2654          0.4601                  0.11890
1                  0.1860          0.2750                  0.08902
2                  0.2430          0.3613                  0.08758
3                  0.2575          0.6638                  0.17300
4                  0.1625          0.2364                  0.07678
..                    ...             ...                      ...
564                0.2216          0.2060                  0.07115
565                0.1628          0.2572                  0.06637
566                0.1418          0.2218                  0.07820
567                0.2650          0.4087                  0.12400
568                0.0000          0.2871                  0.07039

[569 rows x 32 columns]
```

[33]:
```
#display first ten rows
cancer_df.head(10)
```

[33]:
```
         id diagnosis  radius_mean  texture_mean  perimeter_mean  area_mean  \
0    842302         M        17.99         10.38          122.80     1001.0
1    842517         M        20.57         17.77          132.90     1326.0
```

```
2  84300903         M        19.69         21.25          130.00     1203.0
3  84348301         M        11.42         20.38           77.58      386.1
4  84358402         M        20.29         14.34          135.10     1297.0
5    843786         M        12.45         15.70           82.57      477.1
6    844359         M        18.25         19.98          119.60     1040.0
7  84458202         M        13.71         20.83           90.20      577.9
8    844981         M        13.00         21.82           87.50      519.8
9  84501001         M        12.46         24.04           83.97      475.9

   smoothness_mean  compactness_mean  concavity_mean  concave points_mean  \
0          0.11840           0.27760         0.30010              0.14710
1          0.08474           0.07864         0.08690              0.07017
2          0.10960           0.15990         0.19740              0.12790
3          0.14250           0.28390         0.24140              0.10520
4          0.10030           0.13280         0.19800              0.10430
5          0.12780           0.17000         0.15780              0.08089
6          0.09463           0.10900         0.11270              0.07400
7          0.11890           0.16450         0.09366              0.05985
8          0.12730           0.19320         0.18590              0.09353
9          0.11860           0.23960         0.22730              0.08543

   …  radius_worst  texture_worst  perimeter_worst  area_worst  \
0  …         25.38          17.33           184.60      2019.0
1  …         24.99          23.41           158.80      1956.0
2  …         23.57          25.53           152.50      1709.0
3  …         14.91          26.50            98.87       567.7
4  …         22.54          16.67           152.20      1575.0
5  …         15.47          23.75           103.40       741.6
6  …         22.88          27.66           153.20      1606.0
7  …         17.06          28.14           110.60       897.0
8  …         15.49          30.73           106.20       739.3
9  …         15.09          40.68            97.65       711.4

   smoothness_worst  compactness_worst  concavity_worst  concave points_worst  \
0            0.1622             0.6656           0.7119                0.2654
1            0.1238             0.1866           0.2416                0.1860
2            0.1444             0.4245           0.4504                0.2430
3            0.2098             0.8663           0.6869                0.2575
4            0.1374             0.2050           0.4000                0.1625
5            0.1791             0.5249           0.5355                0.1741
6            0.1442             0.2576           0.3784                0.1932
7            0.1654             0.3682           0.2678                0.1556
8            0.1703             0.5401           0.5390                0.2060
9            0.1853             1.0580           1.1050                0.2210

   symmetry_worst  fractal_dimension_worst
0          0.4601                  0.11890
```

```
1          0.2750                  0.08902
2          0.3613                  0.08758
3          0.6638                  0.17300
4          0.2364                  0.07678
5          0.3985                  0.12440
6          0.3063                  0.08368
7          0.3196                  0.11510
8          0.4378                  0.10720
9          0.4366                  0.20750

[10 rows x 32 columns]
```

[35]:
```
#checking for null values
cancer_df.isnull().sum()
```

[35]:
```
id                         0
diagnosis                  0
radius_mean                0
texture_mean               0
perimeter_mean             0
area_mean                  0
smoothness_mean            0
compactness_mean           0
concavity_mean             0
concave points_mean        0
symmetry_mean              0
fractal_dimension_mean     0
radius_se                  0
texture_se                 0
perimeter_se               0
area_se                    0
smoothness_se              0
compactness_se             0
concavity_se               0
concave points_se          0
symmetry_se                0
fractal_dimension_se       0
radius_worst               0
texture_worst              0
perimeter_worst            0
area_worst                 0
smoothness_worst           0
compactness_worst          0
concavity_worst            0
concave points_worst       0
symmetry_worst             0
fractal_dimension_worst    0
```

