# **Breast Cancer regression using MLR**

The decided features were:

a) radius (mean of distances from center to points on the perimeter)

b) texture (standard deviation of gray-scale values)

c) perimeter

d) area

e) smoothness (local variation in radius lengths)

f) compactness (perimeter^2 / area - 1.0)

g) concavity (severity of concave portions of the contour)

h) concave points (number of concave portions of the contour)

i) symmetry

j) fractal dimension ("coastline approximation" - 1)

أ) دائرة نصف قطرها (يعني المسافات من مركز إلى نقاط على المحيط)

ب) الملمس (الانحراف المعياري للقيم النطاق الرمادي)

ج) محيط

د) المنطقة

ه) نعومة (تباين محلي في أطوال دائرة نصف قطرها)

و) الارتياج (محيط ^ 2 / المنطقة - 1.0)

ز) concavity (شدة أجزاء مقعرة من المحيط)

ح) النقاط المقصرة (عدد الأجزاء المقررة من المحيط)

ط) التماثل

ي) البعد كسور ("التقريب الساحلي" - 1)

## **Results:**

On the single linear regression, the results were:

r2 socre is: 0.7234520971661356

mean\_sqrd\_error is== 0.06545512493109215

root\_mean\_squared error of is== 0.2558419921183623.

so we will try to optimized the results using MLR.

We applied MLR on csv file contains some features decided the type of breast mass, and we noticed this results after firstly OLS:

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.0931 0.427 -4.901 0.000 -2.932 -1.254

x1 -0.2279 0.174 -1.311 0.190 -0.569 0.113

x2 0.0046 0.008 0.578 0.563 -0.011 0.020

x3 0.0247 0.025 0.984 0.326 -0.025 0.074

x4 0.0003 0.001 0.652 0.515 -0.001 0.001

x5 -0.2614 2.012 -0.130 0.897 -4.214 3.692

x6 -4.6648 1.314 -3.551 0.000 -7.245 -2.084

x7 1.2424 1.044 1.189 0.235 -0.809 3.294

x8 2.3386 1.980 1.181 0.238 -1.552 6.229

x9 0.0799 0.744 0.107 0.915 -1.382 1.542

x10 5.4997 4.688 1.173 0.241 -3.709 14.709

x11 0.3729 0.309 1.206 0.228 -0.235 0.980

x12 -0.0077 0.037 -0.207 0.836 -0.080 0.065

x13 -0.0178 0.041 -0.433 0.665 -0.099 0.063

x14 -0.0007 0.001 -0.485 0.628 -0.003 0.002

x15 15.6748 6.638 2.361 0.019 2.635 28.714

x16 -0.9354 2.102 -0.445 0.657 -5.064 3.194

x17 -3.9411 1.287 -3.063 0.002 -6.469 -1.414

x18 10.7650 5.462 1.971 0.049 0.035 21.495

x19 1.0495 2.710 0.387 0.699 -4.273 6.372

x20 4.5076 9.751 0.462 0.644 -14.647 23.662

x21 0.2101 0.057 3.655 0.000 0.097 0.323

x22 0.0073 0.007 1.046 0.296 -0.006 0.021

x23 -0.0039 0.006 -0.659 0.510 -0.015 0.008

x24 -0.0011 0.000 -3.297 0.001 -0.002 -0.000

x25 0.7988 1.431 0.558 0.577 -2.011 3.609

x26 0.4237 0.329 1.288 0.198 -0.223 1.070

x27 0.4936 0.262 1.885 0.060 -0.021 1.008

x28 0.3677 0.915 0.402 0.688 -1.429 2.164

x29 0.6279 0.494 1.272 0.204 -0.342 1.598

==============================================================================

Omnibus: 30.729 Durbin-Watson: 1.783

Prob(Omnibus): 0.000 Jarque-Bera (JB): 34.215

Skew: 0.582 Prob(JB): 3.72e-08

Kurtosis: 3.296 Cond. No. 1.33e+06

Which are indicated for the highest value and its impact less than the number of the number x9 that indicates to symmetry\_mean feature based on P>|t| values, so we will re-compute the values of OLS regressor.

After delete x9 or symmetry\_mean feature, the next table show the results

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.0860 0.421 -4.949 0.000 -2.914 -1.258

x1 -0.2267 0.173 -1.308 0.191 -0.567 0.114

x2 0.0047 0.008 0.599 0.550 -0.011 0.020

x3 0.0247 0.025 0.982 0.326 -0.025 0.074

x4 0.0003 0.001 0.646 0.519 -0.001 0.001

x5 -0.2295 1.988 -0.115 0.908 -4.135 3.676

x6 -4.6412 1.294 -3.586 0.000 -7.183 -2.099

x7 1.2468 1.043 1.196 0.232 -0.801 3.295

x8 2.3382 1.979 1.182 0.238 -1.548 6.225

x9 5.4709 4.676 1.170 0.243 -3.715 14.657

x10 0.3775 0.306 1.234 0.218 -0.223 0.978

x11 -0.0069 0.036 -0.191 0.849 -0.078 0.064

x12 -0.0178 0.041 -0.434 0.664 -0.099 0.063

x13 -0.0007 0.001 -0.500 0.617 -0.003 0.002

x14 15.7405 6.604 2.384 0.017 2.768 28.713

x15 -0.9302 2.099 -0.443 0.658 -5.054 3.194

x16 -3.9508 1.282 -3.081 0.002 -6.470 -1.432

x17 10.7992 5.448 1.982 0.048 0.097 21.501

x18 0.9476 2.536 0.374 0.709 -4.034 5.929

x19 4.5374 9.738 0.466 0.641 -14.592 23.667

x20 0.2092 0.057 3.686 0.000 0.098 0.321

x21 0.0072 0.007 1.043 0.297 -0.006 0.021

x22 -0.0039 0.006 -0.658 0.511 -0.015 0.008

x23 -0.0010 0.000 -3.319 0.001 -0.002 -0.000

x24 0.7796 1.418 0.550 0.583 -2.006 3.565

x25 0.4177 0.324 1.290 0.198 -0.219 1.054

x26 0.4943 0.262 1.890 0.059 -0.019 1.008

x27 0.3645 0.913 0.399 0.690 -1.429 2.158

x28 0.6646 0.356 1.868 0.062 -0.034 1.364

==============================================================================

Omnibus: 30.636 Durbin-Watson: 1.782

Prob(Omnibus): 0.000 Jarque-Bera (JB): 34.097

Skew: 0.582 Prob(JB): 3.94e-08

Kurtosis: 3.292 Cond. No. 1.33e+06

We noticed from previous table that the x5 or smoothness\_mean feature was the highest results, so we will delete it.

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.0866 0.421 -4.956 0.000 -2.914 -1.260

x1 -0.2301 0.170 -1.350 0.178 -0.565 0.105

x2 0.0047 0.008 0.594 0.553 -0.011 0.020

x3 0.0251 0.025 1.008 0.314 -0.024 0.074

x4 0.0003 0.001 0.665 0.506 -0.001 0.001

x5 -4.6761 1.257 -3.719 0.000 -7.146 -2.206

x6 1.2648 1.030 1.228 0.220 -0.758 3.288

x7 2.2655 1.874 1.209 0.227 -1.415 5.946

x8 5.4027 4.634 1.166 0.244 -3.701 14.506

x9 0.3751 0.305 1.230 0.219 -0.224 0.974

x10 -0.0075 0.036 -0.210 0.833 -0.078 0.063

x11 -0.0174 0.041 -0.426 0.670 -0.098 0.063

x12 -0.0007 0.001 -0.514 0.608 -0.003 0.002

x13 15.9455 6.355 2.509 0.012 3.463 28.428

x14 -0.9129 2.092 -0.436 0.663 -5.023 3.197

x15 -3.9555 1.281 -3.089 0.002 -6.471 -1.440

x16 10.7266 5.407 1.984 0.048 0.106 21.347

x17 0.9200 2.522 0.365 0.715 -4.034 5.874

x18 4.6272 9.698 0.477 0.633 -14.424 23.678

x19 0.2100 0.056 3.738 0.000 0.100 0.320

x20 0.0073 0.007 1.066 0.287 -0.006 0.021

x21 -0.0039 0.006 -0.670 0.503 -0.015 0.008

x22 -0.0011 0.000 -3.332 0.001 -0.002 -0.000

x23 0.6668 1.027 0.649 0.516 -1.350 2.684

x24 0.4216 0.322 1.310 0.191 -0.210 1.054

x25 0.4922 0.261 1.888 0.059 -0.020 1.004

x26 0.3898 0.886 0.440 0.660 -1.350 2.130

x27 0.6664 0.355 1.876 0.061 -0.031 1.364

==============================================================================

Omnibus: 30.552 Durbin-Watson: 1.783

Prob(Omnibus): 0.000 Jarque-Bera (JB): 33.989

Skew: 0.581 Prob(JB): 4.16e-08

Kurtosis: 3.291 Cond. No. 1.33e+06

Previous results showed that we must ignore the x10 0r the texture\_se in column 12 in the original x

So we must delete it and re- compute the p value.

==============================================================================

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.0865 0.421 -4.960 0.000 -2.913 -1.260

x1 -0.2342 0.169 -1.383 0.167 -0.567 0.098

x2 0.0055 0.007 0.801 0.424 -0.008 0.019

x3 0.0252 0.025 1.017 0.310 -0.024 0.074

x4 0.0004 0.001 0.710 0.478 -0.001 0.001

x5 -4.6740 1.256 -3.721 0.000 -7.142 -2.206

x6 1.2583 1.029 1.223 0.222 -0.762 3.279

x7 2.2264 1.863 1.195 0.233 -1.433 5.886

x8 5.3487 4.623 1.157 0.248 -3.733 14.430

x9 0.3653 0.301 1.213 0.226 -0.226 0.957

x10 -0.0181 0.041 -0.445 0.657 -0.098 0.062

x11 -0.0006 0.001 -0.478 0.633 -0.003 0.002

x12 15.7968 6.310 2.504 0.013 3.403 28.191

x13 -0.9020 2.090 -0.432 0.666 -5.007 3.203

x14 -3.9443 1.278 -3.086 0.002 -6.455 -1.433

x15 10.6641 5.394 1.977 0.049 0.069 21.259

x16 0.7442 2.378 0.313 0.754 -3.927 5.415

x17 4.5825 9.687 0.473 0.636 -14.447 23.612

x18 0.2123 0.055 3.854 0.000 0.104 0.321

x19 0.0063 0.005 1.264 0.207 -0.003 0.016

x20 -0.0039 0.006 -0.661 0.509 -0.015 0.008

x21 -0.0011 0.000 -3.461 0.001 -0.002 -0.000

x22 0.6903 1.020 0.677 0.499 -1.313 2.693

x23 0.4209 0.321 1.309 0.191 -0.211 1.052

x24 0.4922 0.260 1.890 0.059 -0.019 1.004

x25 0.4100 0.880 0.466 0.641 -1.318 2.138

x26 0.6864 0.342 2.008 0.045 0.015 1.358

==============================================================================

Omnibus: 30.471 Durbin-Watson: 1.784

Prob(Omnibus): 0.000 Jarque-Bera (JB): 33.886

Skew: 0.579 Prob(JB): 4.38e-08

Kurtosis: 3.294 Cond. No. 1.33e+06

From previous results we noticed that x16 or column 19 or symmetry\_se feature had the highest p-value, so we must delete it.

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.0896 0.420 -4.973 0.000 -2.915 -1.264

x1 -0.2320 0.169 -1.373 0.170 -0.564 0.100

x2 0.0056 0.007 0.820 0.413 -0.008 0.019

x3 0.0254 0.025 1.024 0.306 -0.023 0.074

x4 0.0003 0.001 0.685 0.494 -0.001 0.001

x5 -4.6645 1.255 -3.717 0.000 -7.129 -2.200

x6 1.2619 1.028 1.228 0.220 -0.757 3.281

x7 2.2374 1.861 1.202 0.230 -1.418 5.893

x8 5.3857 4.618 1.166 0.244 -3.685 14.457

x9 0.3720 0.300 1.240 0.216 -0.217 0.961

x10 -0.0160 0.040 -0.398 0.691 -0.095 0.063

x11 -0.0007 0.001 -0.555 0.579 -0.003 0.002

x12 16.3066 6.091 2.677 0.008 4.342 28.271

x13 -0.7895 2.057 -0.384 0.701 -4.830 3.251

x14 -3.9550 1.277 -3.098 0.002 -6.463 -1.447

x15 10.6186 5.387 1.971 0.049 0.036 21.201

x16 4.4205 9.665 0.457 0.648 -14.566 23.407

x17 0.2105 0.055 3.846 0.000 0.103 0.318

x18 0.0062 0.005 1.252 0.211 -0.004 0.016

x19 -0.0041 0.006 -0.704 0.482 -0.015 0.007

x20 -0.0010 0.000 -3.458 0.001 -0.002 -0.000

x21 0.6269 0.999 0.628 0.530 -1.335 2.588

x22 0.4052 0.317 1.277 0.202 -0.218 1.028

x23 0.4934 0.260 1.897 0.058 -0.018 1.004

x24 0.3962 0.878 0.451 0.652 -1.329 2.121

x25 0.7667 0.226 3.395 0.001 0.323 1.210

==============================================================================

Omnibus: 29.865 Durbin-Watson: 1.781

Prob(Omnibus): 0.000 Jarque-Bera (JB): 33.118

Skew: 0.575 Prob(JB): 6.43e-08

Kurtosis: 3.277 Cond. No. 1.33e+06

We conclude that the highest p-value was the x10 or col 13 or area\_se feature, so we will delete it.

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.0782 0.419 -4.961 0.000 -2.901 -1.255

x1 -0.2182 0.165 -1.320 0.187 -0.543 0.106

x2 0.0058 0.007 0.848 0.397 -0.008 0.019

x3 0.0240 0.025 0.978 0.328 -0.024 0.072

x4 0.0003 0.000 0.611 0.542 -0.001 0.001

x5 -4.6120 1.247 -3.699 0.000 -7.061 -2.163

x6 1.2968 1.023 1.267 0.206 -0.713 3.307

x7 2.2555 1.859 1.213 0.226 -1.396 5.907

x8 4.9306 4.471 1.103 0.271 -3.851 13.713

x9 0.2861 0.208 1.373 0.170 -0.123 0.695

x10 -0.0009 0.001 -0.686 0.493 -0.003 0.002

x11 16.4369 6.077 2.705 0.007 4.499 28.375

x12 -0.8938 2.039 -0.438 0.661 -4.898 3.111

x13 -3.8970 1.267 -3.075 0.002 -6.387 -1.407

x14 9.9641 5.127 1.944 0.052 -0.106 20.034

x15 5.3329 9.383 0.568 0.570 -13.098 23.764

x16 0.2158 0.053 4.068 0.000 0.112 0.320

x17 0.0060 0.005 1.219 0.223 -0.004 0.016

x18 -0.0055 0.005 -1.225 0.221 -0.014 0.003

x19 -0.0010 0.000 -3.483 0.001 -0.002 -0.000

x20 0.6768 0.990 0.684 0.494 -1.268 2.621

x21 0.4150 0.316 1.313 0.190 -0.206 1.036

x22 0.4856 0.259 1.873 0.062 -0.024 0.995

x23 0.4322 0.873 0.495 0.621 -1.282 2.146

x24 0.7673 0.226 3.400 0.001 0.324 1.211

==============================================================================

Omnibus: 30.032 Durbin-Watson: 1.781

Prob(Omnibus): 0.000 Jarque-Bera (JB): 33.330

Skew: 0.576 Prob(JB): 5.79e-08

Kurtosis: 3.281 Cond. No. 1.29e+06

The results were that x12 or col 17 or feature of concave points\_se

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.2217 0.411 -5.401 0.000 -3.030 -1.414

x1 -0.3226 0.163 -1.979 0.048 -0.643 -0.002

x2 0.0070 0.007 1.016 0.310 -0.006 0.020

x3 0.0437 0.024 1.827 0.068 -0.003 0.091

x4 8.08e-05 0.000 0.163 0.871 -0.001 0.001

x5 -5.2595 1.228 -4.281 0.000 -7.673 -2.846

x6 -0.3790 0.881 -0.430 0.667 -2.110 1.352

x7 4.6460 1.715 2.709 0.007 1.278 8.014

x8 6.9539 4.361 1.594 0.111 -1.613 15.521

x9 0.3705 0.209 1.776 0.076 -0.039 0.780

x10 -0.0012 0.001 -0.934 0.351 -0.004 0.001

x11 16.9426 5.804 2.919 0.004 5.542 28.343

x12 -0.5180 4.076 -0.127 0.899 -8.524 7.488

x13 -7.2495 7.810 -0.928 0.354 -22.592 8.093

x14 0.2072 0.053 3.877 0.000 0.102 0.312

x15 0.0055 0.005 1.115 0.265 -0.004 0.015

x16 -0.0061 0.005 -1.336 0.182 -0.015 0.003

x17 -0.0009 0.000 -3.205 0.001 -0.002 -0.000

x18 0.9003 0.953 0.945 0.345 -0.972 2.772

x19 0.5485 0.273 2.010 0.045 0.012 1.085

x20 0.2569 0.251 1.024 0.306 -0.236 0.750

x21 1.0007 0.854 1.172 0.242 -0.676 2.677

x22 0.6876 0.226 3.037 0.003 0.243 1.132

==============================================================================

Omnibus: 27.773 Durbin-Watson: 1.787

Prob(Omnibus): 0.000 Jarque-Bera (JB): 30.631

Skew: 0.563 Prob(JB): 2.23e-07

Kurtosis: 3.155 Cond. No. 1.10e+06

The highest p-value from x12 or col 18 or symmetry\_se feature.

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.2263 0.409 -5.437 0.000 -3.031 -1.422

x1 -0.3273 0.158 -2.066 0.039 -0.639 -0.016

x2 0.0069 0.007 1.013 0.311 -0.007 0.020

x3 0.0441 0.024 1.858 0.064 -0.003 0.091

x4 9.794e-05 0.000 0.206 0.837 -0.001 0.001

x5 -5.2917 1.201 -4.406 0.000 -7.651 -2.933

x6 -0.3909 0.876 -0.446 0.655 -2.111 1.329

x7 4.6962 1.667 2.817 0.005 1.421 7.971

x8 7.0348 4.311 1.632 0.103 -1.433 15.502

x9 0.3610 0.195 1.855 0.064 -0.021 0.743

x10 -0.0011 0.001 -0.931 0.352 -0.004 0.001

x11 16.7323 5.558 3.011 0.003 5.815 27.650

x12 -7.5982 7.306 -1.040 0.299 -21.950 6.753

x13 0.2095 0.050 4.180 0.000 0.111 0.308

x14 0.0056 0.005 1.122 0.262 -0.004 0.015

x15 -0.0061 0.005 -1.356 0.176 -0.015 0.003

x16 -0.0010 0.000 -3.374 0.001 -0.002 -0.000

x17 0.9291 0.925 1.004 0.316 -0.888 2.746

x18 0.5581 0.262 2.128 0.034 0.043 1.073

x19 0.2572 0.251 1.026 0.305 -0.235 0.750

x20 0.9322 0.661 1.409 0.159 -0.367 2.231

x21 0.6902 0.225 3.064 0.002 0.248 1.133

==============================================================================

Omnibus: 27.760 Durbin-Watson: 1.788

Prob(Omnibus): 0.000 Jarque-Bera (JB): 30.614

Skew: 0.563 Prob(JB): 2.25e-07

Kurtosis: 3.155 Cond. No. 1.05e+06

The highest p\_value from x 4 or col 4 or area\_mean feature

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.2708 0.347 -6.540 0.000 -2.953 -1.589

x1 -0.3205 0.155 -2.070 0.039 -0.625 -0.016

x2 0.0069 0.007 1.009 0.314 -0.007 0.020

x3 0.0448 0.023 1.907 0.057 -0.001 0.091

x4 -5.3756 1.129 -4.763 0.000 -7.592 -3.159

x5 -0.3736 0.871 -0.429 0.668 -2.084 1.337

x6 4.6967 1.666 2.820 0.005 1.425 7.969

x7 7.2513 4.177 1.736 0.083 -0.953 15.455

x8 0.3708 0.189 1.966 0.050 0.000 0.741

x9 -0.0012 0.001 -0.950 0.342 -0.004 0.001

x10 16.7515 5.552 3.017 0.003 5.845 27.658

x11 -7.8106 7.226 -1.081 0.280 -22.005 6.384

x12 0.2034 0.040 5.024 0.000 0.124 0.283

x13 0.0056 0.005 1.132 0.258 -0.004 0.015

x14 -0.0062 0.005 -1.373 0.170 -0.015 0.003

x15 -0.0009 0.000 -5.045 0.000 -0.001 -0.001

x16 0.9502 0.918 1.035 0.301 -0.854 2.754

x17 0.5690 0.257 2.218 0.027 0.065 1.073

x18 0.2537 0.250 1.015 0.310 -0.237 0.745

x19 0.9323 0.661 1.411 0.159 -0.366 2.230

x20 0.6956 0.224 3.112 0.002 0.256 1.135

==============================================================================

Omnibus: 27.501 Durbin-Watson: 1.787

Prob(Omnibus): 0.000 Jarque-Bera (JB): 30.313

Skew: 0.561 Prob(JB): 2.62e-07

Kurtosis: 3.144 Cond. No. 8.51e+05

Highest value from x5 or col7 or smoothness\_mean feature

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.2857 0.345 -6.621 0.000 -2.964 -1.608

x1 -0.3093 0.152 -2.028 0.043 -0.609 -0.010

x2 0.0070 0.007 1.028 0.304 -0.006 0.020

x3 0.0430 0.023 1.862 0.063 -0.002 0.088

x4 -5.4546 1.113 -4.903 0.000 -7.640 -3.269

x5 4.2482 1.296 3.278 0.001 1.703 6.794

x6 7.3221 4.170 1.756 0.080 -0.869 15.514

x7 0.3718 0.188 1.974 0.049 0.002 0.742

x8 -0.0012 0.001 -0.991 0.322 -0.004 0.001

x9 16.4676 5.509 2.989 0.003 5.647 27.288

x10 -8.5871 6.991 -1.228 0.220 -22.319 5.145

x11 0.2040 0.040 5.044 0.000 0.125 0.283

x12 0.0055 0.005 1.107 0.269 -0.004 0.015

x13 -0.0061 0.005 -1.360 0.174 -0.015 0.003

x14 -0.0009 0.000 -5.070 0.000 -0.001 -0.001

x15 1.0172 0.904 1.125 0.261 -0.759 2.794

x16 0.6145 0.233 2.633 0.009 0.156 1.073

x17 0.1630 0.133 1.224 0.221 -0.099 0.425

x18 1.0137 0.633 1.602 0.110 -0.229 2.256

x19 0.6889 0.223 3.091 0.002 0.251 1.127

==============================================================================

Omnibus: 27.873 Durbin-Watson: 1.785

Prob(Omnibus): 0.000 Jarque-Bera (JB): 30.747

Skew: 0.564 Prob(JB): 2.11e-07

Kurtosis: 3.159 Cond. No. 8.38e+05

===============================================================x8 or col 14 was the highest one area\_se feature

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.3454 0.340 -6.900 0.000 -3.013 -1.678

x1 -0.2905 0.151 -1.920 0.055 -0.588 0.007

x2 0.0077 0.007 1.138 0.256 -0.006 0.021

x3 0.0384 0.023 1.698 0.090 -0.006 0.083

x4 -5.3232 1.105 -4.819 0.000 -7.493 -3.153

x5 4.5693 1.255 3.642 0.000 2.105 7.034

x6 7.4978 4.166 1.800 0.072 -0.686 15.682

x7 0.1996 0.073 2.749 0.006 0.057 0.342

x8 17.6337 5.381 3.277 0.001 7.063 28.204

x9 -7.7555 6.940 -1.117 0.264 -21.388 5.877

x10 0.2214 0.036 6.074 0.000 0.150 0.293

x11 0.0051 0.005 1.034 0.302 -0.005 0.015

x12 -0.0057 0.004 -1.267 0.206 -0.014 0.003

x13 -0.0010 0.000 -7.865 0.000 -0.001 -0.001

x14 0.8300 0.884 0.939 0.348 -0.907 2.567

x15 0.5810 0.231 2.516 0.012 0.127 1.035

x16 0.1649 0.133 1.239 0.216 -0.097 0.426

x17 0.9323 0.627 1.486 0.138 -0.300 2.164

x18 0.6922 0.223 3.107 0.002 0.255 1.130

==============================================================================

Omnibus: 28.925 Durbin-Watson: 1.778

Prob(Omnibus): 0.000 Jarque-Bera (JB): 31.985

Skew: 0.571 Prob(JB): 1.13e-07

Kurtosis: 3.213 Cond. No. 8.37e+05

==============================================================================

X14 or col 25 was the highest p- value

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.3480 0.340 -6.908 0.000 -3.016 -1.680

x1 -0.2990 0.151 -1.980 0.048 -0.596 -0.002

x2 0.0060 0.007 0.919 0.358 -0.007 0.019

x3 0.0392 0.023 1.734 0.083 -0.005 0.084

x4 -5.3739 1.103 -4.871 0.000 -7.541 -3.207

x5 4.6034 1.254 3.671 0.000 2.140 7.067

x6 8.8274 3.918 2.253 0.025 1.132 16.523

x7 0.1888 0.072 2.634 0.009 0.048 0.330

x8 20.5169 4.418 4.644 0.000 11.839 29.195

x9 -10.6956 6.192 -1.727 0.085 -22.859 1.468

x10 0.2297 0.035 6.497 0.000 0.160 0.299

x11 0.0063 0.005 1.317 0.188 -0.003 0.016

x12 -0.0064 0.004 -1.434 0.152 -0.015 0.002

x13 -0.0010 0.000 -7.966 0.000 -0.001 -0.001

x14 0.5977 0.230 2.596 0.010 0.145 1.050

x15 0.1655 0.133 1.243 0.214 -0.096 0.427

x16 1.0747 0.609 1.766 0.078 -0.121 2.270

x17 0.7020 0.223 3.155 0.002 0.265 1.139

==============================================================================

Omnibus: 29.212 Durbin-Watson: 1.778

Prob(Omnibus): 0.000 Jarque-Bera (JB): 32.325

Skew: 0.572 Prob(JB): 9.57e-08

Kurtosis: 3.231 Cond. No. 7.03e+05

X2 or col2 was the highest

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.3256 0.339 -6.861 0.000 -2.991 -1.660

x1 -0.2968 0.151 -1.966 0.050 -0.593 -0.000

x2 0.0393 0.023 1.738 0.083 -0.005 0.084

x3 -5.2793 1.098 -4.807 0.000 -7.437 -3.122

x4 4.5589 1.253 3.639 0.000 2.098 7.020

x5 8.5364 3.904 2.186 0.029 0.867 16.206

x6 0.1994 0.071 2.818 0.005 0.060 0.338

x7 20.7224 4.411 4.697 0.000 12.057 29.388

x8 -10.1999 6.168 -1.654 0.099 -22.315 1.915

x9 0.2279 0.035 6.457 0.000 0.159 0.297

x10 0.0103 0.002 5.442 0.000 0.007 0.014

x11 -0.0064 0.004 -1.435 0.152 -0.015 0.002

x12 -0.0010 0.000 -7.974 0.000 -0.001 -0.001

x13 0.5876 0.230 2.556 0.011 0.136 1.039

x14 0.1667 0.133 1.253 0.211 -0.095 0.428

x15 1.0224 0.606 1.688 0.092 -0.168 2.212

x16 0.6785 0.221 3.070 0.002 0.244 1.113

==============================================================================

Omnibus: 29.105 Durbin-Watson: 1.778

Prob(Omnibus): 0.000 Jarque-Bera (JB): 32.192

Skew: 0.571 Prob(JB): 1.02e-07

Kurtosis: 3.229 Cond. No. 7.00e+05

X14 or col 27 was the highest

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.2933 0.338 -6.782 0.000 -2.958 -1.629

x1 -0.3303 0.149 -2.222 0.027 -0.622 -0.038

x2 0.0444 0.022 1.998 0.046 0.001 0.088

x3 -5.5765 1.073 -5.198 0.000 -7.684 -3.469

x4 4.7060 1.248 3.771 0.000 2.254 7.158

x5 8.4194 3.905 2.156 0.032 0.749 16.090

x6 0.1962 0.071 2.773 0.006 0.057 0.335

x7 20.3795 4.405 4.626 0.000 11.727 29.033

x8 -8.4042 6.002 -1.400 0.162 -20.194 3.385

x9 0.2285 0.035 6.473 0.000 0.159 0.298

x10 0.0104 0.002 5.523 0.000 0.007 0.014

x11 -0.0068 0.004 -1.532 0.126 -0.015 0.002

x12 -0.0010 0.000 -7.910 0.000 -0.001 -0.001

x13 0.7341 0.198 3.706 0.000 0.345 1.123

x14 1.2379 0.581 2.130 0.034 0.096 2.379

x15 0.6811 0.221 3.080 0.002 0.247 1.116

==============================================================================

Omnibus: 29.970 Durbin-Watson: 1.792

Prob(Omnibus): 0.000 Jarque-Bera (JB): 33.266

Skew: 0.579 Prob(JB): 5.98e-08

Kurtosis: 3.253 Cond. No. 6.84e+05

X8 or col20 was the highest one

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.1382 0.320 -6.687 0.000 -2.766 -1.510

x1 -0.3348 0.149 -2.251 0.025 -0.627 -0.043

x2 0.0443 0.022 1.988 0.047 0.001 0.088

x3 -5.5166 1.073 -5.142 0.000 -7.624 -3.409

x4 4.8774 1.243 3.924 0.000 2.436 7.319

x5 5.9959 3.504 1.711 0.088 -0.886 12.878

x6 0.1724 0.069 2.508 0.012 0.037 0.307

x7 19.2963 4.341 4.446 0.000 10.770 27.822

x8 0.2291 0.035 6.483 0.000 0.160 0.298

x9 0.0104 0.002 5.498 0.000 0.007 0.014

x10 -0.0065 0.004 -1.474 0.141 -0.015 0.002

x11 -0.0010 0.000 -7.843 0.000 -0.001 -0.001

x12 0.6888 0.196 3.522 0.000 0.305 1.073

x13 1.2188 0.581 2.096 0.037 0.077 2.361

x14 0.7433 0.217 3.428 0.001 0.317 1.169

==============================================================================

Omnibus: 31.150 Durbin-Watson: 1.797

Prob(Omnibus): 0.000 Jarque-Bera (JB): 34.754

Skew: 0.588 Prob(JB): 2.84e-08

Kurtosis: 3.289 Cond. No. 4.64e+05

X10 or col 23

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.2286 0.314 -7.094 0.000 -2.846 -1.612

x1 -0.2147 0.125 -1.723 0.085 -0.459 0.030

x2 0.0257 0.018 1.399 0.163 -0.010 0.062

x3 -5.2470 1.058 -4.958 0.000 -7.326 -3.168

x4 5.1714 1.228 4.210 0.000 2.759 7.584

x5 6.5116 3.490 1.866 0.063 -0.344 13.367

x6 0.1675 0.069 2.437 0.015 0.032 0.302

x7 19.5185 4.342 4.495 0.000 10.989 28.048

x8 0.1894 0.023 8.272 0.000 0.144 0.234

x9 0.0105 0.002 5.565 0.000 0.007 0.014

x10 -0.0010 0.000 -7.918 0.000 -0.001 -0.001

x11 0.6386 0.193 3.312 0.001 0.260 1.017

x12 1.1115 0.578 1.925 0.055 -0.023 2.246

x13 0.7485 0.217 3.449 0.001 0.322 1.175

==============================================================================

Omnibus: 29.269 Durbin-Watson: 1.789

Prob(Omnibus): 0.000 Jarque-Bera (JB): 32.365

Skew: 0.567 Prob(JB): 9.38e-08

Kurtosis: 3.279 Cond. No. 4.61e+05

X2 0r col3 was the highest

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -2.2117 0.314 -7.039 0.000 -2.829 -1.595

x1 -0.0419 0.016 -2.583 0.010 -0.074 -0.010

x2 -4.3182 0.825 -5.236 0.000 -5.938 -2.698

x3 5.4834 1.209 4.536 0.000 3.109 7.858

x4 5.4727 3.413 1.604 0.109 -1.231 12.176

x5 0.1774 0.068 2.593 0.010 0.043 0.312

x6 18.6536 4.302 4.336 0.000 10.204 27.104

x7 0.1783 0.021 8.295 0.000 0.136 0.220

x8 0.0108 0.002 5.747 0.000 0.007 0.014

x9 -0.0010 0.000 -7.867 0.000 -0.001 -0.001

x10 0.6393 0.193 3.313 0.001 0.260 1.018

x11 1.0434 0.576 1.812 0.071 -0.088 2.175

x12 0.7293 0.217 3.365 0.001 0.304 1.155

==============================================================================

Omnibus: 27.603 Durbin-Watson: 1.774

Prob(Omnibus): 0.000 Jarque-Bera (JB): 30.285

Skew: 0.551 Prob(JB): 2.65e-07

Kurtosis: 3.254 Cond. No. 4.57e+05

X4 or col 10

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -1.7761 0.158 -11.235 0.000 -2.087 -1.466

x1 -0.0517 0.015 -3.427 0.001 -0.081 -0.022

x2 -3.5638 0.678 -5.254 0.000 -4.896 -2.231

x3 5.5634 1.210 4.599 0.000 3.187 7.939

x4 0.1737 0.068 2.537 0.011 0.039 0.308

x5 19.3491 4.286 4.514 0.000 10.930 27.768

x6 0.1743 0.021 8.153 0.000 0.132 0.216

x7 0.0105 0.002 5.615 0.000 0.007 0.014

x8 -0.0009 0.000 -7.691 0.000 -0.001 -0.001

x9 0.6100 0.192 3.171 0.002 0.232 0.988

x10 1.0291 0.577 1.785 0.075 -0.104 2.162

x11 0.7418 0.217 3.420 0.001 0.316 1.168

==============================================================================

Omnibus: 26.548 Durbin-Watson: 1.787

Prob(Omnibus): 0.000 Jarque-Bera (JB): 28.978

Skew: 0.538 Prob(JB): 5.10e-07

Kurtosis: 3.254 Cond. No. 4.51e+05

X10 was the highest

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -1.7868 0.158 -11.289 0.000 -2.098 -1.476

x1 -0.0595 0.014 -4.109 0.000 -0.088 -0.031

x2 -3.8461 0.661 -5.819 0.000 -5.144 -2.548

x3 6.9660 0.921 7.562 0.000 5.157 8.775

x4 0.1465 0.067 2.190 0.029 0.015 0.278

x5 19.2426 4.294 4.481 0.000 10.808 27.677

x6 0.1873 0.020 9.302 0.000 0.148 0.227

x7 0.0103 0.002 5.520 0.000 0.007 0.014

x8 -0.0010 0.000 -8.290 0.000 -0.001 -0.001

x9 0.7773 0.168 4.619 0.000 0.447 1.108

x10 0.7561 0.217 3.481 0.001 0.329 1.183

==============================================================================

Omnibus: 29.452 Durbin-Watson: 1.803

Prob(Omnibus): 0.000 Jarque-Bera (JB): 32.637

Skew: 0.560 Prob(JB): 8.19e-08

Kurtosis: 3.349 Cond. No. 4.51e+05

X4 is the highest

coef std err t P>|t| [0.025 0.975]

------------------------------------------------------------------------------

const -1.7606 0.158 -11.117 0.000 -2.072 -1.450

x1 -0.0638 0.014 -4.432 0.000 -0.092 -0.035

x2 -3.4989 0.644 -5.434 0.000 -4.764 -2.234

x3 7.0506 0.924 7.634 0.000 5.237 8.865

x4 21.7445 4.153 5.235 0.000 13.586 29.903

x5 0.1884 0.020 9.326 0.000 0.149 0.228

x6 0.0102 0.002 5.447 0.000 0.007 0.014

x7 -0.0009 0.000 -7.969 0.000 -0.001 -0.001

x8 0.6685 0.161 4.143 0.000 0.352 0.985

x9 0.7267 0.218 3.341 0.001 0.299 1.154

==============================================================================

Omnibus: 34.039 Durbin-Watson: 1.784

Prob(Omnibus): 0.000 Jarque-Bera (JB): 39.069

Skew: 0.587 Prob(JB): 3.28e-09

Kurtosis: 3.517 Cond. No. 4.35e+05

Now, all x’s less than 0.025 , so we re-split data and then evaluate it.

The evaluation results were:

r2 socre is: 0.768052920531675

mean\_sqrd\_error is: 0.056202012407337305

root\_mean\_squared error of is: 0.23706963619860158.

Table 1 Comparisons between the Evaluation results ( SLR,MLR)

|  |  |  |
| --- | --- | --- |
| **Metrics** | **With SLR** | **With MLR** |
| **r2 socre** | 0.7234520971661356 | 0.768052920531675 |
| **mean\_sqrd\_error** | 0.06545512493109215 | 0.056202012407337305 |
| **root\_mean\_squared error** | 0.2558419921183623 | 0.23706963619860158. |

## **Conclusion:**

We concluded that the most feature that effected on the prediction of the breast masses were 10 features from 30 features were the:

* radius\_mean
* texture\_mean
* concavity\_mean
* symmetry\_mean
* concavity\_se
* perimeter\_worst
* area\_worst
* concavity\_worst
* symmetry\_worst
* fractal\_dimension\_worst.