Session 12 – Non- Linear Models

Assignment - 1

1. Use the given link Data Set.

Answer the below questions:

1. Perform ANOVA test on the discriminant analysis scores of nuclear localization signals of both nuclear and non-nuclear proteins by class variables (Target).
2. names(yeast) <- c("seq","mcg", "gvh", "alm", "mit", "erl", "pox", "vac", "nuc", "class")
3. > head(yeast)
4. # A tibble: 6 x 10
5. seq mcg gvh alm mit erl pox vac nuc class
6. *<chr>* *<dbl>* *<dbl>* *<dbl>* *<dbl>* *<dbl>* *<dbl>* *<dbl>* *<dbl>* *<chr>*
7. 1 ADT1\_YEAST 0.580 0.61 0.47 0.13 0.5 0 0.48 0.22 MIT
8. 2 ADT2\_YEAST 0.43 0.67 0.48 0.27 0.5 0 0.53 0.22 MIT
9. 3 ADT3\_YEAST 0.64 0.62 0.49 0.15 0.5 0 0.53 0.22 MIT
10. 4 AAR2\_YEAST 0.580 0.44 0.570 0.13 0.5 0 0.54 0.22 NUC
11. 5 AATM\_YEAST 0.42 0.44 0.48 0.54 0.5 0 0.48 0.22 MIT
12. 6 AATC\_YEAST 0.51 0.4 0.56 0.17 0.5 0.5 0.49 0.22 CYT
13. > dim(yeast)
14. [1] 1484 10
15. > sapply(yeast, function(x) sum(is.na(yeast)))
16. seq mcg gvh alm mit erl pox vac nuc class
17. 0 0 0 0 0 0 0 0 0 0

HO : scores across the class variables is same

x <- aov(nuc~class, data = yeast)

summary(x)

# since p value is less than 0.05, scores across the class variables is

# not same at 95% confidence level

b. Which class is significantly different from others?

TukeyHSD(x)

Tukey multiple comparisons of means

95% family-wise confidence level

Fit: aov(formula = nuc ~ class, data = yeast)

$`class`

diff lwr upr p adj

ERL-CYT -0.0112354212 -0.15411077 0.131639929 0.9999999

EXC-CYT -0.0306639926 -0.08636974 0.025041758 0.7698696

ME1-CYT 0.0109463970 -0.03918352 0.061076319 0.9995610

ME2-CYT -0.0117452251 -0.05862827 0.035137825 0.9986751

ME3-CYT 0.0127645788 -0.01617637 0.041705527 0.9281817

MIT-CYT -0.0162518146 -0.04139001 0.008886385 0.5645771

NUC-CYT 0.0744429005 0.05314812 0.095737681 0.0000000

POX-CYT -0.0232354212 -0.09580890 0.049338059 0.9914395

VAC-CYT -0.0065687545 -0.06643503 0.053297518 0.9999988

EXC-ERL -0.0194285714 -0.17135063 0.132493492 0.9999953

ME1-ERL 0.0221818182 -0.12778548 0.172149120 0.9999834

ME2-ERL -0.0005098039 -0.14942321 0.148403604 1.0000000

ME3-ERL 0.0240000000 -0.12027322 0.168273217 0.9999547

MIT-ERL -0.0050163934 -0.14857513 0.138542346 1.0000000

NUC-ERL 0.0856783217 -0.05725750 0.228614148 0.6696330

POX-ERL -0.0120000000 -0.17088390 0.146883897 1.0000000

VAC-ERL 0.0046666667 -0.14882976 0.158163094 1.0000000

ME1-EXC 0.0416103896 -0.03036153 0.113582306 0.7147855

ME2-EXC 0.0189187675 -0.05083056 0.088668090 0.9975341

ME3-EXC 0.0434285714 -0.01577047 0.102627610 0.3735707

MIT-EXC 0.0144121780 -0.04302367 0.071848027 0.9986581

NUC-EXC 0.1051068931 0.04924622 0.160967571 0.0000001

POX-EXC 0.0074285714 -0.08164363 0.096500777 0.9999999

VAC-EXC 0.0240952381 -0.05496751 0.103157987 0.9940385

ME2-ME1 -0.0226916221 -0.08807393 0.042690687 0.9847913

ME3-ME1 0.0018181818 -0.05216704 0.055803406 1.0000000

MIT-ME1 -0.0271982116 -0.07924391 0.024847491 0.8199565

NUC-ME1 0.0634965035 0.01319448 0.113798528 0.0026743

POX-ME1 -0.0341818182 -0.11987742 0.051513783 0.9613426

VAC-ME1 -0.0175151515 -0.09275342 0.057723114 0.9992623

ME3-ME2 0.0245098039 -0.02647466 0.075494270 0.8829855

MIT-ME2 -0.0045065895 -0.05343272 0.044419543 0.9999997

NUC-ME2 0.0861881256 0.03912110 0.133255152 0.0000004

POX-ME2 -0.0114901961 -0.09532782 0.072347429 0.9999914

VAC-ME2 0.0051764706 -0.06793856 0.078291502 1.0000000

MIT-ME3 -0.0290163934 -0.06116176 0.003128972 0.1179577

NUC-ME3 0.0616783217 0.03244028 0.090916364 0.0000000

POX-ME3 -0.0360000000 -0.11128815 0.039288150 0.8863429

VAC-ME3 -0.0193333333 -0.08246309 0.043796422 0.9938197

NUC-MIT 0.0906947151 0.06521504 0.116174388 0.0000000

POX-MIT -0.0069836066 -0.08089339 0.066926177 0.9999997

VAC-MIT 0.0096830601 -0.05179634 0.071162460 0.9999716

POX-NUC -0.0976783217 -0.17037079 -0.024985855 0.0009144

VAC-NUC -0.0810116550 -0.14102211 -0.021001196 0.0008397

VAC-POX 0.0166666667 -0.07506499 0.108398327 0.9999044

# ERL is significantly different from others