```
dtype: int64
```

[37]:
```
#check for datatypes
cancer_df.dtypes
```

[37]:
```
id                         int64
diagnosis                 object
radius_mean              float64
texture_mean             float64
perimeter_mean           float64
area_mean                float64
smoothness_mean          float64
compactness_mean         float64
concavity_mean           float64
concave points_mean      float64
symmetry_mean            float64
fractal_dimension_mean   float64
radius_se                float64
texture_se               float64
perimeter_se             float64
area_se                  float64
smoothness_se            float64
compactness_se           float64
concavity_se             float64
concave points_se        float64
symmetry_se              float64
fractal_dimension_se     float64
radius_worst             float64
texture_worst            float64
perimeter_worst          float64
area_worst               float64
smoothness_worst         float64
compactness_worst        float64
concavity_worst          float64
concave points_worst     float64
symmetry_worst           float64
fractal_dimension_worst  float64
dtype: object
```

[39]:
```
cancer_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 569 entries, 0 to 568
Data columns (total 32 columns):
 #   Column                   Non-Null Count  Dtype
---  ------                   --------------  -----
 0   id                       569 non-null    int64
 1   diagnosis                569 non-null    object
```

```
 2   radius_mean              569 non-null    float64
 3   texture_mean             569 non-null    float64
 4   perimeter_mean           569 non-null    float64
 5   area_mean                569 non-null    float64
 6   smoothness_mean          569 non-null    float64
 7   compactness_mean         569 non-null    float64
 8   concavity_mean           569 non-null    float64
 9   concave points_mean      569 non-null    float64
 10  symmetry_mean            569 non-null    float64
 11  fractal_dimension_mean   569 non-null    float64
 12  radius_se                569 non-null    float64
 13  texture_se               569 non-null    float64
 14  perimeter_se             569 non-null    float64
 15  area_se                  569 non-null    float64
 16  smoothness_se            569 non-null    float64
 17  compactness_se           569 non-null    float64
 18  concavity_se             569 non-null    float64
 19  concave points_se        569 non-null    float64
 20  symmetry_se              569 non-null    float64
 21  fractal_dimension_se     569 non-null    float64
 22  radius_worst             569 non-null    float64
 23  texture_worst            569 non-null    float64
 24  perimeter_worst          569 non-null    float64
 25  area_worst               569 non-null    float64
 26  smoothness_worst         569 non-null    float64
 27  compactness_worst        569 non-null    float64
 28  concavity_worst          569 non-null    float64
 29  concave points_worst     569 non-null    float64
 30  symmetry_worst           569 non-null    float64
 31  fractal_dimension_worst  569 non-null    float64
dtypes: float64(30), int64(1), object(1)
memory usage: 142.4+ KB
```

[41]:
```
#descriptive analysis
cancer_df.describe()
```

[41]:

	id	radius_mean	texture_mean	perimeter_mean	area_mean	\
count	5.690000e+02	569.000000	569.000000	569.000000	569.000000	
mean	3.037183e+07	14.127292	19.289649	91.969033	654.889104	
std	1.250206e+08	3.524049	4.301036	24.298981	351.914129	
min	8.670000e+03	6.981000	9.710000	43.790000	143.500000	
25%	8.692180e+05	11.700000	16.170000	75.170000	420.300000	
50%	9.060240e+05	13.370000	18.840000	86.240000	551.100000	
75%	8.813129e+06	15.780000	21.800000	104.100000	782.700000	
max	9.113205e+08	28.110000	39.280000	188.500000	2501.000000	

	smoothness_mean	compactness_mean	concavity_mean	concave points_mean	\

```
count       569.000000        569.000000      569.000000           569.000000
mean          0.096360          0.104341        0.088799             0.048919
std           0.014064          0.052813        0.079720             0.038803
min           0.052630          0.019380        0.000000             0.000000
25%           0.086370          0.064920        0.029560             0.020310
50%           0.095870          0.092630        0.061540             0.033500
75%           0.105300          0.130400        0.130700             0.074000
max           0.163400          0.345400        0.426800             0.201200

       symmetry_mean  …  radius_worst  texture_worst  perimeter_worst  \
count     569.000000  …    569.000000     569.000000       569.000000
mean        0.181162  …     16.269190      25.677223       107.261213
std         0.027414  …      4.833242       6.146258        33.602542
min         0.106000  …      7.930000      12.020000        50.410000
25%         0.161900  …     13.010000      21.080000        84.110000
50%         0.179200  …     14.970000      25.410000        97.660000
75%         0.195700  …     18.790000      29.720000       125.400000
max         0.304000  …     36.040000      49.540000       251.200000

        area_worst  smoothness_worst  compactness_worst  concavity_worst  \
count   569.000000        569.000000         569.000000       569.000000
mean    880.583128          0.132369           0.254265         0.272188
std     569.356993          0.022832           0.157336         0.208624
min     185.200000          0.071170           0.027290         0.000000
25%     515.300000          0.116600           0.147200         0.114500
50%     686.500000          0.131300           0.211900         0.226700
75%    1084.000000          0.146000           0.339100         0.382900
max    4254.000000          0.222600           1.058000         1.252000

       concave points_worst  symmetry_worst  fractal_dimension_worst
count            569.000000      569.000000               569.000000
mean               0.114606        0.290076                 0.083946
std                0.065732        0.061867                 0.018061
min                0.000000        0.156500                 0.055040
25%                0.064930        0.250400                 0.071460
50%                0.099930        0.282200                 0.080040
75%                0.161400        0.317900                 0.092080
max                0.291000        0.663800                 0.207500

[8 rows x 31 columns]
```

[47]:
```
#encoding the categorical data
cancer_df['diagnosis']=cancer_df['diagnosis'].map({'M':1,'B':0})
cancer_df['diagnosis']
```

[47]:
```
0      1
1      1
```

```
2      1
3      1
4      1
      ..
564    1
565    1
566    1
567    1
568    0
Name: diagnosis, Length: 569, dtype: int64
```

[49]:
```
cancer_df.corr()
```

[49]:

	id	diagnosis	radius_mean	texture_mean	\
id	1.000000	0.039769	0.074626	0.099770	
diagnosis	0.039769	1.000000	0.730029	0.415185	
radius_mean	0.074626	0.730029	1.000000	0.323782	
texture_mean	0.099770	0.415185	0.323782	1.000000	
perimeter_mean	0.073159	0.742636	0.997855	0.329533	
area_mean	0.096893	0.708984	0.987357	0.321086	
smoothness_mean	-0.012968	0.358560	0.170581	-0.023389	
compactness_mean	0.000096	0.596534	0.506124	0.236702	
concavity_mean	0.050080	0.696360	0.676764	0.302418	
concave points_mean	0.044158	0.776614	0.822529	0.293464	
symmetry_mean	-0.022114	0.330499	0.147741	0.071401	
fractal_dimension_mean	-0.052511	-0.012838	-0.311631	-0.076437	
radius_se	0.143048	0.567134	0.679090	0.275869	
texture_se	-0.007526	-0.008303	-0.097317	0.386358	
perimeter_se	0.137331	0.556141	0.674172	0.281673	
area_se	0.177742	0.548236	0.735864	0.259845	
smoothness_se	0.096781	-0.067016	-0.222600	0.006614	
compactness_se	0.033961	0.292999	0.206000	0.191975	
concavity_se	0.055239	0.253730	0.194204	0.143293	
concave points_se	0.078768	0.408042	0.376169	0.163851	
symmetry_se	-0.017306	-0.006522	-0.104321	0.009127	
fractal_dimension_se	0.025725	0.077972	-0.042641	0.054458	
radius_worst	0.082405	0.776454	0.969539	0.352573	
texture_worst	0.064720	0.456903	0.297008	0.912045	
perimeter_worst	0.079986	0.782914	0.965137	0.358040	
area_worst	0.107187	0.733825	0.941082	0.343546	
smoothness_worst	0.010338	0.421465	0.119616	0.077503	
compactness_worst	-0.002968	0.590998	0.413463	0.277830	
concavity_worst	0.023203	0.659610	0.526911	0.301025	
concave points_worst	0.035174	0.793566	0.744214	0.295316	
symmetry_worst	-0.044224	0.416294	0.163953	0.105008	
fractal_dimension_worst	-0.029866	0.323872	0.007066	0.119205	

	perimeter_mean	area_mean	smoothness_mean	\
id	0.073159	0.096893	-0.012968	
diagnosis	0.742636	0.708984	0.358560	
radius_mean	0.997855	0.987357	0.170581	
texture_mean	0.329533	0.321086	-0.023389	
perimeter_mean	1.000000	0.986507	0.207278	
area_mean	0.986507	1.000000	0.177028	
smoothness_mean	0.207278	0.177028	1.000000	
compactness_mean	0.556936	0.498502	0.659123	
concavity_mean	0.716136	0.685983	0.521984	
concave points_mean	0.850977	0.823269	0.553695	
symmetry_mean	0.183027	0.151293	0.557775	
fractal_dimension_mean	-0.261477	-0.283110	0.584792	
radius_se	0.691765	0.732562	0.301467	
texture_se	-0.086761	-0.066280	0.068406	
perimeter_se	0.693135	0.726628	0.296092	
area_se	0.744983	0.800086	0.246552	
smoothness_se	-0.202694	-0.166777	0.332375	
compactness_se	0.250744	0.212583	0.318943	
concavity_se	0.228082	0.207660	0.248396	
concave points_se	0.407217	0.372320	0.380676	
symmetry_se	-0.081629	-0.072497	0.200774	
fractal_dimension_se	-0.005523	-0.019887	0.283607	
radius_worst	0.969476	0.962746	0.213120	
texture_worst	0.303038	0.287489	0.036072	
perimeter_worst	0.970387	0.959120	0.238853	
area_worst	0.941550	0.959213	0.206718	
smoothness_worst	0.150549	0.123523	0.805324	
compactness_worst	0.455774	0.390410	0.472468	
concavity_worst	0.563879	0.512606	0.434926	
concave points_worst	0.771241	0.722017	0.503053	
symmetry_worst	0.189115	0.143570	0.394309	
fractal_dimension_worst	0.051019	0.003738	0.499316	

	compactness_mean	concavity_mean	\
id	0.000096	0.050080	
diagnosis	0.596534	0.696360	
radius_mean	0.506124	0.676764	
texture_mean	0.236702	0.302418	
perimeter_mean	0.556936	0.716136	
area_mean	0.498502	0.685983	
smoothness_mean	0.659123	0.521984	
compactness_mean	1.000000	0.883121	
concavity_mean	0.883121	1.000000	
concave points_mean	0.831135	0.921391	
symmetry_mean	0.602641	0.500667	
fractal_dimension_mean	0.565369	0.336783	

```
radius_se                       0.497473        0.631925
texture_se                      0.046205        0.076218
perimeter_se                    0.548905        0.660391
area_se                         0.455653        0.617427
smoothness_se                   0.135299        0.098564
compactness_se                  0.738722        0.670279
concavity_se                    0.570517        0.691270
concave points_se               0.642262        0.683260
symmetry_se                     0.229977        0.178009
fractal_dimension_se            0.507318        0.449301
radius_worst                    0.535315        0.688236
texture_worst                   0.248133        0.299879
perimeter_worst                 0.590210        0.729565
area_worst                      0.509604        0.675987
smoothness_worst                0.565541        0.448822
compactness_worst               0.865809        0.754968
concavity_worst                 0.816275        0.884103
concave points_worst            0.815573        0.861323
symmetry_worst                  0.510223        0.409464
fractal_dimension_worst         0.687382        0.514930

                        concave points_mean  …  radius_worst  \
id                                 0.044158  …      0.082405
diagnosis                          0.776614  …      0.776454
radius_mean                        0.822529  …      0.969539
texture_mean                       0.293464  …      0.352573
perimeter_mean                     0.850977  …      0.969476
area_mean                          0.823269  …      0.962746
smoothness_mean                    0.553695  …      0.213120
compactness_mean                   0.831135  …      0.535315
concavity_mean                     0.921391  …      0.688236
concave points_mean                1.000000  …      0.830318
symmetry_mean                      0.462497  …      0.185728
fractal_dimension_mean             0.166917  …     -0.253691
radius_se                          0.698050  …      0.715065
texture_se                         0.021480  …     -0.111690
perimeter_se                       0.710650  …      0.697201
area_se                            0.690299  …      0.757373
smoothness_se                      0.027653  …     -0.230691
compactness_se                     0.490424  …      0.204607
concavity_se                       0.439167  …      0.186904
concave points_se                  0.615634  …      0.358127
symmetry_se                        0.095351  …     -0.128121
fractal_dimension_se               0.257584  …     -0.037488
radius_worst                       0.830318  …      1.000000
texture_worst                      0.292752  …      0.359921
perimeter_worst                    0.855923  …      0.993708
```

```
area_worst                      0.809630  …      0.984015
smoothness_worst                0.452753  …      0.216574
compactness_worst               0.667454  …      0.475820
concavity_worst                 0.752399  …      0.573975
concave points_worst            0.910155  …      0.787424
symmetry_worst                  0.375744  …      0.243529
fractal_dimension_worst         0.368661  …      0.093492

                        texture_worst  perimeter_worst  area_worst  \
id                           0.064720         0.079986    0.107187
diagnosis                    0.456903         0.782914    0.733825
radius_mean                  0.297008         0.965137    0.941082
texture_mean                 0.912045         0.358040    0.343546
perimeter_mean               0.303038         0.970387    0.941550
area_mean                    0.287489         0.959120    0.959213
smoothness_mean              0.036072         0.238853    0.206718
compactness_mean             0.248133         0.590210    0.509604
concavity_mean               0.299879         0.729565    0.675987
concave points_mean          0.292752         0.855923    0.809630
symmetry_mean                0.090651         0.219169    0.177193
fractal_dimension_mean      -0.051269        -0.205151   -0.231854
radius_se                    0.194799         0.719684    0.751548
texture_se                   0.409003        -0.102242   -0.083195
perimeter_se                 0.200371         0.721031    0.730713
area_se                      0.196497         0.761213    0.811408
smoothness_se               -0.074743        -0.217304   -0.182195
compactness_se               0.143003         0.260516    0.199371
concavity_se                 0.100241         0.226680    0.188353
concave points_se            0.086741         0.394999    0.342271
symmetry_se                 -0.077473        -0.103753   -0.110343
fractal_dimension_se        -0.003195        -0.001000   -0.022736
radius_worst                 0.359921         0.993708    0.984015
texture_worst                1.000000         0.365098    0.345842
perimeter_worst              0.365098         1.000000    0.977578
area_worst                   0.345842         0.977578    1.000000
smoothness_worst             0.225429         0.236775    0.209145
compactness_worst            0.360832         0.529408    0.438296
concavity_worst              0.368366         0.618344    0.543331
concave points_worst         0.359755         0.816322    0.747419
symmetry_worst               0.233027         0.269493    0.209146
fractal_dimension_worst      0.219122         0.138957    0.079647

                        smoothness_worst  compactness_worst  concavity_worst  \
id                              0.010338          -0.002968         0.023203
diagnosis                       0.421465           0.590998         0.659610
radius_mean                     0.119616           0.413463         0.526911
texture_mean                    0.077503           0.277830         0.301025
```

perimeter_mean	0.150549	0.455774	0.563879
area_mean	0.123523	0.390410	0.512606
smoothness_mean	0.805324	0.472468	0.434926
compactness_mean	0.565541	0.865809	0.816275
concavity_mean	0.448822	0.754968	0.884103
concave points_mean	0.452753	0.667454	0.752399
symmetry_mean	0.426675	0.473200	0.433721
fractal_dimension_mean	0.504942	0.458798	0.346234
radius_se	0.141919	0.287103	0.380585
texture_se	-0.073658	-0.092439	-0.068956
perimeter_se	0.130054	0.341919	0.418899
area_se	0.125389	0.283257	0.385100
smoothness_se	0.314457	-0.055558	-0.058298
compactness_se	0.227394	0.678780	0.639147
concavity_se	0.168481	0.484858	0.662564
concave points_se	0.215351	0.452888	0.549592
symmetry_se	-0.012662	0.060255	0.037119
fractal_dimension_se	0.170568	0.390159	0.379975
radius_worst	0.216574	0.475820	0.573975
texture_worst	0.225429	0.360832	0.368366
perimeter_worst	0.236775	0.529408	0.618344
area_worst	0.209145	0.438296	0.543331
smoothness_worst	1.000000	0.568187	0.518523
compactness_worst	0.568187	1.000000	0.892261
concavity_worst	0.518523	0.892261	1.000000
concave points_worst	0.547691	0.801080	0.855434
symmetry_worst	0.493838	0.614441	0.532520
fractal_dimension_worst	0.617624	0.810455	0.686511

	concave points_worst	symmetry_worst	\
id	0.035174	-0.044224	
diagnosis	0.793566	0.416294	
radius_mean	0.744214	0.163953	
texture_mean	0.295316	0.105008	
perimeter_mean	0.771241	0.189115	
area_mean	0.722017	0.143570	
smoothness_mean	0.503053	0.394309	
compactness_mean	0.815573	0.510223	
concavity_mean	0.861323	0.409464	
concave points_mean	0.910155	0.375744	
symmetry_mean	0.430297	0.699826	
fractal_dimension_mean	0.175325	0.334019	
radius_se	0.531062	0.094543	
texture_se	-0.119638	-0.128215	
perimeter_se	0.554897	0.109930	
area_se	0.538166	0.074126	
smoothness_se	-0.102007	-0.107342	

compactness_se	0.483208	0.277878
concavity_se	0.440472	0.197788
concave points_se	0.602450	0.143116
symmetry_se	-0.030413	0.389402
fractal_dimension_se	0.215204	0.111094
radius_worst	0.787424	0.243529
texture_worst	0.359755	0.233027
perimeter_worst	0.816322	0.269493
area_worst	0.747419	0.209146
smoothness_worst	0.547691	0.493838
compactness_worst	0.801080	0.614441
concavity_worst	0.855434	0.532520
concave points_worst	1.000000	0.502528
symmetry_worst	0.502528	1.000000
fractal_dimension_worst	0.511114	0.537848

	fractal_dimension_worst
id	-0.029866
diagnosis	0.323872
radius_mean	0.007066
texture_mean	0.119205
perimeter_mean	0.051019
area_mean	0.003738
smoothness_mean	0.499316
compactness_mean	0.687382
concavity_mean	0.514930
concave points_mean	0.368661
symmetry_mean	0.438413
fractal_dimension_mean	0.767297
radius_se	0.049559
texture_se	-0.045655
perimeter_se	0.085433
area_se	0.017539
smoothness_se	0.101480
compactness_se	0.590973
concavity_se	0.439329
concave points_se	0.310655
symmetry_se	0.078079
fractal_dimension_se	0.591328
radius_worst	0.093492
texture_worst	0.219122
perimeter_worst	0.138957
area_worst	0.079647
smoothness_worst	0.617624
compactness_worst	0.810455
concavity_worst	0.686511
concave points_worst	0.511114

```
symmetry_worst                          0.537848
fractal_dimension_worst                 1.000000

[32 rows x 32 columns]
```

[51]:
```
#checking for balance
cancer_df['diagnosis'].value_counts()
```

[51]:
```
diagnosis
0    357
1    212
Name: count, dtype: int64
```

[61]:
```
#split the dataset into features and labels
X=cancer_df[['concave points_mean','radius_worst','perimeter_worst','concave␣
 ↪points_worst']]
y=cancer_df['diagnosis']
```

[63]:
```
print(X)
```

```
     concave points_mean  radius_worst  perimeter_worst  concave points_worst
0                0.14710        25.380           184.60                0.2654
1                0.07017        24.990           158.80                0.1860
2                0.12790        23.570           152.50                0.2430
3                0.10520        14.910            98.87                0.2575
4                0.10430        22.540           152.20                0.1625
..                   …             …                …                     …
564              0.13890        25.450           166.10                0.2216
565              0.09791        23.690           155.00                0.1628
566              0.05302        18.980           126.70                0.1418
567              0.15200        25.740           184.60                0.2650
568              0.00000         9.456            59.16                0.0000

[569 rows x 4 columns]
```

[65]:
```
print(y)
```

```
0      1
1      1
2      1
3      1
4      1
      ..
564    1
565    1
566    1
567    1
568    0
```

```
Name: diagnosis, Length: 569, dtype: int64
```

[67]:
```
#split the data
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.
↪2,random_state=42)
```

[69]:
```
#standardize features
scaler=StandardScaler()
X_train_scaled=scaler.fit_transform(X_train)
X_test_scaled=scaler.transform(X_test)
```

[71]:
```
from imblearn.over_sampling import SMOTE
smote = SMOTE(random_state=42)
X_train_resampled, y_train_resampled = smote.fit_resample(X_train_scaled,␣
↪y_train)
```

[73]:
```
#create and train the model
model=LogisticRegression()
model.fit(X_train_scaled,y_train)
```

[73]:
```
LogisticRegression()
```

[75]:
```
#make predictions
y_pred=model.predict(X_test_scaled)
```

[77]:
```
#evaluate the model
accuracy=accuracy_score(y_test,y_pred)
print("Accuracy:",accuracy)
```

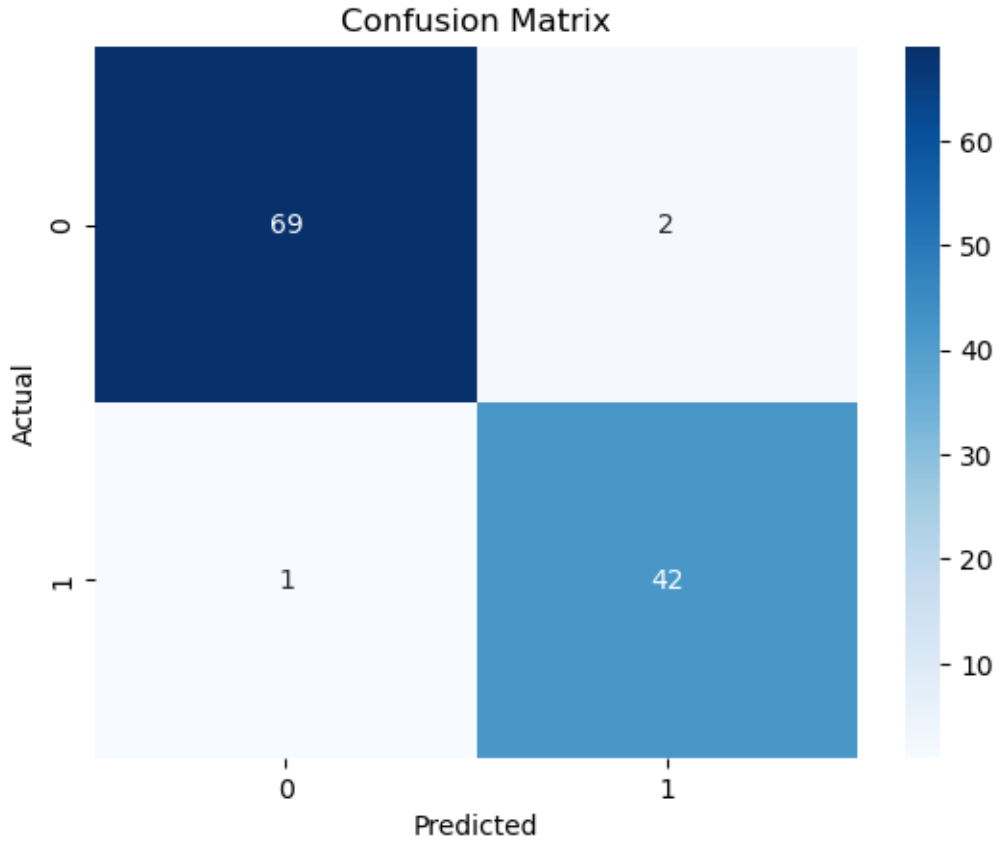
```
Accuracy: 0.9736842105263158
```

[79]:
```
from sklearn.metrics import classification_report
print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

           0       0.99      0.97      0.98        71
           1       0.95      0.98      0.97        43

    accuracy                           0.97       114
   macro avg       0.97      0.97      0.97       114
weighted avg       0.97      0.97      0.97       114
```

[81]:
```
from sklearn.metrics import confusion_matrix


cm = confusion_matrix(y_test, y_pred)
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues')
```

```
plt.xlabel("Predicted")
plt.ylabel("Actual")
plt.title("Confusion Matrix")
plt.show()
```



[ ]: