Breast_Cancer_Prediction_using_ml_model

February 3, 2021

[1]:
```
import warnings
warnings.filterwarnings('ignore')
```

[2]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[3]:
```
df = pd.read_csv("https://raw.githubusercontent.com/ingledarshan/AIML-B2/main/
↪data.csv")
```

[67]:
```
df.head(10)
```

[67]:

	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	\
0	1	17.99	10.38	122.80	1001.0	
1	1	20.57	17.77	132.90	1326.0	
2	1	19.69	21.25	130.00	1203.0	
3	1	11.42	20.38	77.58	386.1	
4	1	20.29	14.34	135.10	1297.0	
5	1	12.45	15.70	82.57	477.1	
6	1	18.25	19.98	119.60	1040.0	
7	1	13.71	20.83	90.20	577.9	
8	1	13.00	21.82	87.50	519.8	
9	1	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	

```
   symmetry_mean  …  radius_worst  texture_worst  perimeter_worst  \
0         0.2419  …         25.38          17.33           184.60
1         0.1812  …         24.99          23.41           158.80
2         0.2069  …         23.57          25.53           152.50
3         0.2597  …         14.91          26.50            98.87
4         0.1809  …         22.54          16.67           152.20
5         0.2087  …         15.47          23.75           103.40
6         0.1794  …         22.88          27.66           153.20
7         0.2196  …         17.06          28.14           110.60
8         0.2350  …         15.49          30.73           106.20
9         0.2030  …         15.09          40.68            97.65

   area_worst  smoothness_worst  compactness_worst  concavity_worst  \
0      2019.0            0.1622             0.6656           0.7119
1      1956.0            0.1238             0.1866           0.2416
2      1709.0            0.1444             0.4245           0.4504
3       567.7            0.2098             0.8663           0.6869
4      1575.0            0.1374             0.2050           0.4000
5       741.6            0.1791             0.5249           0.5355
6      1606.0            0.1442             0.2576           0.3784
7       897.0            0.1654             0.3682           0.2678
8       739.3            0.1703             0.5401           0.5390
9       711.4            0.1853             1.0580           1.1050

   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
5                0.1741          0.3985                  0.12440
6                0.1932          0.3063                  0.08368
7                0.1556          0.3196                  0.11510
8                0.2060          0.4378                  0.10720
9                0.2210          0.4366                  0.20750

[10 rows x 31 columns]
```

[5]:
```
df.columns
```

[5]:
```
Index(['id', 'diagnosis', 'radius_mean', 'texture_mean', 'perimeter_mean',
       'area_mean', 'smoothness_mean', 'compactness_mean', 'concavity_mean',
       'concave points_mean', 'symmetry_mean', 'fractal_dimension_mean',
       'radius_se', 'texture_se', 'perimeter_se', 'area_se', 'smoothness_se',
       'compactness_se', 'concavity_se', 'concave points_se', 'symmetry_se',
       'fractal_dimension_se', 'radius_worst', 'texture_worst',
       'perimeter_worst', 'area_worst', 'smoothness_worst',
```

```
       'compactness_worst', 'concavity_worst', 'concave points_worst',
       'symmetry_worst', 'fractal_dimension_worst', 'Unnamed: 32'],
      dtype='object')
```

[6]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 569 entries, 0 to 568
Data columns (total 33 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
 32  Unnamed: 32              0 non-null      float64
dtypes: float64(31), int64(1), object(1)
memory usage: 146.8+ KB
```

[7]:
```
df['Unnamed: 32']
```

[7]:
```
0     NaN
1     NaN
2     NaN
3     NaN
4     NaN
       ..
564   NaN
565   NaN
566   NaN
567   NaN
568   NaN
Name: Unnamed: 32, Length: 569, dtype: float64
```

[8]:
```
df = df.drop("Unnamed: 32", axis=1)
```

[71]:
```
df.head(40)
```

[71]:

	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	\
0	1	17.990	10.38	122.80	1001.0	
1	1	20.570	17.77	132.90	1326.0	
2	1	19.690	21.25	130.00	1203.0	
3	1	11.420	20.38	77.58	386.1	
4	1	20.290	14.34	135.10	1297.0	
5	1	12.450	15.70	82.57	477.1	
6	1	18.250	19.98	119.60	1040.0	
7	1	13.710	20.83	90.20	577.9	
8	1	13.000	21.82	87.50	519.8	
9	1	12.460	24.04	83.97	475.9	
10	1	16.020	23.24	102.70	797.8	
11	1	15.780	17.89	103.60	781.0	
12	1	19.170	24.80	132.40	1123.0	
13	1	15.850	23.95	103.70	782.7	
14	1	13.730	22.61	93.60	578.3	
15	1	14.540	27.54	96.73	658.8	
16	1	14.680	20.13	94.74	684.5	
17	1	16.130	20.68	108.10	798.8	
18	1	19.810	22.15	130.00	1260.0	
19	0	13.540	14.36	87.46	566.3	
20	0	13.080	15.71	85.63	520.0	
21	0	9.504	12.44	60.34	273.9	
22	1	15.340	14.26	102.50	704.4	
23	1	21.160	23.04	137.20	1404.0	
24	1	16.650	21.38	110.00	904.6	
25	1	17.140	16.40	116.00	912.7	
26	1	14.580	21.53	97.41	644.8	

27	1	18.610	20.25	122.10	1094.0
28	1	15.300	25.27	102.40	732.4
29	1	17.570	15.05	115.00	955.1
30	1	18.630	25.11	124.80	1088.0
31	1	11.840	18.70	77.93	440.6
32	1	17.020	23.98	112.80	899.3
33	1	19.270	26.47	127.90	1162.0
34	1	16.130	17.88	107.00	807.2
35	1	16.740	21.59	110.10	869.5
36	1	14.250	21.72	93.63	633.0
37	0	13.030	18.42	82.61	523.8
38	1	14.990	25.20	95.54	698.8
39	1	13.480	20.82	88.40	559.2

	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	
10	0.08206	0.06669	0.03299	0.03323	
11	0.09710	0.12920	0.09954	0.06606	
12	0.09740	0.24580	0.20650	0.11180	
13	0.08401	0.10020	0.09938	0.05364	
14	0.11310	0.22930	0.21280	0.08025	
15	0.11390	0.15950	0.16390	0.07364	
16	0.09867	0.07200	0.07395	0.05259	
17	0.11700	0.20220	0.17220	0.10280	
18	0.09831	0.10270	0.14790	0.09498	
19	0.09779	0.08129	0.06664	0.04781	
20	0.10750	0.12700	0.04568	0.03110	
21	0.10240	0.06492	0.02956	0.02076	
22	0.10730	0.21350	0.20770	0.09756	
23	0.09428	0.10220	0.10970	0.08632	
24	0.11210	0.14570	0.15250	0.09170	
25	0.11860	0.22760	0.22290	0.14010	
26	0.10540	0.18680	0.14250	0.08783	
27	0.09440	0.10660	0.14900	0.07731	
28	0.10820	0.16970	0.16830	0.08751	
29	0.09847	0.11570	0.09875	0.07953	
30	0.10640	0.18870	0.23190	0.12440	
31	0.11090	0.15160	0.12180	0.05182	

```
32          0.11970           0.14960         0.24170              0.12030
33          0.09401           0.17190         0.16570              0.07593
34          0.10400           0.15590         0.13540              0.07752
35          0.09610           0.13360         0.13480              0.06018
36          0.09823           0.10980         0.13190              0.05598
37          0.08983           0.03766         0.02562              0.02923
38          0.09387           0.05131         0.02398              0.02899
39          0.10160           0.12550         0.10630              0.05439

    symmetry_mean  …  radius_worst  texture_worst  perimeter_worst  \
0          0.2419  …         25.38          17.33           184.60
1          0.1812  …         24.99          23.41           158.80
2          0.2069  …         23.57          25.53           152.50
3          0.2597  …         14.91          26.50            98.87
4          0.1809  …         22.54          16.67           152.20
5          0.2087  …         15.47          23.75           103.40
6          0.1794  …         22.88          27.66           153.20
7          0.2196  …         17.06          28.14           110.60
8          0.2350  …         15.49          30.73           106.20
9          0.2030  …         15.09          40.68            97.65
10         0.1528  …         19.19          33.88           123.80
11         0.1842  …         20.42          27.28           136.50
12         0.2397  …         20.96          29.94           151.70
13         0.1847  …         16.84          27.66           112.00
14         0.2069  …         15.03          32.01           108.80
15         0.2303  …         17.46          37.13           124.10
16         0.1586  …         19.07          30.88           123.40
17         0.2164  …         20.96          31.48           136.80
18         0.1582  …         27.32          30.88           186.80
19         0.1885  …         15.11          19.26            99.70
20         0.1967  …         14.50          20.49            96.09
21         0.1815  …         10.23          15.66            65.13
22         0.2521  …         18.07          19.08           125.10
23         0.1769  …         29.17          35.59           188.00
24         0.1995  …         26.46          31.56           177.00
25         0.3040  …         22.25          21.40           152.40
26         0.2252  …         17.62          33.21           122.40
27         0.1697  …         21.31          27.26           139.90
28         0.1926  …         20.27          36.71           149.30
29         0.1739  …         20.01          19.52           134.90
30         0.2183  …         23.15          34.01           160.50
31         0.2301  …         16.82          28.12           119.40
32         0.2248  …         20.88          32.09           136.10
33         0.1853  …         24.15          30.90           161.40
34         0.1998  …         20.21          27.26           132.70
35         0.1896  …         20.01          29.02           133.50
36         0.1885  …         15.89          30.36           116.20
```

```
37         0.1467  …         13.30          22.81           84.46
38         0.1565  …         14.99          25.20           95.54
39         0.1720  …         15.53          26.02          107.30

    area_worst  smoothness_worst  compactness_worst  concavity_worst  \
0       2019.0           0.16220            0.66560          0.71190
1       1956.0           0.12380            0.18660          0.24160
2       1709.0           0.14440            0.42450          0.45040
3        567.7           0.20980            0.86630          0.68690
4       1575.0           0.13740            0.20500          0.40000
5        741.6           0.17910            0.52490          0.53550
6       1606.0           0.14420            0.25760          0.37840
7        897.0           0.16540            0.36820          0.26780
8        739.3           0.17030            0.54010          0.53900
9        711.4           0.18530            1.05800          1.10500
10      1150.0           0.11810            0.15510          0.14590
11      1299.0           0.13960            0.56090          0.39650
12      1332.0           0.10370            0.39030          0.36390
13       876.5           0.11310            0.19240          0.23220
14       697.7           0.16510            0.77250          0.69430
15       943.2           0.16780            0.65770          0.70260
16      1138.0           0.14640            0.18710          0.29140
17      1315.0           0.17890            0.42330          0.47840
18      2398.0           0.15120            0.31500          0.53720
19       711.2           0.14400            0.17730          0.23900
20       630.5           0.13120            0.27760          0.18900
21       314.9           0.13240            0.11480          0.08867
22       980.9           0.13900            0.59540          0.63050
23      2615.0           0.14010            0.26000          0.31550
24      2215.0           0.18050            0.35780          0.46950
25      1461.0           0.15450            0.39490          0.38530
26       896.9           0.15250            0.66430          0.55390
27      1403.0           0.13380            0.21170          0.34460
28      1269.0           0.16410            0.61100          0.63350
29      1227.0           0.12550            0.28120          0.24890
30      1670.0           0.14910            0.42570          0.61330
31       888.7           0.16370            0.57750          0.69560
32      1344.0           0.16340            0.35590          0.55880
33      1813.0           0.15090            0.65900          0.60910
34      1261.0           0.14460            0.58040          0.52740
35      1229.0           0.15630            0.38350          0.54090
36       799.6           0.14460            0.42380          0.51860
37       545.9           0.09701            0.04619          0.04833
38       698.8           0.09387            0.05131          0.02398
39       740.4           0.16100            0.42250          0.50300

    concave points_worst  symmetry_worst  fractal_dimension_worst
```

0	0.26540	0.4601	0.11890
1	0.18600	0.2750	0.08902
2	0.24300	0.3613	0.08758
3	0.25750	0.6638	0.17300
4	0.16250	0.2364	0.07678
5	0.17410	0.3985	0.12440
6	0.19320	0.3063	0.08368
7	0.15560	0.3196	0.11510
8	0.20600	0.4378	0.10720
9	0.22100	0.4366	0.20750
10	0.09975	0.2948	0.08452
11	0.18100	0.3792	0.10480
12	0.17670	0.3176	0.10230
13	0.11190	0.2809	0.06287
14	0.22080	0.3596	0.14310
15	0.17120	0.4218	0.13410
16	0.16090	0.3029	0.08216
17	0.20730	0.3706	0.11420
18	0.23880	0.2768	0.07615
19	0.12880	0.2977	0.07259
20	0.07283	0.3184	0.08183
21	0.06227	0.2450	0.07773
22	0.23930	0.4667	0.09946
23	0.20090	0.2822	0.07526
24	0.20950	0.3613	0.09564
25	0.25500	0.4066	0.10590
26	0.27010	0.4264	0.12750
27	0.14900	0.2341	0.07421
28	0.20240	0.4027	0.09876
29	0.14560	0.2756	0.07919
30	0.18480	0.3444	0.09782
31	0.15460	0.4761	0.14020
32	0.18470	0.3530	0.08482
33	0.17850	0.3672	0.11230
34	0.18640	0.4270	0.12330
35	0.18130	0.4863	0.08633
36	0.14470	0.3591	0.10140
37	0.05013	0.1987	0.06169
38	0.02899	0.1565	0.05504
39	0.22580	0.2807	0.10710

[40 rows x 31 columns]

[10]:
```
df.columns
```

[10]:
```
Index(['id', 'diagnosis', 'radius_mean', 'texture_mean', 'perimeter_mean',
       'area_mean', 'smoothness_mean', 'compactness_mean', 'concavity_mean',
```

```
       'concave points_mean', 'symmetry_mean', 'fractal_dimension_mean',
       'radius_se', 'texture_se', 'perimeter_se', 'area_se', 'smoothness_se',
       'compactness_se', 'concavity_se', 'concave points_se', 'symmetry_se',
       'fractal_dimension_se', 'radius_worst', 'texture_worst',
       'perimeter_worst', 'area_worst', 'smoothness_worst',
       'compactness_worst', 'concavity_worst', 'concave points_worst',
       'symmetry_worst', 'fractal_dimension_worst'],
      dtype='object')
```

[11]:
```
df.drop('id', axis=1, inplace=True)
# df = df.drop('id', axis=1)
```

[12]:
```
df.columns
```

[12]:
```
Index(['diagnosis', 'radius_mean', 'texture_mean', 'perimeter_mean',
       'area_mean', 'smoothness_mean', 'compactness_mean', 'concavity_mean',
       'concave points_mean', 'symmetry_mean', 'fractal_dimension_mean',
       'radius_se', 'texture_se', 'perimeter_se', 'area_se', 'smoothness_se',
       'compactness_se', 'concavity_se', 'concave points_se', 'symmetry_se',
       'fractal_dimension_se', 'radius_worst', 'texture_worst',
       'perimeter_worst', 'area_worst', 'smoothness_worst',
       'compactness_worst', 'concavity_worst', 'concave points_worst',
       'symmetry_worst', 'fractal_dimension_worst'],
      dtype='object')
```

[13]:
```
type(df.columns)
```

[13]:
```
pandas.core.indexes.base.Index
```

[14]:
```
l = list(df.columns)
print(l)
```

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

[15]:
```
features_mean = l[1:11]

features_se = l[11:21]

features_worst = l[21:]
```

[16]:
```
print(features_mean)
```

```
['radius_mean', 'texture_mean', 'perimeter_mean', 'area_mean',
'smoothness_mean', 'compactness_mean', 'concavity_mean', 'concave points_mean',
'symmetry_mean', 'fractal_dimension_mean']
```

[17]:
```
print(features_se)
```

```
['radius_se', 'texture_se', 'perimeter_se', 'area_se', 'smoothness_se',
'compactness_se', 'concavity_se', 'concave points_se', 'symmetry_se',
'fractal_dimension_se']
```

[18]:
```
print(features_worst)
```

```
['radius_worst', 'texture_worst', 'perimeter_worst', 'area_worst',
'smoothness_worst', 'compactness_worst', 'concavity_worst', 'concave
points_worst', 'symmetry_worst', 'fractal_dimension_worst']
```

[19]:
```
df.head(2)
```

[19]:

	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	\
0	M	17.99	10.38	122.8	1001.0	
1	M	20.57	17.77	132.9	1326.0	

	smoothness_mean	compactness_mean	concavity_mean	concave points_mean	\
0	0.11840	0.27760	0.3001	0.14710	
1	0.08474	0.07864	0.0869	0.07017	

	symmetry_mean	…	radius_worst	texture_worst	perimeter_worst	\
0	0.2419	…	25.38	17.33	184.6	
1	0.1812	…	24.99	23.41	158.8	

	area_worst	smoothness_worst	compactness_worst	concavity_worst	\
0	2019.0	0.1622	0.6656	0.7119	
1	1956.0	0.1238	0.1866	0.2416	

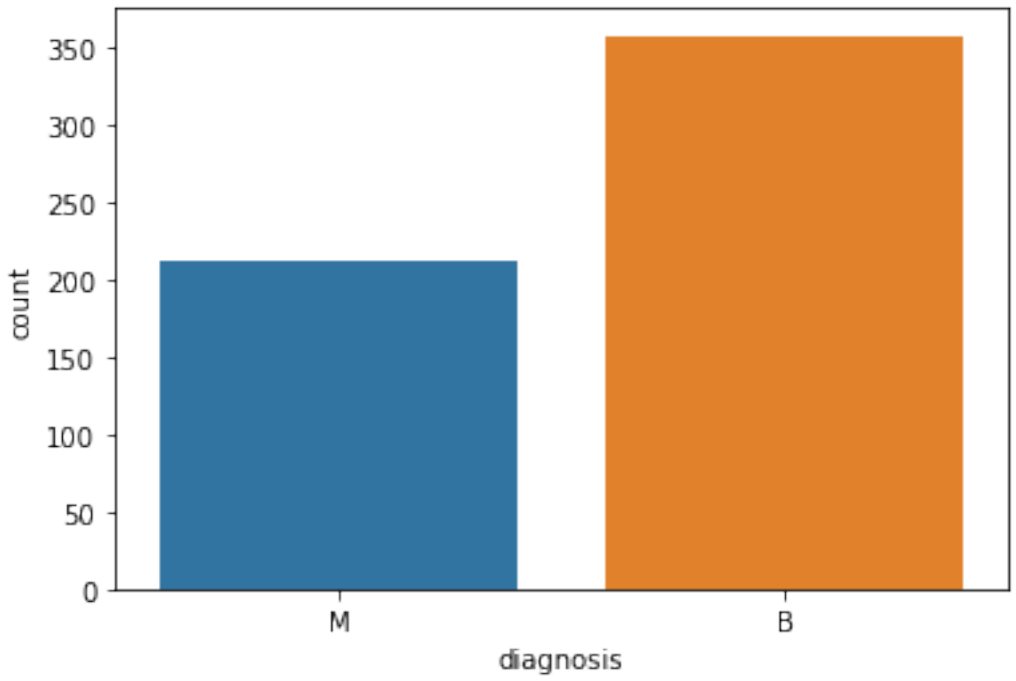
	concave points_worst	symmetry_worst	fractal_dimension_worst
0	0.2654	0.4601	0.11890
1	0.1860	0.2750	0.08902

[2 rows x 31 columns]

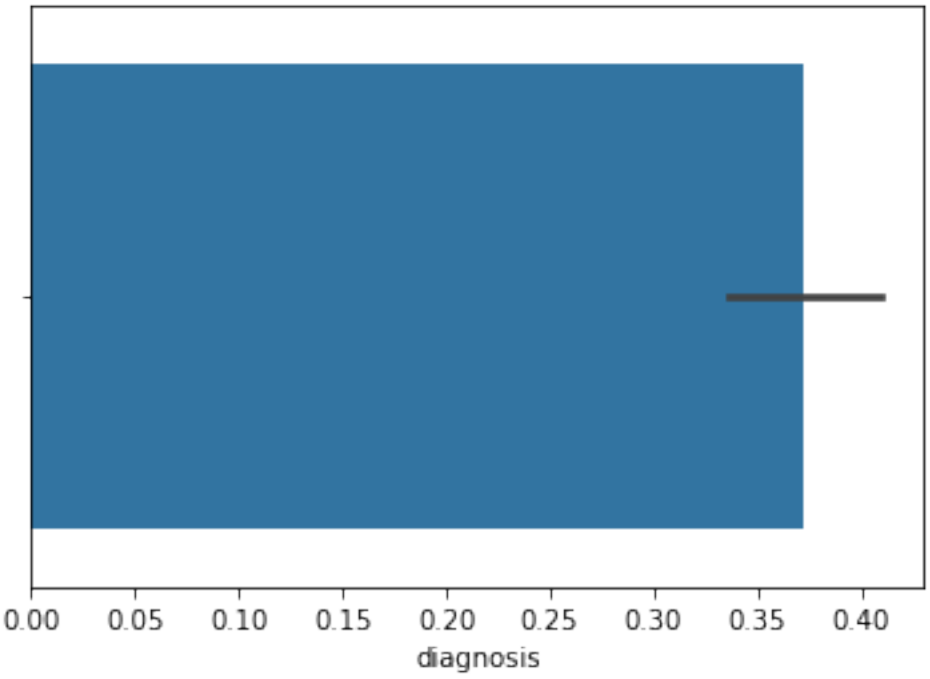
[20]:
```
df['diagnosis'].unique()
# M= Malignant, B= Benign
```

[20]:
```
array(['M', 'B'], dtype=object)
```

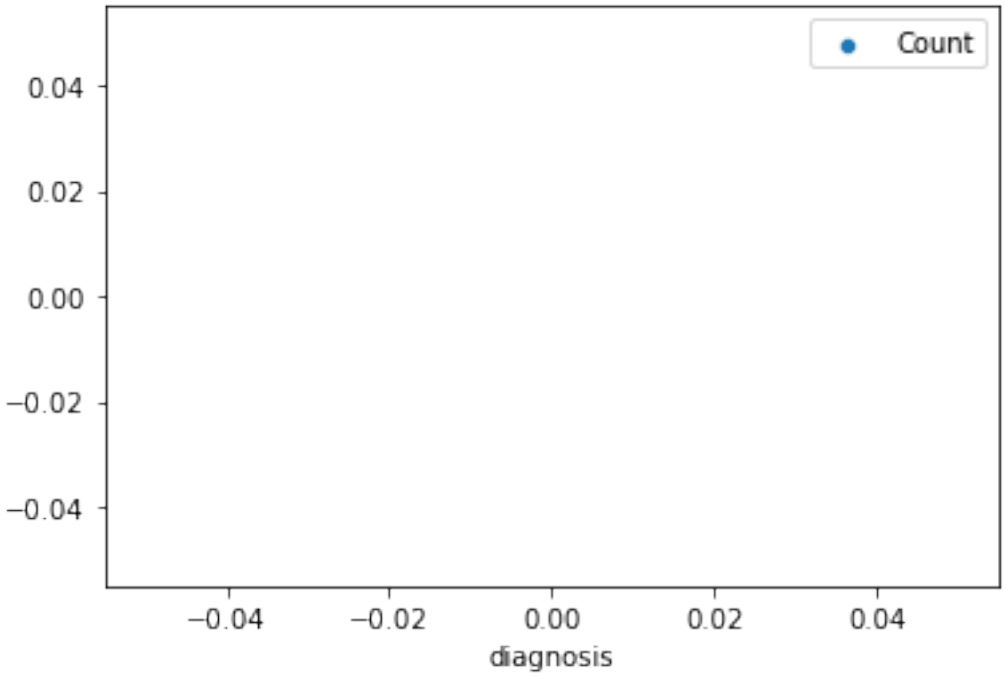
[21]:
```
sns.countplot(df['diagnosis'], label="Count",);
```



[72]:
```
sns.barplot(df['diagnosis'], label="Count",);
```



[76]:
```
sns.scatterplot(df['diagnosis'], label="Count",);
```



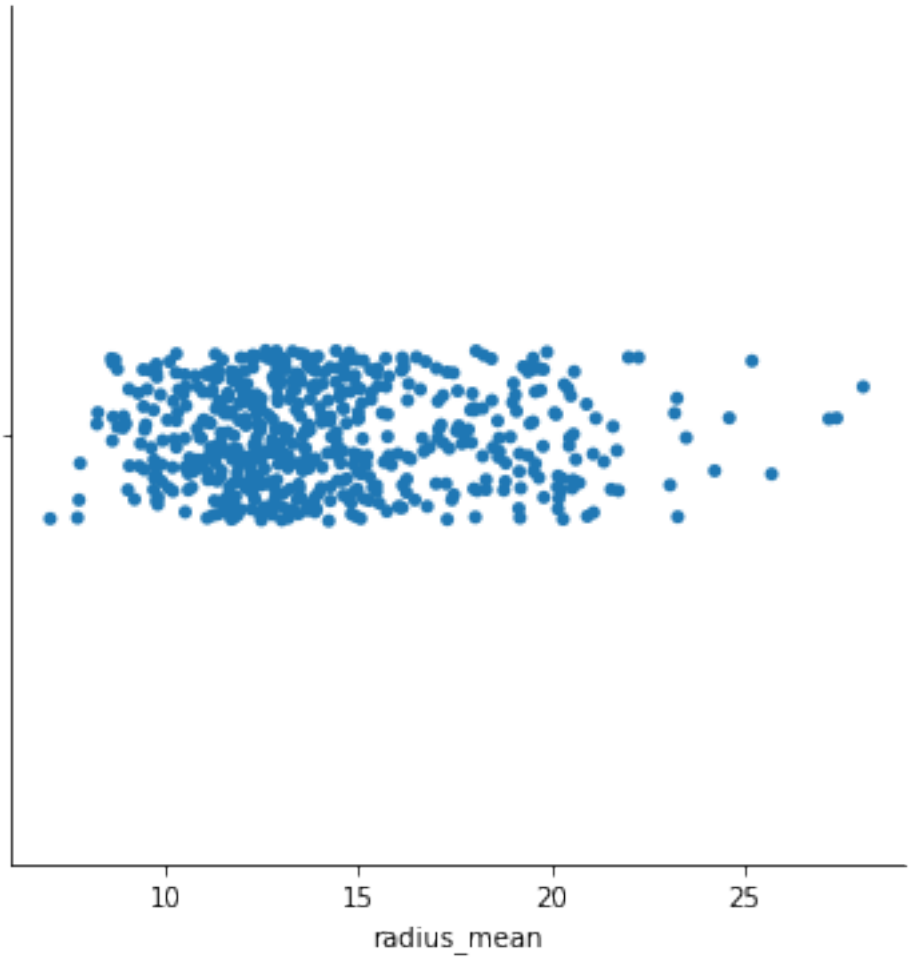
[22]:
```
df['diagnosis'].value_counts()
```

[22]:
```
B    357
M    212
Name: diagnosis, dtype: int64
```

[23]:
```
df.shape
```

[23]:
```
(569, 31)
```

[77]:
```
sns.catplot(x='radius_mean',data=df);
```



1 Explore the data

[24]:
```
df.describe()
# summary of all the numeric columns
```

[24]:

	radius_mean	texture_mean	perimeter_mean	area_mean	\
count	569.000000	569.000000	569.000000	569.000000	
mean	14.127292	19.289649	91.969033	654.889104	
std	3.524049	4.301036	24.298981	351.914129	
min	6.981000	9.710000	43.790000	143.500000	
25%	11.700000	16.170000	75.170000	420.300000	
50%	13.370000	18.840000	86.240000	551.100000	
75%	15.780000	21.800000	104.100000	782.700000	
max	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  fractal_dimension_mean  …  radius_worst  \
count     569.000000              569.000000  …    569.000000
mean        0.181162                0.062798  …     16.269190
std         0.027414                0.007060  …      4.833242
min         0.106000                0.049960  …      7.930000
25%         0.161900                0.057700  …     13.010000
50%         0.179200                0.061540  …     14.970000
75%         0.195700                0.066120  …     18.790000
max         0.304000                0.097440  …     36.040000

       texture_worst  perimeter_worst   area_worst  smoothness_worst  \
count     569.000000       569.000000   569.000000        569.000000
mean       25.677223       107.261213   880.583128          0.132369
std         6.146258        33.602542   569.356993          0.022832
min        12.020000        50.410000   185.200000          0.071170
25%        21.080000        84.110000   515.300000          0.116600
50%        25.410000        97.660000   686.500000          0.131300
75%        29.720000       125.400000  1084.000000          0.146000
max        49.540000       251.200000  4254.000000          0.222600

       compactness_worst  concavity_worst  concave points_worst  \
count         569.000000       569.000000            569.000000
mean            0.254265         0.272188              0.114606
std             0.157336         0.208624              0.065732
min             0.027290         0.000000              0.000000
25%             0.147200         0.114500              0.064930
50%             0.211900         0.226700              0.099930
75%             0.339100         0.382900              0.161400
max             1.058000         1.252000              0.291000

       symmetry_worst  fractal_dimension_worst
count      569.000000               569.000000
mean         0.290076                 0.083946
std          0.061867                 0.018061
min          0.156500                 0.055040
25%          0.250400                 0.071460
50%          0.282200                 0.080040
75%          0.317900                 0.092080
```

```
max          0.663800                 0.207500

[8 rows x 30 columns]
```

[25]:
```
len(df.columns)
```

[25]:
```
31
```

[26]:
```
# Correlation Plot
corr = df.corr()
corr
```

[26]:

	radius_mean	texture_mean	perimeter_mean	area_mean	\
radius_mean	1.000000	0.323782	0.997855	0.987357	
texture_mean	0.323782	1.000000	0.329533	0.321086	
perimeter_mean	0.997855	0.329533	1.000000	0.986507	
area_mean	0.987357	0.321086	0.986507	1.000000	
smoothness_mean	0.170581	-0.023389	0.207278	0.177028	
compactness_mean	0.506124	0.236702	0.556936	0.498502	
concavity_mean	0.676764	0.302418	0.716136	0.685983	
concave points_mean	0.822529	0.293464	0.850977	0.823269	
symmetry_mean	0.147741	0.071401	0.183027	0.151293	
fractal_dimension_mean	-0.311631	-0.076437	-0.261477	-0.283110	
radius_se	0.679090	0.275869	0.691765	0.732562	
texture_se	-0.097317	0.386358	-0.086761	-0.066280	
perimeter_se	0.674172	0.281673	0.693135	0.726628	
area_se	0.735864	0.259845	0.744983	0.800086	
smoothness_se	-0.222600	0.006614	-0.202694	-0.166777	
compactness_se	0.206000	0.191975	0.250744	0.212583	
concavity_se	0.194204	0.143293	0.228082	0.207660	
concave points_se	0.376169	0.163851	0.407217	0.372320	
symmetry_se	-0.104321	0.009127	-0.081629	-0.072497	
fractal_dimension_se	-0.042641	0.054458	-0.005523	-0.019887	
radius_worst	0.969539	0.352573	0.969476	0.962746	
texture_worst	0.297008	0.912045	0.303038	0.287489	
perimeter_worst	0.965137	0.358040	0.970387	0.959120	
area_worst	0.941082	0.343546	0.941550	0.959213	
smoothness_worst	0.119616	0.077503	0.150549	0.123523	
compactness_worst	0.413463	0.277830	0.455774	0.390410	
concavity_worst	0.526911	0.301025	0.563879	0.512606	
concave points_worst	0.744214	0.295316	0.771241	0.722017	
symmetry_worst	0.163953	0.105008	0.189115	0.143570	
fractal_dimension_worst	0.007066	0.119205	0.051019	0.003738	

	smoothness_mean	compactness_mean	concavity_mean	\
radius_mean	0.170581	0.506124	0.676764	
texture_mean	-0.023389	0.236702	0.302418	

perimeter_mean	0.207278	0.556936	0.716136
area_mean	0.177028	0.498502	0.685983
smoothness_mean	1.000000	0.659123	0.521984
compactness_mean	0.659123	1.000000	0.883121
concavity_mean	0.521984	0.883121	1.000000
concave points_mean	0.553695	0.831135	0.921391
symmetry_mean	0.557775	0.602641	0.500667
fractal_dimension_mean	0.584792	0.565369	0.336783
radius_se	0.301467	0.497473	0.631925
texture_se	0.068406	0.046205	0.076218
perimeter_se	0.296092	0.548905	0.660391
area_se	0.246552	0.455653	0.617427
smoothness_se	0.332375	0.135299	0.098564
compactness_se	0.318943	0.738722	0.670279
concavity_se	0.248396	0.570517	0.691270
concave points_se	0.380676	0.642262	0.683260
symmetry_se	0.200774	0.229977	0.178009
fractal_dimension_se	0.283607	0.507318	0.449301
radius_worst	0.213120	0.535315	0.688236
texture_worst	0.036072	0.248133	0.299879
perimeter_worst	0.238853	0.590210	0.729565
area_worst	0.206718	0.509604	0.675987
smoothness_worst	0.805324	0.565541	0.448822
compactness_worst	0.472468	0.865809	0.754968
concavity_worst	0.434926	0.816275	0.884103
concave points_worst	0.503053	0.815573	0.861323
symmetry_worst	0.394309	0.510223	0.409464
fractal_dimension_worst	0.499316	0.687382	0.514930

	concave points_mean	symmetry_mean	\
radius_mean	0.822529	0.147741	
texture_mean	0.293464	0.071401	
perimeter_mean	0.850977	0.183027	
area_mean	0.823269	0.151293	
smoothness_mean	0.553695	0.557775	
compactness_mean	0.831135	0.602641	
concavity_mean	0.921391	0.500667	
concave points_mean	1.000000	0.462497	
symmetry_mean	0.462497	1.000000	
fractal_dimension_mean	0.166917	0.479921	
radius_se	0.698050	0.303379	
texture_se	0.021480	0.128053	
perimeter_se	0.710650	0.313893	
area_se	0.690299	0.223970	
smoothness_se	0.027653	0.187321	
compactness_se	0.490424	0.421659	
concavity_se	0.439167	0.342627	

```
concave points_se                  0.615634       0.393298
symmetry_se                        0.095351       0.449137
fractal_dimension_se               0.257584       0.331786
radius_worst                       0.830318       0.185728
texture_worst                      0.292752       0.090651
perimeter_worst                    0.855923       0.219169
area_worst                         0.809630       0.177193
smoothness_worst                   0.452753       0.426675
compactness_worst                  0.667454       0.473200
concavity_worst                    0.752399       0.433721
concave points_worst               0.910155       0.430297
symmetry_worst                     0.375744       0.699826
fractal_dimension_worst            0.368661       0.438413

                         fractal_dimension_mean  …  radius_worst  \
radius_mean                           -0.311631  …      0.969539
texture_mean                          -0.076437  …      0.352573
perimeter_mean                        -0.261477  …      0.969476
area_mean                             -0.283110  …      0.962746
smoothness_mean                        0.584792  …      0.213120
compactness_mean                       0.565369  …      0.535315
concavity_mean                         0.336783  …      0.688236
concave points_mean                    0.166917  …      0.830318
symmetry_mean                          0.479921  …      0.185728
fractal_dimension_mean                 1.000000  …     -0.253691
radius_se                              0.000111  …      0.715065
texture_se                             0.164174  …     -0.111690
perimeter_se                           0.039830  …      0.697201
area_se                               -0.090170  …      0.757373
smoothness_se                          0.401964  …     -0.230691
compactness_se                         0.559837  …      0.204607
concavity_se                           0.446630  …      0.186904
concave points_se                      0.341198  …      0.358127
symmetry_se                            0.345007  …     -0.128121
fractal_dimension_se                   0.688132  …     -0.037488
radius_worst                          -0.253691  …      1.000000
texture_worst                         -0.051269  …      0.359921
perimeter_worst                       -0.205151  …      0.993708
area_worst                            -0.231854  …      0.984015
smoothness_worst                       0.504942  …      0.216574
compactness_worst                      0.458798  …      0.475820
concavity_worst                        0.346234  …      0.573975
concave points_worst                   0.175325  …      0.787424
symmetry_worst                         0.334019  …      0.243529
fractal_dimension_worst                0.767297  …      0.093492

                         texture_worst  perimeter_worst  area_worst  \
```

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	\
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	\
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	

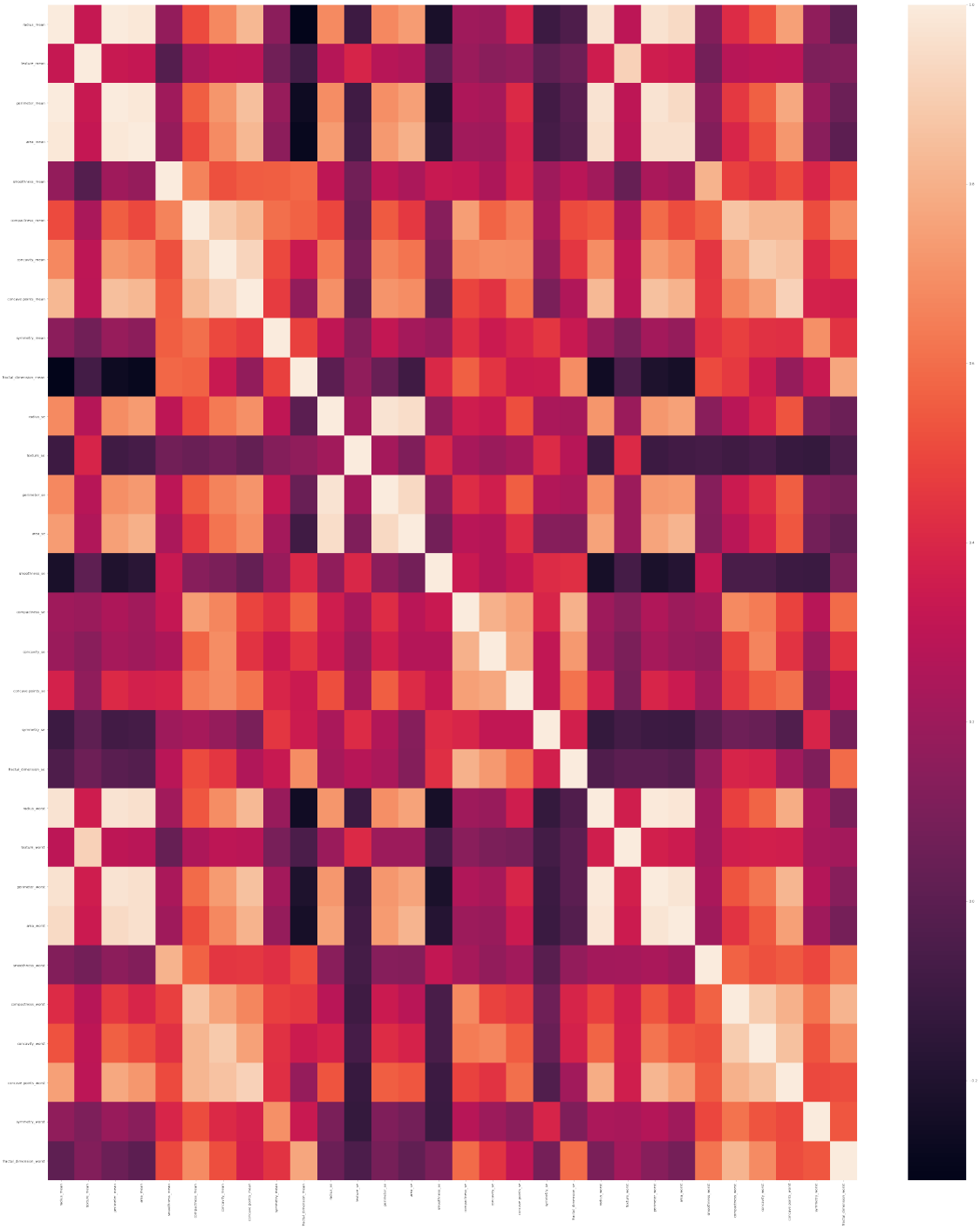
```
                         fractal_dimension_worst
radius_mean                             0.007066
texture_mean                            0.119205
perimeter_mean                          0.051019
area_mean                               0.003738
smoothness_mean                         0.499316
compactness_mean                        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

[30 rows x 30 columns]
```

[27]:
```
corr.shape
```

[27]:
```
(30, 30)
```

[82]:
```
plt.figure(figsize=(52,62))
sns.heatmap(corr);
```



[29]:
```
df.head()
```

[29]:

	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	\
0	M	17.99	10.38	122.80	1001.0	
1	M	20.57	17.77	132.90	1326.0	

```
2         M        19.69         21.25          130.00     1203.0
3         M        11.42         20.38           77.58      386.1
4         M        20.29         14.34          135.10     1297.0

   smoothness_mean  compactness_mean  concavity_mean  concave points_mean  \
0          0.11840           0.27760          0.3001              0.14710
1          0.08474           0.07864          0.0869              0.07017
2          0.10960           0.15990          0.1974              0.12790
3          0.14250           0.28390          0.2414              0.10520
4          0.10030           0.13280          0.1980              0.10430

   symmetry_mean  …  radius_worst  texture_worst  perimeter_worst  \
0         0.2419  …         25.38          17.33           184.60
1         0.1812  …         24.99          23.41           158.80
2         0.2069  …         23.57          25.53           152.50
3         0.2597  …         14.91          26.50            98.87
4         0.1809  …         22.54          16.67           152.20

   area_worst  smoothness_worst  compactness_worst  concavity_worst  \
0      2019.0            0.1622             0.6656           0.7119
1      1956.0            0.1238             0.1866           0.2416
2      1709.0            0.1444             0.4245           0.4504
3       567.7            0.2098             0.8663           0.6869
4      1575.0            0.1374             0.2050           0.4000

   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

[5 rows x 31 columns]
```

[30]:
```
df['diagnosis'] = df['diagnosis'].map({'M':1, 'B':0})
```

[31]:
```
df.head()
```

[31]:
```
   diagnosis  radius_mean  texture_mean  perimeter_mean  area_mean  \
0          1        17.99         10.38          122.80     1001.0
1          1        20.57         17.77          132.90     1326.0
2          1        19.69         21.25          130.00     1203.0
3          1        11.42         20.38           77.58      386.1
4          1        20.29         14.34          135.10     1297.0

   smoothness_mean  compactness_mean  concavity_mean  concave points_mean  \
0          0.11840           0.27760          0.3001              0.14710
```

```
1          0.08474           0.07864          0.0869               0.07017
2          0.10960           0.15990          0.1974               0.12790
3          0.14250           0.28390          0.2414               0.10520
4          0.10030           0.13280          0.1980               0.10430

   symmetry_mean  …  radius_worst  texture_worst  perimeter_worst  \
0         0.2419  …         25.38          17.33           184.60
1         0.1812  …         24.99          23.41           158.80
2         0.2069  …         23.57          25.53           152.50
3         0.2597  …         14.91          26.50            98.87
4         0.1809  …         22.54          16.67           152.20

   area_worst  smoothness_worst  compactness_worst  concavity_worst  \
0      2019.0            0.1622             0.6656           0.7119
1      1956.0            0.1238             0.1866           0.2416
2      1709.0            0.1444             0.4245           0.4504
3       567.7            0.2098             0.8663           0.6869
4      1575.0            0.1374             0.2050           0.4000

   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

[5 rows x 31 columns]
```

[32]:
```
df['diagnosis'].unique()
```

[32]:
```
array([1, 0])
```

[33]:
```
X = df.drop('diagnosis', axis=1)
X.head()
```

[33]:
```
   radius_mean  texture_mean  perimeter_mean  area_mean  smoothness_mean  \
0        17.99         10.38          122.80     1001.0          0.11840
1        20.57         17.77          132.90     1326.0          0.08474
2        19.69         21.25          130.00     1203.0          0.10960
3        11.42         20.38           77.58      386.1          0.14250
4        20.29         14.34          135.10     1297.0          0.10030

   compactness_mean  concavity_mean  concave points_mean  symmetry_mean  \
0           0.27760          0.3001              0.14710         0.2419
1           0.07864          0.0869              0.07017         0.1812
2           0.15990          0.1974              0.12790         0.2069
3           0.28390          0.2414              0.10520         0.2597
```

```
4           0.13280          0.1980               0.10430         0.1809

   fractal_dimension_mean  …  radius_worst  texture_worst  perimeter_worst  \
0                 0.07871  …         25.38          17.33           184.60
1                 0.05667  …         24.99          23.41           158.80
2                 0.05999  …         23.57          25.53           152.50
3                 0.09744  …         14.91          26.50            98.87
4                 0.05883  …         22.54          16.67           152.20

   area_worst  smoothness_worst  compactness_worst  concavity_worst  \
0      2019.0            0.1622             0.6656           0.7119
1      1956.0            0.1238             0.1866           0.2416
2      1709.0            0.1444             0.4245           0.4504
3       567.7            0.2098             0.8663           0.6869
4      1575.0            0.1374             0.2050           0.4000

   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

[5 rows x 30 columns]
```

[34]:
```
y = df['diagnosis']
y.head()
```

[34]:
```
0    1
1    1
2    1
3    1
4    1
Name: diagnosis, dtype: int64
```

[35]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3)
```

[36]:
```
df.shape
```

[36]:
```
(569, 31)
```

[37]:
```
X_train.shape
```

[37]:
```
(398, 30)
```

[38]:
```
X_test.shape
```

[38]: (171, 30)

[39]:
```
y_train.shape
```

[39]: (398,)

[40]:
```
y_test.shape
```

[40]: (171,)

[41]:
```
X_train.head(1)
```

[41]:
```
     radius_mean  texture_mean  perimeter_mean  area_mean  smoothness_mean  \
440        10.97          17.2           71.73      371.5          0.08915

     compactness_mean  concavity_mean  concave points_mean  symmetry_mean  \
440            0.1113         0.09457              0.03613         0.1489

     fractal_dimension_mean  …  radius_worst  texture_worst  \
440                  0.0664  …         12.36          26.87

     perimeter_worst  area_worst  smoothness_worst  compactness_worst  \
440            90.14       476.4            0.1391             0.4082

     concavity_worst  concave points_worst  symmetry_worst  \
440           0.4779                0.1555           0.254

     fractal_dimension_worst
440                  0.09532

[1 rows x 30 columns]
```

[42]:
```
from sklearn.preprocessing import StandardScaler
ss = StandardScaler()
X_train = ss.fit_transform(X_train)
X_test = ss.transform(X_test)
```

[43]:
```
X_train
```

[43]:
```
array([[-0.86220204, -0.49899522, -0.80165717, …,  0.62972758,
        -0.6246113 ,  0.5776332 ],
       [ 0.16453422, -0.84880234,  0.05709313, …, -0.87959881,
        -0.6419361 , -1.48081868],
       [-1.32395128,  0.59907482, -1.32305591, …, -1.38322431,
        -0.80652171, -0.56988532],
       …,
       [-0.09214985, -0.39706467, -0.14457997, …, -0.37364818,
```

```
 -0.66272587, -0.8021787 ],
[-0.77476021, -1.38162114, -0.74959927, …, -0.44588255,
 -1.38343758, -0.89659817],
[-0.64500783, -1.08046269, -0.61392082, …, -0.14376926,
 -0.51373258,  0.76378978]])
```

2 Machine Learning Models

3 Logistic Regression

[44]:
```
from sklearn.linear_model import LogisticRegression
lr = LogisticRegression()
lr.fit(X_train, y_train)
```

[44]:
```
LogisticRegression()
```

[45]:
```
y_pred = lr.predict(X_test)
```

[46]:
```
y_pred
```

[46]:
```
array([0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0,
       1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1,
       0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0,
       0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0,
       1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 0,
       1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1])
```

[47]:
```
y_test
```

[47]:
```
528    0
448    0
394    0
523    0
473    0
      ..
416    0
285    0
351    1
471    0
373    1
Name: diagnosis, Length: 171, dtype: int64
```

[48]:
```
from sklearn.metrics import accuracy_score
print(accuracy_score(y_test, y_pred))
```

```
0.9415204678362573
```

[49]:
```
lr_acc = accuracy_score(y_test, y_pred)
print(lr_acc)
```

```
0.9415204678362573
```

[50]:
```
results = pd.DataFrame()
results
```

[50]:
```
Empty DataFrame
Columns: []
Index: []
```

[51]:
```
tempResults = pd.DataFrame({'Algorithm':['Logistic Regression Method'],␣
↪'Accuracy':[lr_acc]})
results = pd.concat( [results, tempResults] )
results = results[['Algorithm','Accuracy']]
results
```

[51]:

	Algorithm	Accuracy
0	Logistic Regression Method	0.94152

4 Decision Tree Classifier

[52]:
```
from sklearn.tree import DecisionTreeClassifier
dtc = DecisionTreeClassifier()
dtc.fit(X_train, y_train)
```

[52]:
```
DecisionTreeClassifier()
```

[53]:
```
y_pred = dtc.predict(X_test)
y_pred
```

[53]:
```
array([0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0,
       1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1,
       0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 0,
       1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1])
```

[54]:
```
from sklearn.metrics import accuracy_score
print(accuracy_score(y_test, y_pred))
```

```
0.935672514619883
```

[55]:
```
dtc_acc = accuracy_score(y_test, y_pred)
print(dtc_acc)
```

```
0.935672514619883
```

[56]:
```
tempResults = pd.DataFrame({'Algorithm':['Decision tree Classifier Method'],␣
↪'Accuracy':[dtc_acc]})
results = pd.concat( [results, tempResults] )
results = results[['Algorithm','Accuracy']]
results
```

[56]:

	Algorithm	Accuracy
0	Logistic Regression Method	0.941520
0	Decision tree Classifier Method	0.935673

5 Random Forest Classifier

[57]:
```
from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier()
rfc.fit(X_train, y_train)
```

[57]:
```
RandomForestClassifier()
```

[58]:
```
y_pred = rfc.predict(X_test)
y_pred
```

[58]:
```
array([0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0,
       1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1,
       0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0,
       1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0,
       1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1])
```

[59]:
```
from sklearn.metrics import accuracy_score
print(accuracy_score(y_test, y_pred))
```

```
0.9824561403508771
```

[60]:
```
rfc_acc = accuracy_score(y_test, y_pred)
print(rfc_acc)
```

```
0.9824561403508771
```

[61]:
```
tempResults = pd.DataFrame({'Algorithm':['Random Forest Classifier Method'],␣
↪'Accuracy':[rfc_acc]})
results = pd.concat( [results, tempResults] )
results = results[['Algorithm','Accuracy']]
results
```

[61]:

	Algorithm	Accuracy
0	Logistic Regression Method	0.941520
0	Decision tree Classifier Method	0.935673
0	Random Forest Classifier Method	0.982456

6 Support Vector Classifier

[62]:
```
from sklearn import svm
svc = svm.SVC()
svc.fit(X_train,y_train)
```

[62]:
```
SVC()
```

[63]:
```
y_pred = svc.predict(X_test)
y_pred
```

[63]:
```
array([0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 0, 0,
       1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1,
       0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0,
       1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0,
       1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1])
```

[64]:
```
from sklearn.metrics import accuracy_score
print(accuracy_score(y_test, y_pred))
```

```
0.9649122807017544
```

[65]:
```
svc_acc = accuracy_score(y_test, y_pred)
print(svc_acc)
```

```
0.9649122807017544
```

[66]:
```
tempResults = pd.DataFrame({'Algorithm':['Support Vector Classifier Method'],␣
↪'Accuracy':[svc_acc]})
results = pd.concat( [results, tempResults] )
results = results[['Algorithm','Accuracy']]
results
```

[66]:

	Algorithm	Accuracy
0	Logistic Regression Method	0.941520
0	Decision tree Classifier Method	0.935673
0	Random Forest Classifier Method	0.982456
0	Support Vector Classifier Method	0.964912

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]: