# Predicting readmission probability for diabetes in patients

STAT 471/571/701, Fall 2017

*Akhil Ganti*

## Executive Summary

As diabetes is a chronic medical condition that affects a large number of Americans every year, improper management can lead to patients being repeatedly readmitted to hospitals. These readmissions constitute an especially grievous problem that is indicative of the healthcare system not being able to provide adequate support to patients as well as large costs that are incurred as a result. It is highly important then to be able to identify factors that can aid in predicting which patients are more likely to be readmitted and act accordingly.

Sourced from the Center for Clinical and Translational Research at Virginia Commonwealth University, the data used covers diabetes patients across 130 hopsitals in the time from 1998 to 2008 and has over 100,000 unique hospital admissions and about 70,000 corresponding unique patients. I use this data to identify the factors predicting whether or not a given patient will be readmitted within 30 days as well as to create a classification rule for this prediction.

The priamry method used in my analysis focused on the selection of the most important features through regularized logistic regression by 10-fold cross validation. After creating and testing between several models, the final one chosen (constructed instead through elastic net regularization) consisted of the following variables: number of emergency visits by the patient in the year prior to the current encounter, number of inpatient visits by the patient in the year prior to the current encounter, total number of diagnoses entered for the patient, results of the insulin test, whether any diabetes medication was prescribed, where the patient was discharged to after treatment, age, the ICD9 codes for the primary and tertiary diagnoses of the patient, and whether the patient was a repeat admit.

### Concerns & Limitations

The limitations of this analysis stem both from flaws in the dataset as well as in the methods used to produce the final model. As this data only covers 130 hospitals, it is possible that it does not represent the total population fully and accurately, so certain conclusions drawn may in fact be specific only to the sample used here. In addition, due to the inherent variety among these hospitals, there is the risk that the data itself was not collected uniformly across, which would also potentially skew the final results. A final concern occurs due to the time period in which the data was collected. Since the time span was across nine years, it's more likely than not that procedural changes occurred in hospitals that would have altered the collection of the data itself.

## Summary of Data

As mentioned above, the data has over 100,000 unique hospital admissions and about 70,000 unique patients, with approximately 50 features. However, for the purpose of this analysis, the provided cleaned dataset is used instead, which contains only 31 variables. While there are no missing values, in the race column, 2273 patients are marked as "?", and in the gender column, 3 are marked as "Unknown/Invalid". Given that these entries comprise only about 2% of the whole dataset, I decided to omit them altogether. In addition, I created an additional variable called "not_unique", which indicates whether a patient, based on his/her ID,

is a repeat admit. After transforming the response variable into an indicator variable, we can see that 11169 patients were readmitted in less than 30 days (11.22% of the observations), while the other 88323 were either not readmitted again or were readmitted after 30 days. Note that in creating the models, all variables except "encounter_id" and "patient_nbr" are used. See figure 1 in the appendix for a summary of the (cleaned) dataset.

### Exploratory Data Analysis

We will look at plots of some of the different features (both numeric and factor) against the readmitted indictator to get a better sense of the data as a whole. We first look at some of the demographic information, namely race and gender. See figures 2 and 3 in the appendix.

It seems that neither race nor gender has any distinguishing characteristics for readmitted patients. One variable that might be expected to be separating, however, is the number of distinct medications prescribed in the current encounter, since it's reasonable to think that more prescriptions implies worse health and consequently a higher chance of readmission. See figure 4 in the appendix.

Surprisingly, although the mean of the positive response is slightly higher, the distributions appear extremely close. We next look at the "not_unique" variable that was created, since it is reasonable to expect that patients who are repeat admits are more likely to be readmitted in less than 30 days. See figure 5 in the appendix.

As expected, there is a measurable difference in the response based solely on this variable. The next variable that we look at it is the total number of diagnoses entered for a patient. The line of reasoning here is that, with a higher number of diagnoses, there is likely to be more uncertaintiy surrounding the patient's condition, which might be a contributing factor to the patient's quick readmission. See figure 6 in the appendix.

The graph confirms the expectation with a higher mean and smaller distribution for the positive responses. The final variable we look at is age, since it is expected that older patients are more likely to have higher readmission probabilities due to overall health complications that are associated with aging. See figure 7 in the appendix.

Surprisingly, there does not seem to be a distinguishing difference in the response based solely on age.

## Analyses

### Factor Identification

As mentioned in the executive summary, LASSO logistic regression (through the glmnet package) by 10-fold cross-validation was used to select a subset of the most predictive variables. The default deviance type.measure parameter was used as well, while the other measurement possibilites will be used later in determining a final model. See figure 8 in the appendix for a plot of the binomial deviance against log(lambda) and figure 9 for the Anova table of the logistic model constructed from the chosen variables (with s = "lambda.1se").

We can see that all of the variables are significant at the 0.05 level except for time_in_hospital and number_diagnoses. While time_in_hospital will be removed to construct a new model, I chose to keep number_diagnoses based on the fact that it's still significant at the 0.1 level as well as because it seems to have some distinguishing power based on the plot in figure 6. The resulting model consists of the variables num_medications, number_emergency, number_inpatient, number_diagnoses, metformin, insulin, diabetesMed, disch_disp_modified, age_mod, diag1_mod, diag2_mod, diag3_mod, and not_unique. From the ANOVA table for this model (figure 10), we see that all variables are significant. A summary of the coefficients themselves can be seen in figure 11 with a corresponding ROC curve in figure 12.

Looking at some of the coefficients themselves, we see that the probability of readmission is positively correlated with the number of medications, number of emergency visits, number of inpatient visits, whether

the patient was prescribed diabetes medications, and number of diagnoses. All of these make sense since they are all either indicative of worse overall health (the first four) or uncertainty about the patient's condition (the fifth). Unsurprisngly, being a repeat admit also increases the chance of readmission within 30 days.

In addition, based on the age coefficients, there seems to be the overall trend that older people are more likely to be readmitted, though it's interesting to note that people between 60 and 79 years have a greater probability of readmission than people older than 80 years, though this paticular result could possibly be attributed to the large difference in sample size for each category.

Where a patient is discharged to after treatment also is positively correlated with readmission - those who aren't discharged to their homes have a higher chance of readmission than those who are, and this is easily explained since a patient not being discharged to his/her home implies a relatively more serious condition.

## Model Selection and Classification

As mentioned above, two additional models were created by running cv.glmnet using the "mce" and "auc" type.measure values instead of "deviance." While "lambda.1se" was used to extract coefficients from the AUC model, "lambda.min" had to be used for the MCE model since, interestingly, it produced only the intercept under "lambda.1se." Note that the AUC model produced the same variables both before and after removing insignificant variables as the full-LASSO deviance model (abbreviated as FLD onwards). The MCE against log(lambda) curve can be seen in figure 13 in the appendix and the AUC against log(lambda) can be seen in figure 14, with the respective ROC curves based on the resulting models in the subsequent two figures.

In addition, one more model was created using the deviance as type.measure but with elastic net regularization instead, setting alpha to 0.99. Similar to the FLD model, there were some insignificant variables (see figure 17) that were removed in constructing the model, the ANOVA of which is in figure 18. It should be noted here that, in contrast to the FLD model, the variables num_medications, metformin, and diag2_mod were zeroed out.

We now have four candidate models from which to select the final one: the FLD deviance model, the MCE model, the AUC model, and the elastic net deviance model. However, given that the AUC model is equivalent to the FLD model, we really only have three models to consider. Using the given ratio that it costs approximately twice as much to misclassify a readmission than to misclassify a non-readmission, the resulting Bayesian classificaiton rule is to predict as readmission if the logistic probability is greater than $\frac{1}{3}$. Calculating the MCEs for these models, we get:

FLD Model:

```
## [1] 0.2226611
```

MCE Model:

```
## [1] 0.2226812
```

Elastic Net Deviance Model:

```
## [1] 0.22245
```

Oddly, the MCE model has a higher MCE than the FLD model, though this is most likely because of my choice to remove the insignificant variables in the latter as well as due to incorporated unequal loss. In addition, given that these values are so close to each other, the difference is most likely not meaningful. However, it's important to note here that the elastic net model has the lowest MCE. We now look at the AUC values:

FLD Model:

```
## Area under the curve: 0.659
```

MCE Model:

```
## Area under the curve: 0.6466
```

Elastic Net Deviance Model:

```
## Area under the curve: 0.6571
```

As expected, the FLD model (which the AUC model is equivalent to) has the highest AUC value, with the MCE model having the lowest AUC. Even though the elastic net model is only slightly better than the FLD model in terms of MCE and worse by a greater amount in terms of AUC, I nonetheless choose the elastic net model to be the final model due to its comparable performance and greater parsimony.

# Conclusion

The final model, constructed through elastic net regularization and based on binomial deviance, consisted of the variables number_emergency, number_inpatient, number_diagnoses, insulin, diabetesMed, disch_disp_modified, age_mod, diag1_mod, diag3_mod, and not_unique. This model was able to achieve a Bayesian misclassification error of 22.245% and an AUC of 0.6571. While this is marginally worse than the other tested models, overall it is a good model due to being relatively more parsimonious. Based on this, my initial recommendation would be to utilize this model to predict whether a given new patient will be readmitted, and in the meantime collect more data that relates more directly to the occurrence of diabetes in patients to help improve the model's accuracy in future iterations. In addition, further research could be conducted to see if there are any significant interactions between variables or if any meaningful transformations could be performed to increase predictive power.

# Appendix

Note that because of the large amount of time taken to compute the different variables (e.g. fit1.cv, etc.), all code (which can be seen below after figure 18) was run in a separate file and the variables were saved in .rds files, which were then loaded above to speed up the knitting of this .rmd file.

### Figure 1 - Summary of Dataset

```
##   encounter_id        patient_nbr                     race
##  Min.   :    12522   Min.   :      135   AfricanAmerican:19210
##  1st Qu.: 84616156   1st Qu.: 23460039   Asian          :  641
##  Median :152592672   Median : 45982994   Caucasian      :76099
##  Mean   :165026773   Mean   : 54563668   Hispanic       : 2037
##  3rd Qu.:229541397   3rd Qu.: 87797016   Other          : 1505
##  Max.   :443867222   Max.   :189502619
##
##     gender      time_in_hospital num_lab_procedures num_procedures
##  Female:53575   Min.   : 1.000   Min.   :  1.00     Min.   :0.000
##  Male  :45917   1st Qu.: 2.000   1st Qu.: 31.00     1st Qu.:0.000
##                 Median : 4.000   Median : 44.00     Median :1.000
##                 Mean   : 4.398   Mean   : 43.07     Mean   :1.341
##                 3rd Qu.: 6.000   3rd Qu.: 57.00     3rd Qu.:2.000
##                 Max.   :14.000   Max.   :132.00     Max.   :6.000
##
##  num_medications number_outpatient number_emergency  number_inpatient
##  Min.   : 1.00   Min.   : 0.0000   Min.   : 0.0000   Min.   : 0.0000
##  1st Qu.:10.00   1st Qu.: 0.0000   1st Qu.: 0.0000   1st Qu.: 0.0000
##  Median :15.00   Median : 0.0000   Median : 0.0000   Median : 0.0000
##  Mean   :16.03   Mean   : 0.3734   Mean   : 0.2013   Mean   : 0.6431
```

```
##  3rd Qu.:20.00   3rd Qu.: 0.0000   3rd Qu.: 0.0000   3rd Qu.: 1.0000
##  Max.   :81.00   Max.   :42.0000   Max.   :76.0000   Max.   :21.0000
##
##  number_diagnoses max_glu_serum A1Cresult     metformin     glimepiride
##  Min.   : 1.000   >200: 1466    >7  : 3730   Down  :  562   Down  :  186
##  1st Qu.: 6.000   >300: 1253    >8  : 7961   No    :79971   No    :94458
##  Median : 8.000   None:94202    None:82896   Steady:17920   Steady: 4527
##  Mean   : 7.439   Norm: 2571    Norm: 4905   Up    : 1039   Up    :  321
##  3rd Qu.: 9.000
##  Max.   :16.000
##
##   glipizide      glyburide     pioglitazone   rosiglitazone
##  Down  :  544   Down  :  546   Down  :  117   Down  :   84
##  No    :87093   No    :89113   No    :92316   No    :93265
##  Steady:11101   Steady: 9039   Steady: 6829   Steady: 5968
##  Up    :  754   Up    :  794   Up    :  230   Up    :  175
##
##
##
##    insulin      change     diabetesMed
##  Down  :12039   Ch:45910   No :23001
##  No    :46495   No:53582   Yes:76491
##  Steady:29902
##  Up    :11056
##
##
##
##                                     disch_disp_modified
##  Discharged to home                              :58767
##  Discharged to home with Home Health Service:12698
##  Discharged/Transferred to SNF                   :13614
##  Other                                           :14413
##
##
##
##                      adm_src_mod       adm_typ_mod     age_mod
##  Emergency Room            :56632   Elective :18507   0-19 :  842
##  Other                     : 7308   Emergency:52900   20-59:31670
##  Physician Referral        :28854   Other    :10299   60-79:47456
##  Transfer from Home Health: 6698   Urgent   :17786   80+  :19524
##
##
##
##    diag1_mod       diag2_mod       diag3_mod     readmitted not_unique
##  Other  :45951   Other  :36612   Other  :41387   0:88323    0:69667
##  428    : 6739   276    : 6638   250    :11227   1:11169    1:29825
##  414    : 6406   428    : 6522   401    : 8097
##  786    : 3938   250    : 5879   276    : 5100
##  410    : 3518   427    : 4933   428    : 4496
##  486    : 3425   401    : 3641   427    : 3871
##  (Other):29515   (Other):35267   (Other):25314
```

**Figure 2 - readmitted vs race**

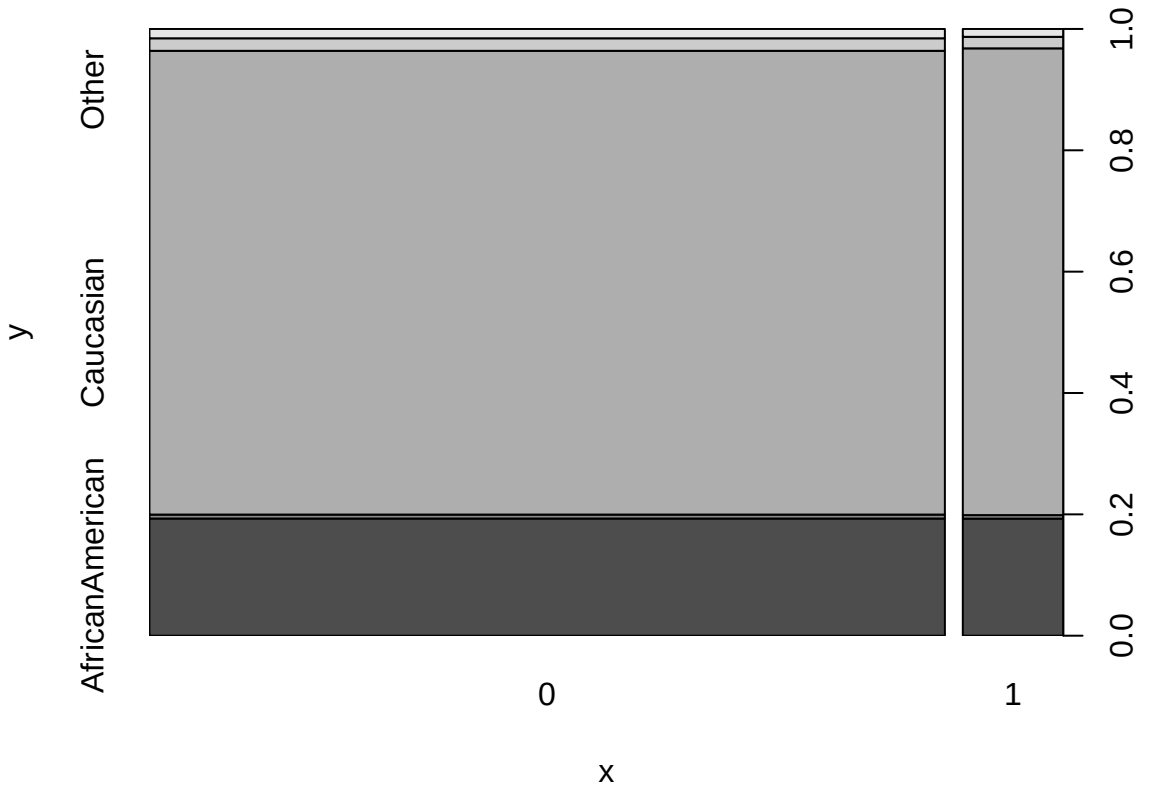

**Figure 3 - readmitted vs gender**

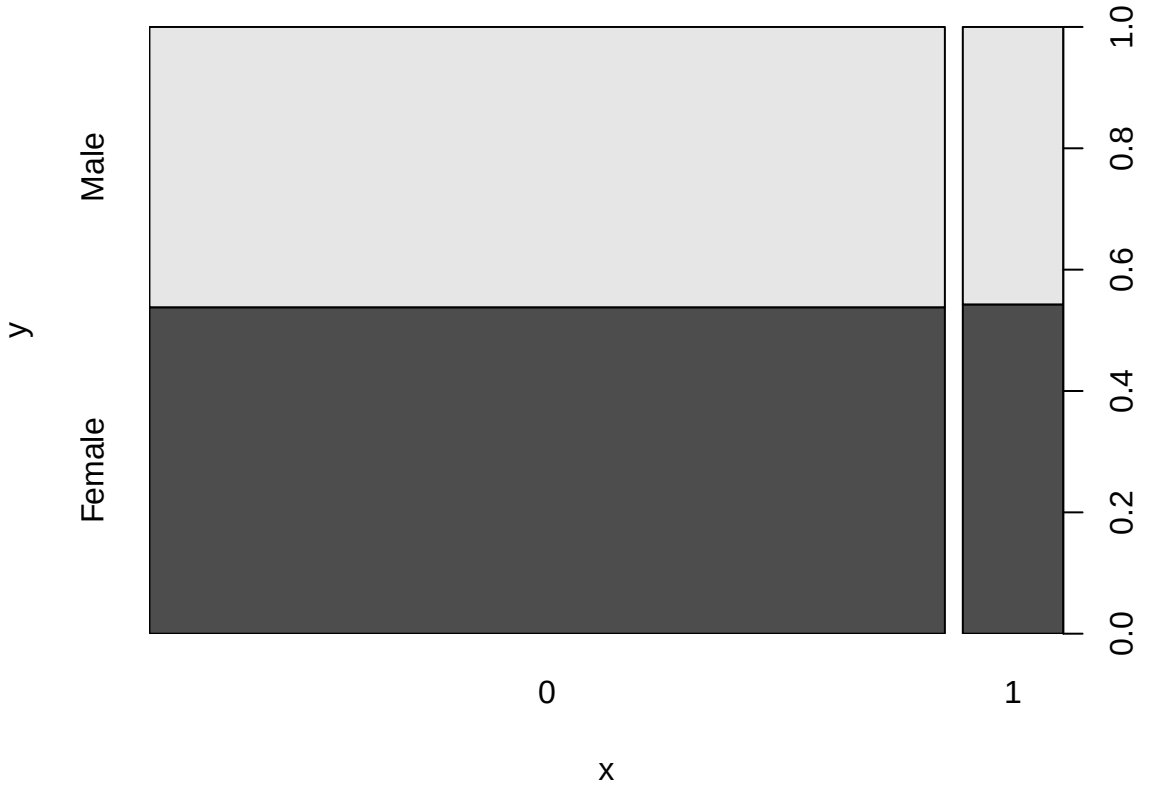


**Figure 4 - readmitted vs num_medications**

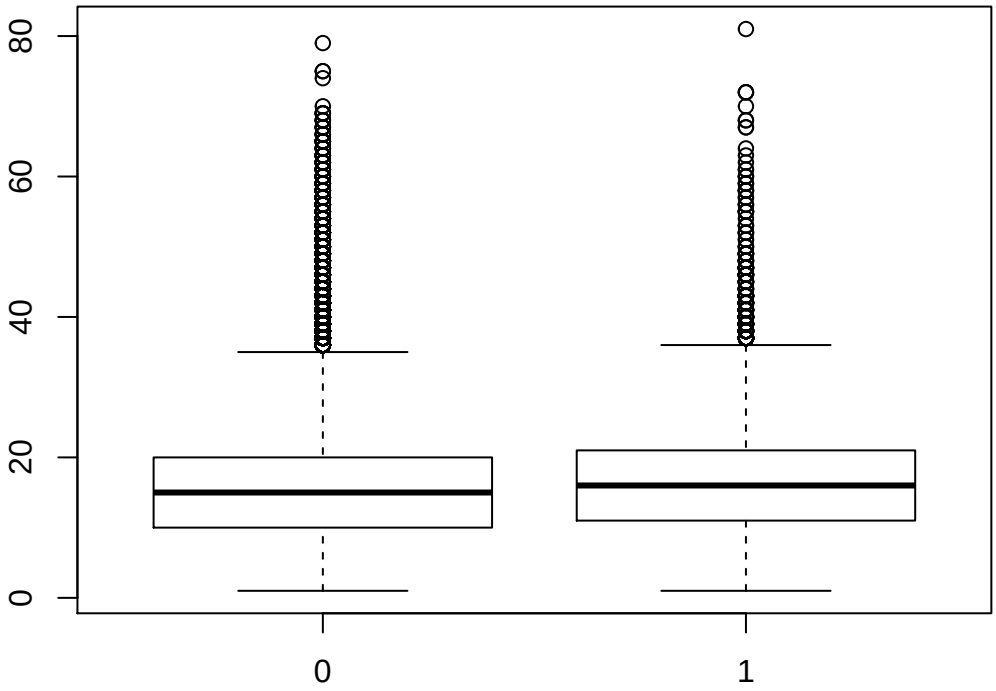

**Figure 5 - readmitted vs not_unique**

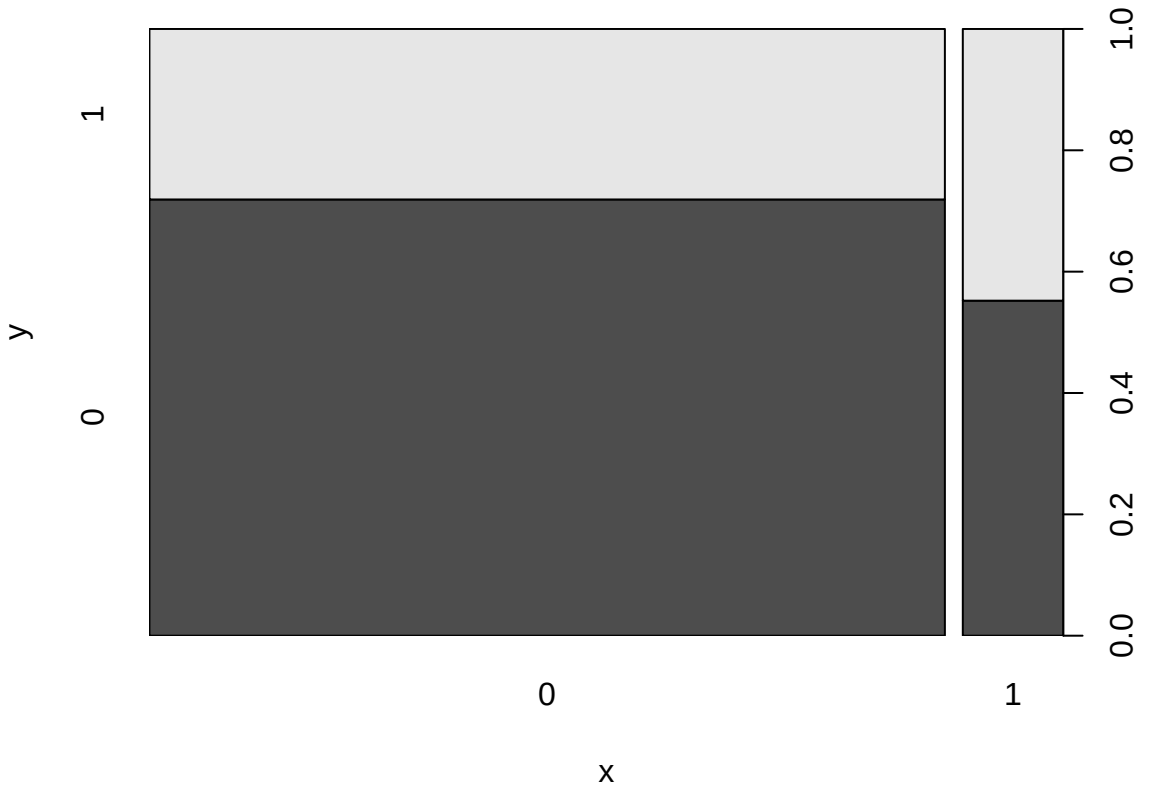


**Figure 6 - readmitted vs number_diagnoses**

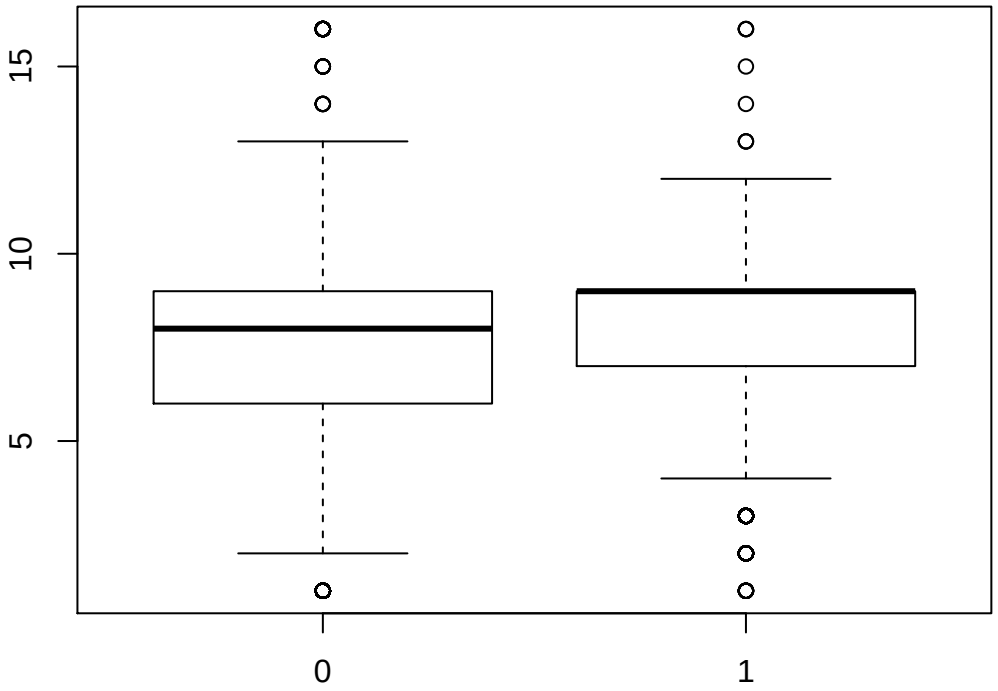

**Figure 7 - readmitted vs age_mod**

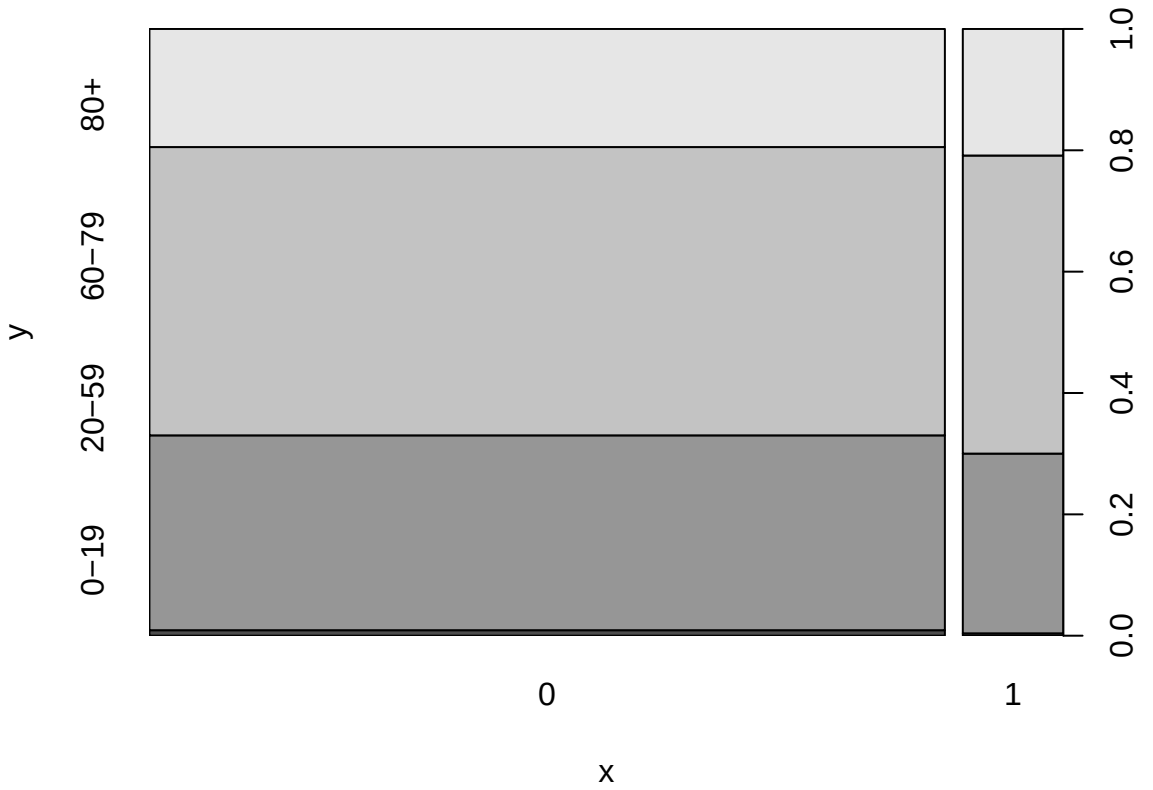

**Figure 8 - Binomial Deviance vs log(Lambda)**

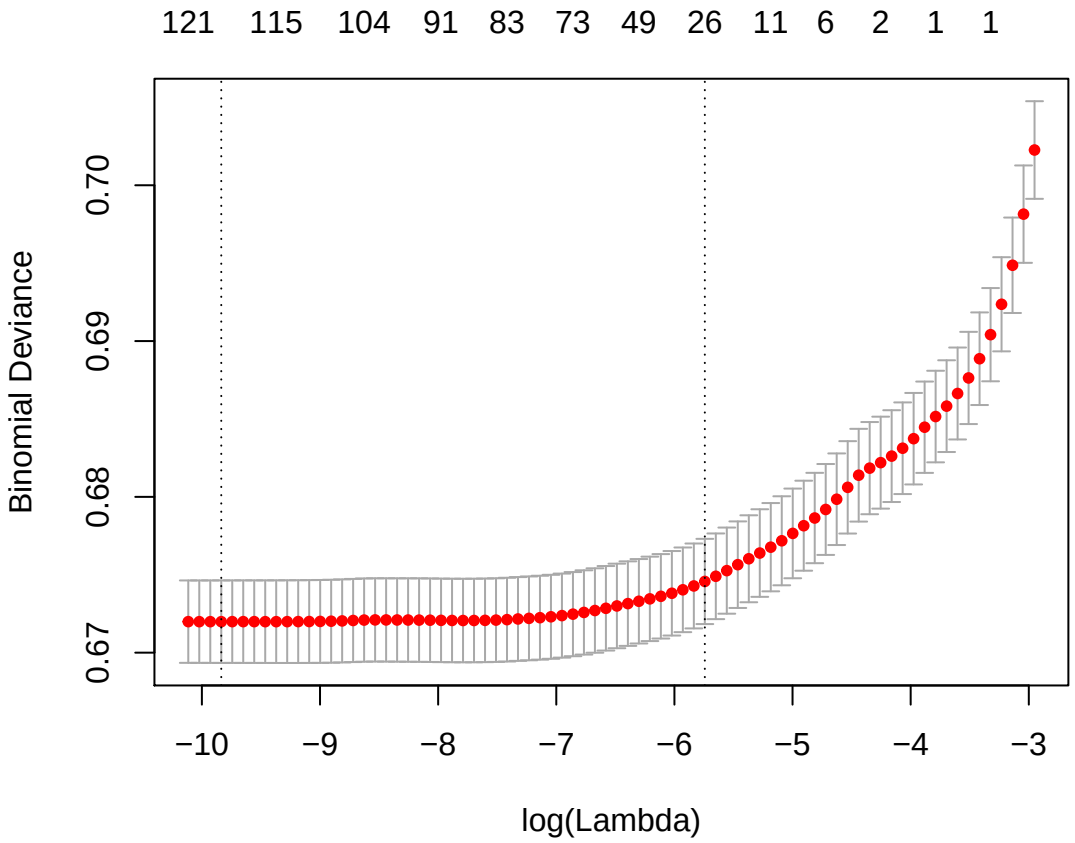


**Figure 9 - ANOVA with insignificant variables**

```
## Analysis of Deviance Table (Type II tests)
##
## Response: readmitted
##                       LR Chisq Df Pr(>Chisq)
## time_in_hospital         1.25  1  0.2635002
## num_medications          5.34  1  0.0209013 *
## number_emergency        18.49  1  1.704e-05 ***
## number_inpatient       757.16  1  < 2.2e-16 ***
## number_diagnoses         2.82  1  0.0929568 .
## metformin               16.61  3  0.0008518 ***
## insulin                 12.09  3  0.0070723 **
## diabetesMed             35.44  1  2.634e-09 ***
## disch_disp_modified    215.56  3  < 2.2e-16 ***
## age_mod                 37.54  3  3.532e-08 ***
## diag1_mod              200.26 23  < 2.2e-16 ***
## diag2_mod               72.30 24  9.736e-07 ***
## diag3_mod               89.13 20  1.052e-10 ***
## not_unique             110.79  1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Figure 10 - ANOVA with no insignificant variables

```
## Analysis of Deviance Table (Type II tests)
##
## Response: readmitted
##                      LR Chisq Df Pr(>Chisq)
## num_medications         9.47  1  0.0020872 **
## number_emergency       18.18  1  2.005e-05 ***
## number_inpatient      759.42  1  < 2.2e-16 ***
## number_diagnoses        2.92  1  0.0874612 .
## metformin              16.71  3  0.0008108 ***
## insulin                12.03  3  0.0072824 **
## diabetesMed            35.33  1  2.785e-09 ***
## disch_disp_modified   226.58  3  < 2.2e-16 ***
## age_mod                37.77  3  3.163e-08 ***
## diag1_mod             201.96 23  < 2.2e-16 ***
## diag2_mod              72.82 24  8.119e-07 ***
## diag3_mod              90.20 20  6.844e-11 ***
## not_unique            110.28  1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Figure 11 - Summary of FLD Model

```
##
## Call:
## glm(formula = readmitted ~ num_medications + number_emergency +
##     number_inpatient + number_diagnoses + metformin + insulin +
##     diabetesMed + disch_disp_modified + age_mod + diag1_mod +
##     diag2_mod + diag3_mod + not_unique, family = "binomial",
##     data = cleaned_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2259  -0.5087  -0.4271  -0.3637   2.6335
##
## Coefficients:
##                                                                  Estimate
## (Intercept)                                                     -3.184870
## num_medications                                                  0.004406
## number_emergency                                                 0.036187
## number_inpatient                                                 0.212930
## number_diagnoses                                                 0.011913
## metforminNo                                                     -0.149256
## metforminSteady                                                 -0.249735
## metforminUp                                                     -0.392882
## insulinNo                                                       -0.087030
## insulinSteady                                                   -0.115629
## insulinUp                                                       -0.081941
## diabetesMedYes                                                   0.198635
## disch_disp_modifiedDischarged to home with Home Health Service  0.177922
## disch_disp_modifiedDischarged/Transferred to SNF                 0.351350
## disch_disp_modifiedOther                                         0.391672
```

```
## age_mod20-59                                          0.563509
## age_mod60-79                                          0.680854
## age_mod80+                                            0.653475
## diag1_mod250.8                                       -0.519734
## diag1_mod276                                         -0.202528
## diag1_mod38                                          -0.521653
## diag1_mod410                                         -0.332432
## diag1_mod414                                         -0.377711
## diag1_mod427                                         -0.463372
## diag1_mod428                                         -0.213839
## diag1_mod434                                          0.107506
## diag1_mod435                                         -0.582265
## diag1_mod486                                         -0.671122
## diag1_mod491                                         -0.326670
## diag1_mod493                                         -0.493636
## diag1_mod518                                         -0.771191
## diag1_mod577                                         -0.145077
## diag1_mod584                                         -0.332924
## diag1_mod599                                         -0.546103
## diag1_mod682                                         -0.599064
## diag1_mod715                                         -0.258969
## diag1_mod780                                         -0.496788
## diag1_mod786                                         -0.584426
## diag1_mod820                                         -0.058564
## diag1_mod996                                         -0.350088
## diag1_modOther                                       -0.325367
## diag2_mod250.01                                       0.375098
## diag2_mod250.02                                       0.167761
## diag2_mod276                                          0.194388
## diag2_mod285                                         -0.134070
## diag2_mod401                                         -0.065633
## diag2_mod403                                          0.232497
## diag2_mod411                                          0.139244
## diag2_mod413                                         -0.027346
## diag2_mod414                                          0.014443
## diag2_mod424                                          0.106253
## diag2_mod425                                          0.154271
## diag2_mod427                                          0.105137
## diag2_mod428                                          0.189255
## diag2_mod486                                         -0.058105
## diag2_mod491                                          0.270893
## diag2_mod496                                          0.082853
## diag2_mod518                                         -0.130802
## diag2_mod584                                          0.027134
## diag2_mod585                                          0.164777
## diag2_mod599                                          0.034573
## diag2_mod682                                          0.225077
## diag2_mod707                                          0.205613
## diag2_mod780                                          0.008509
## diag2_modOther                                        0.164105
## diag3_mod250                                          0.090167
## diag3_mod250.02                                       0.343156
## diag3_mod250.6                                        0.644717
## diag3_mod272                                          0.010828
```

```
## diag3_mod276                                                    0.173222
## diag3_mod285                                                    0.055267
## diag3_mod401                                                    0.080027
## diag3_mod403                                                    0.420387
## diag3_mod414                                                    0.071752
## diag3_mod424                                                    0.125526
## diag3_mod425                                                    0.162767
## diag3_mod427                                                    0.182674
## diag3_mod428                                                    0.193671
## diag3_mod496                                                    0.302916
## diag3_mod585                                                    0.385904
## diag3_mod599                                                    0.165886
## diag3_mod707                                                    0.301244
## diag3_mod780                                                    0.194282
## diag3_modOther                                                  0.246050
## diag3_modV45                                                    0.020080
## not_unique1                                                     0.264847
##                                                               Std. Error
## (Intercept)                                                     0.245326
## num_medications                                                 0.001426
## number_emergency                                                0.008482
## number_inpatient                                                0.007590
## number_diagnoses                                                0.006980
## metforminNo                                                     0.131803
## metforminSteady                                                 0.133312
## metforminUp                                                     0.173201
## insulinNo                                                       0.036465
## insulinSteady                                                   0.033245
## insulinUp                                                       0.039609
## diabetesMedYes                                                  0.033416
## disch_disp_modifiedDischarged to home with Home Health Service  0.031996
## disch_disp_modifiedDischarged/Transferred to SNF                0.031737
## disch_disp_modifiedOther                                        0.029368
## age_mod20-59                                                    0.163636
## age_mod60-79                                                    0.163986
## age_mod80+                                                      0.165264
## diag1_mod250.8                                                  0.113576
## diag1_mod276                                                    0.106075
## diag1_mod38                                                     0.113592
## diag1_mod410                                                    0.100249
## diag1_mod414                                                    0.098586
## diag1_mod427                                                    0.105697
## diag1_mod428                                                    0.088991
## diag1_mod434                                                    0.102366
## diag1_mod435                                                    0.144876
## diag1_mod486                                                    0.101593
## diag1_mod491                                                    0.104528
## diag1_mod493                                                    0.135907
## diag1_mod518                                                    0.133722
## diag1_mod577                                                    0.123555
## diag1_mod584                                                    0.112330
## diag1_mod599                                                    0.115360
## diag1_mod682                                                    0.112780
## diag1_mod715                                                    0.112258
```

```
## diag1_mod780                                          0.111894
## diag1_mod786                                          0.102225
## diag1_mod820                                          0.119511
## diag1_mod996                                          0.105354
## diag1_modOther                                        0.081232
## diag2_mod250.01                                       0.095098
## diag2_mod250.02                                       0.090639
## diag2_mod276                                          0.068442
## diag2_mod285                                          0.107111
## diag2_mod401                                          0.086613
## diag2_mod403                                          0.080630
## diag2_mod411                                          0.098824
## diag2_mod413                                          0.131914
## diag2_mod414                                          0.092426
## diag2_mod424                                          0.117720
## diag2_mod425                                          0.101057
## diag2_mod427                                          0.072754
## diag2_mod428                                          0.068660
## diag2_mod486                                          0.107840
## diag2_mod491                                          0.099346
## diag2_mod496                                          0.079724
## diag2_mod518                                          0.108823
## diag2_mod584                                          0.097807
## diag2_mod585                                          0.089700
## diag2_mod599                                          0.080650
## diag2_mod682                                          0.099667
## diag2_mod707                                          0.088745
## diag2_mod780                                          0.104995
## diag2_modOther                                        0.057641
## diag3_mod250                                          0.124522
## diag3_mod250.02                                       0.147005
## diag3_mod250.6                                        0.147069
## diag3_mod272                                          0.150789
## diag3_mod276                                          0.129756
## diag3_mod285                                          0.155311
## diag3_mod401                                          0.125430
## diag3_mod403                                          0.135677
## diag3_mod414                                          0.134732
## diag3_mod424                                          0.160133
## diag3_mod425                                          0.154954
## diag3_mod427                                          0.132649
## diag3_mod428                                          0.131019
## diag3_mod496                                          0.135826
## diag3_mod585                                          0.138633
## diag3_mod599                                          0.141661
## diag3_mod707                                          0.145906
## diag3_mod780                                          0.150089
## diag3_modOther                                        0.121865
## diag3_modV45                                          0.154597
## not_unique1                                           0.025039
##                                                      z value
## (Intercept)                                         -12.982
## num_medications                                       3.091
## number_emergency                                      4.266
```

```
## number_inpatient                                                28.055
## number_diagnoses                                                 1.707
## metforminNo                                                     -1.132
## metforminSteady                                                 -1.873
## metforminUp                                                     -2.268
## insulinNo                                                       -2.387
## insulinSteady                                                   -3.478
## insulinUp                                                       -2.069
## diabetesMedYes                                                   5.944
## disch_disp_modifiedDischarged to home with Home Health Service   5.561
## disch_disp_modifiedDischarged/Transferred to SNF                11.071
## disch_disp_modifiedOther                                        13.336
## age_mod20-59                                                     3.444
## age_mod60-79                                                     4.152
## age_mod80+                                                       3.954
## diag1_mod250.8                                                  -4.576
## diag1_mod276                                                    -1.909
## diag1_mod38                                                     -4.592
## diag1_mod410                                                    -3.316
## diag1_mod414                                                    -3.831
## diag1_mod427                                                    -4.384
## diag1_mod428                                                    -2.403
## diag1_mod434                                                     1.050
## diag1_mod435                                                    -4.019
## diag1_mod486                                                    -6.606
## diag1_mod491                                                    -3.125
## diag1_mod493                                                    -3.632
## diag1_mod518                                                    -5.767
## diag1_mod577                                                    -1.174
## diag1_mod584                                                    -2.964
## diag1_mod599                                                    -4.734
## diag1_mod682                                                    -5.312
## diag1_mod715                                                    -2.307
## diag1_mod780                                                    -4.440
## diag1_mod786                                                    -5.717
## diag1_mod820                                                    -0.490
## diag1_mod996                                                    -3.323
## diag1_modOther                                                  -4.005
## diag2_mod250.01                                                  3.944
## diag2_mod250.02                                                  1.851
## diag2_mod276                                                     2.840
## diag2_mod285                                                    -1.252
## diag2_mod401                                                    -0.758
## diag2_mod403                                                     2.884
## diag2_mod411                                                     1.409
## diag2_mod413                                                    -0.207
## diag2_mod414                                                     0.156
## diag2_mod424                                                     0.903
## diag2_mod425                                                     1.527
## diag2_mod427                                                     1.445
## diag2_mod428                                                     2.756
## diag2_mod486                                                    -0.539
## diag2_mod491                                                     2.727
## diag2_mod496                                                     1.039
```

```
## diag2_mod518                                                    -1.202
## diag2_mod584                                                     0.277
## diag2_mod585                                                     1.837
## diag2_mod599                                                     0.429
## diag2_mod682                                                     2.258
## diag2_mod707                                                     2.317
## diag2_mod780                                                     0.081
## diag2_modOther                                                   2.847
## diag3_mod250                                                     0.724
## diag3_mod250.02                                                  2.334
## diag3_mod250.6                                                   4.384
## diag3_mod272                                                     0.072
## diag3_mod276                                                     1.335
## diag3_mod285                                                     0.356
## diag3_mod401                                                     0.638
## diag3_mod403                                                     3.098
## diag3_mod414                                                     0.533
## diag3_mod424                                                     0.784
## diag3_mod425                                                     1.050
## diag3_mod427                                                     1.377
## diag3_mod428                                                     1.478
## diag3_mod496                                                     2.230
## diag3_mod585                                                     2.784
## diag3_mod599                                                     1.171
## diag3_mod707                                                     2.065
## diag3_mod780                                                     1.294
## diag3_modOther                                                   2.019
## diag3_modV45                                                     0.130
## not_unique1                                                     10.577
##                                                                Pr(>|z|)
## (Intercept)                                                     < 2e-16
## num_medications                                                0.001995
## number_emergency                                               1.99e-05
## number_inpatient                                                < 2e-16
## number_diagnoses                                               0.087874
## metforminNo                                                    0.257459
## metforminSteady                                                0.061025
## metforminUp                                                    0.023307
## insulinNo                                                      0.017001
## insulinSteady                                                  0.000505
## insulinUp                                                      0.038568
## diabetesMedYes                                                 2.78e-09
## disch_disp_modifiedDischarged to home with Home Health Service 2.69e-08
## disch_disp_modifiedDischarged/Transferred to SNF                < 2e-16
## disch_disp_modifiedOther                                        < 2e-16
## age_mod20-59                                                   0.000574
## age_mod60-79                                                   3.30e-05
## age_mod80+                                                     7.68e-05
## diag1_mod250.8                                                 4.74e-06
## diag1_mod276                                                   0.056224
## diag1_mod38                                                    4.38e-06
## diag1_mod410                                                   0.000913
## diag1_mod414                                                   0.000127
## diag1_mod427                                                   1.17e-05
```

```
## diag1_mod428                                    0.016265
## diag1_mod434                                    0.293621
## diag1_mod435                                    5.84e-05
## diag1_mod486                                    3.95e-11
## diag1_mod491                                    0.001777
## diag1_mod493                                    0.000281
## diag1_mod518                                    8.06e-09
## diag1_mod577                                    0.240319
## diag1_mod584                                    0.003039
## diag1_mod599                                    2.20e-06
## diag1_mod682                                    1.09e-07
## diag1_mod715                                    0.021060
## diag1_mod780                                    9.00e-06
## diag1_mod786                                    1.08e-08
## diag1_mod820                                    0.624115
## diag1_mod996                                    0.000891
## diag1_mod0ther                                  6.19e-05
## diag2_mod250.01                                 8.00e-05
## diag2_mod250.02                                 0.064189
## diag2_mod276                                    0.004509
## diag2_mod285                                    0.210681
## diag2_mod401                                    0.448586
## diag2_mod403                                    0.003933
## diag2_mod411                                    0.158834
## diag2_mod413                                    0.835777
## diag2_mod414                                    0.875827
## diag2_mod424                                    0.366741
## diag2_mod425                                    0.126866
## diag2_mod427                                    0.148429
## diag2_mod428                                    0.005844
## diag2_mod486                                    0.590017
## diag2_mod491                                    0.006396
## diag2_mod496                                    0.298693
## diag2_mod518                                    0.229378
## diag2_mod584                                    0.781458
## diag2_mod585                                    0.066213
## diag2_mod599                                    0.668153
## diag2_mod682                                    0.023928
## diag2_mod707                                    0.020509
## diag2_mod780                                    0.935408
## diag2_mod0ther                                  0.004413
## diag3_mod250                                    0.468999
## diag3_mod250.02                                 0.019579
## diag3_mod250.6                                  1.17e-05
## diag3_mod272                                    0.942754
## diag3_mod276                                    0.181883
## diag3_mod285                                    0.721955
## diag3_mod401                                    0.523461
## diag3_mod403                                    0.001945
## diag3_mod414                                    0.594343
## diag3_mod424                                    0.433107
## diag3_mod425                                    0.293525
## diag3_mod427                                    0.168475
## diag3_mod428                                    0.139356
```

```
## diag3_mod496                                                    0.025736
## diag3_mod585                                                    0.005375
## diag3_mod599                                                    0.241596
## diag3_mod707                                                    0.038956
## diag3_mod780                                                    0.195512
## diag3_modOther                                                  0.043483
## diag3_modV45                                                    0.896655
## not_unique1                                                      < 2e-16
##
## (Intercept)                                                     ***
## num_medications                                                 **
## number_emergency                                                ***
## number_inpatient                                                ***
## number_diagnoses                                                .
## metforminNo
## metforminSteady                                                 .
## metforminUp                                                     *
## insulinNo                                                       *
## insulinSteady                                                   ***
## insulinUp                                                       *
## diabetesMedYes                                                  ***
## disch_disp_modifiedDischarged to home with Home Health Service ***
## disch_disp_modifiedDischarged/Transferred to SNF                ***
## disch_disp_modifiedOther                                        ***
## age_mod20-59                                                    ***
## age_mod60-79                                                    ***
## age_mod80+                                                      ***
## diag1_mod250.8                                                  ***
## diag1_mod276                                                    .
## diag1_mod38                                                     ***
## diag1_mod410                                                    ***
## diag1_mod414                                                    ***
## diag1_mod427                                                    ***
## diag1_mod428                                                    *
## diag1_mod434
## diag1_mod435                                                    ***
## diag1_mod486                                                    ***
## diag1_mod491                                                    **
## diag1_mod493                                                    ***
## diag1_mod518                                                    ***
## diag1_mod577
## diag1_mod584                                                    **
## diag1_mod599                                                    ***
## diag1_mod682                                                    ***
## diag1_mod715                                                    *
## diag1_mod780                                                    ***
## diag1_mod786                                                    ***
## diag1_mod820
## diag1_mod996                                                    ***
## diag1_modOther                                                  ***
## diag2_mod250.01                                                 ***
## diag2_mod250.02                                                 .
## diag2_mod276                                                    **
## diag2_mod285
```

```
## diag2_mod401
## diag2_mod403                                                    **
## diag2_mod411
## diag2_mod413
## diag2_mod414
## diag2_mod424
## diag2_mod425
## diag2_mod427
## diag2_mod428                                                    **
## diag2_mod486
## diag2_mod491                                                    **
## diag2_mod496
## diag2_mod518
## diag2_mod584
## diag2_mod585                                                    .
## diag2_mod599
## diag2_mod682                                                    *
## diag2_mod707                                                    *
## diag2_mod780
## diag2_modOther                                                  **
## diag3_mod250
## diag3_mod250.02                                                 *
## diag3_mod250.6                                                  ***
## diag3_mod272
## diag3_mod276
## diag3_mod285
## diag3_mod401
## diag3_mod403                                                    **
## diag3_mod414
## diag3_mod424
## diag3_mod425
## diag3_mod427
## diag3_mod428
## diag3_mod496                                                    *
## diag3_mod585                                                    **
## diag3_mod599
## diag3_mod707                                                    *
## diag3_mod780
## diag3_modOther                                                  *
## diag3_modV45
## not_unique1                                                     ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 69886  on 99491  degrees of freedom
## Residual deviance: 66670  on 99406  degrees of freedom
## AIC: 66842
##
## Number of Fisher Scoring iterations: 5
```

**Figure 12 - ROC Curve of FLD Model**

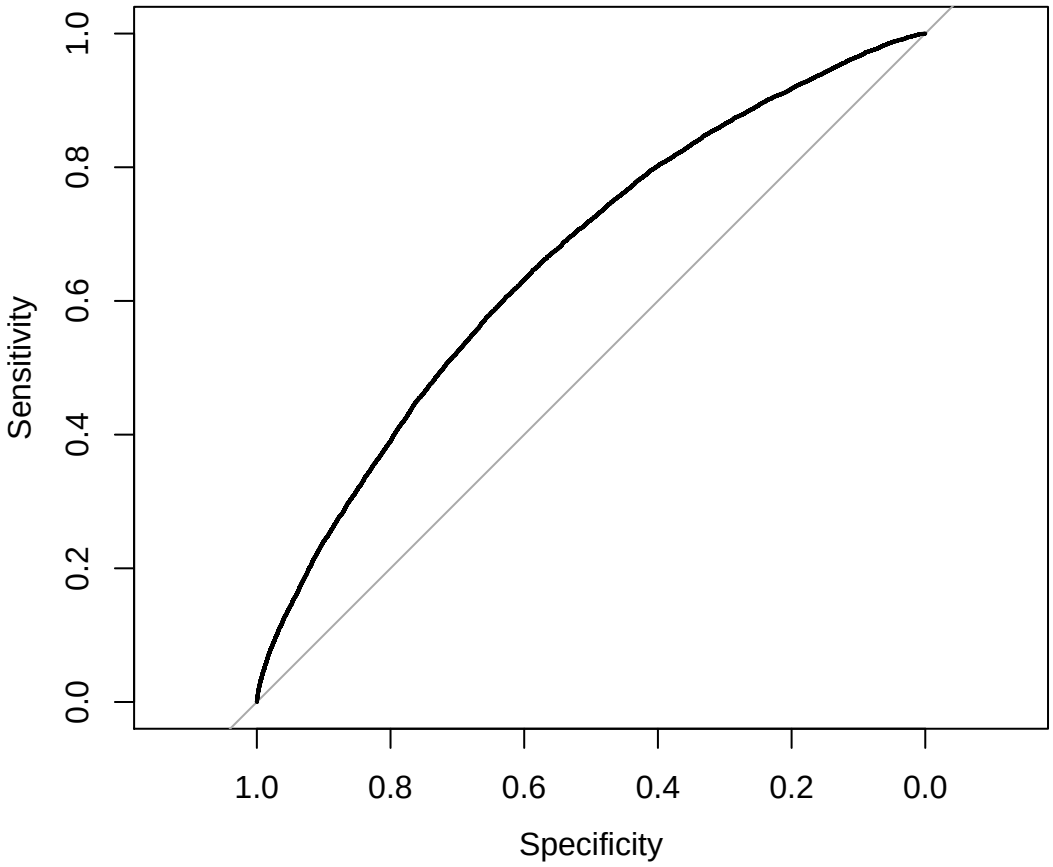

**Figure 13 - Misclassification Error vs log(Lambda)**

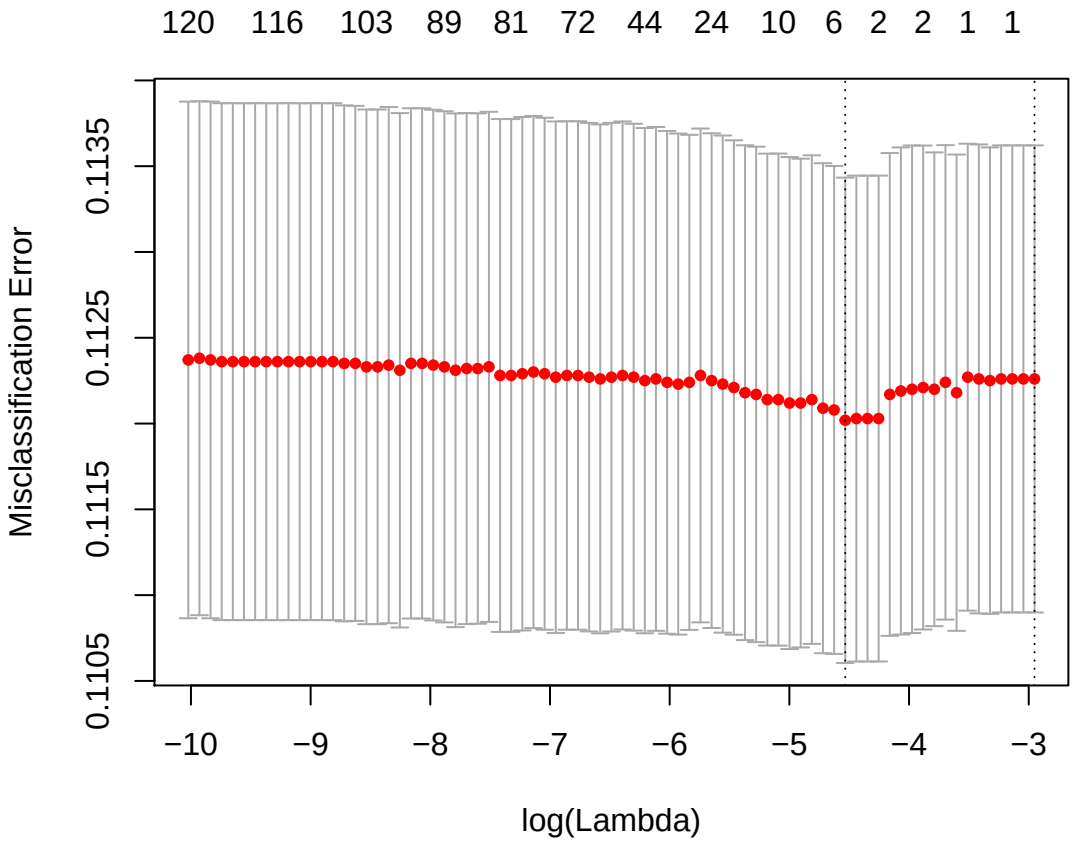

**Figure 14 - AUC vs log(Lambda)**

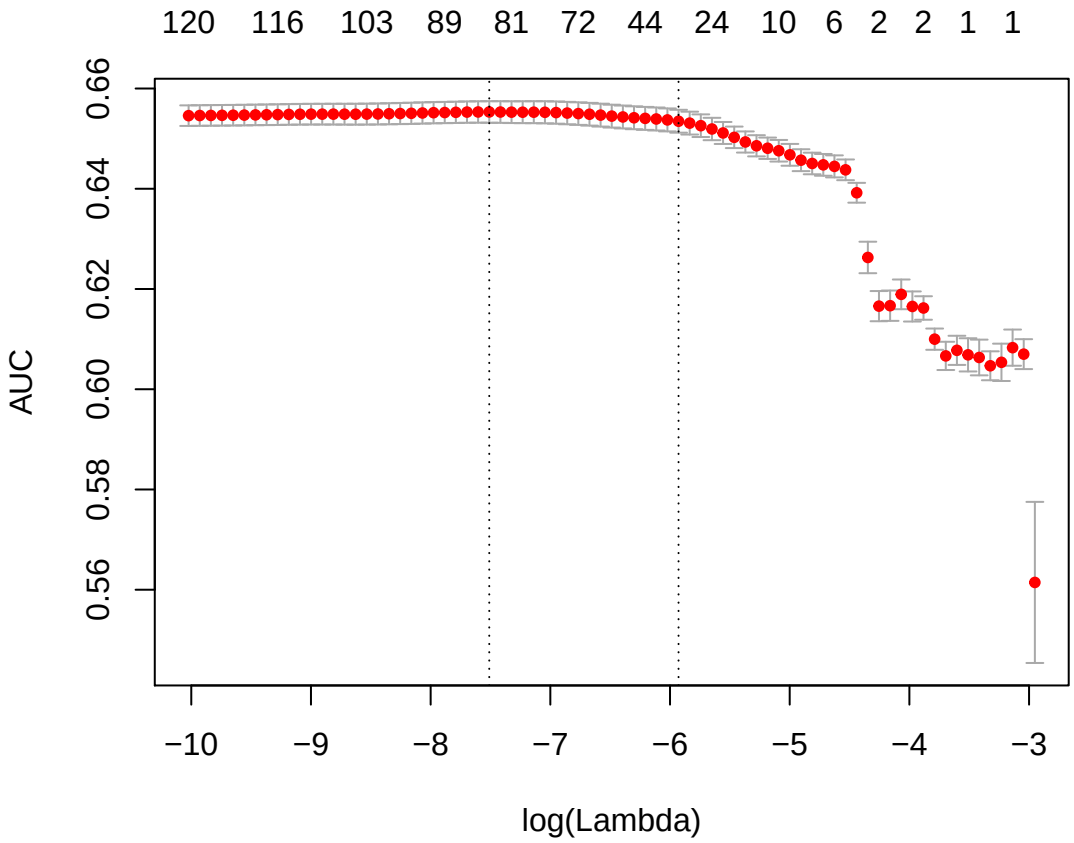

**Figure 15 - ROC Curve of MCE Model**

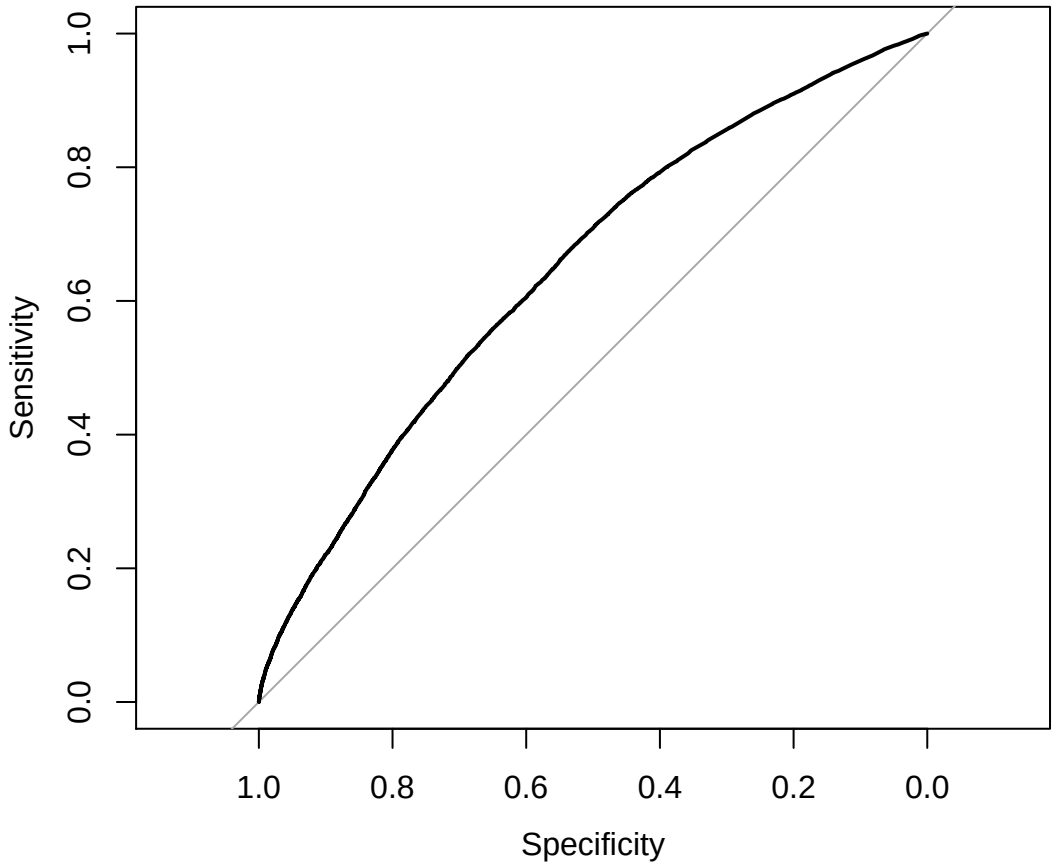

**Figure 16 - ROC Curve of AUC Model**

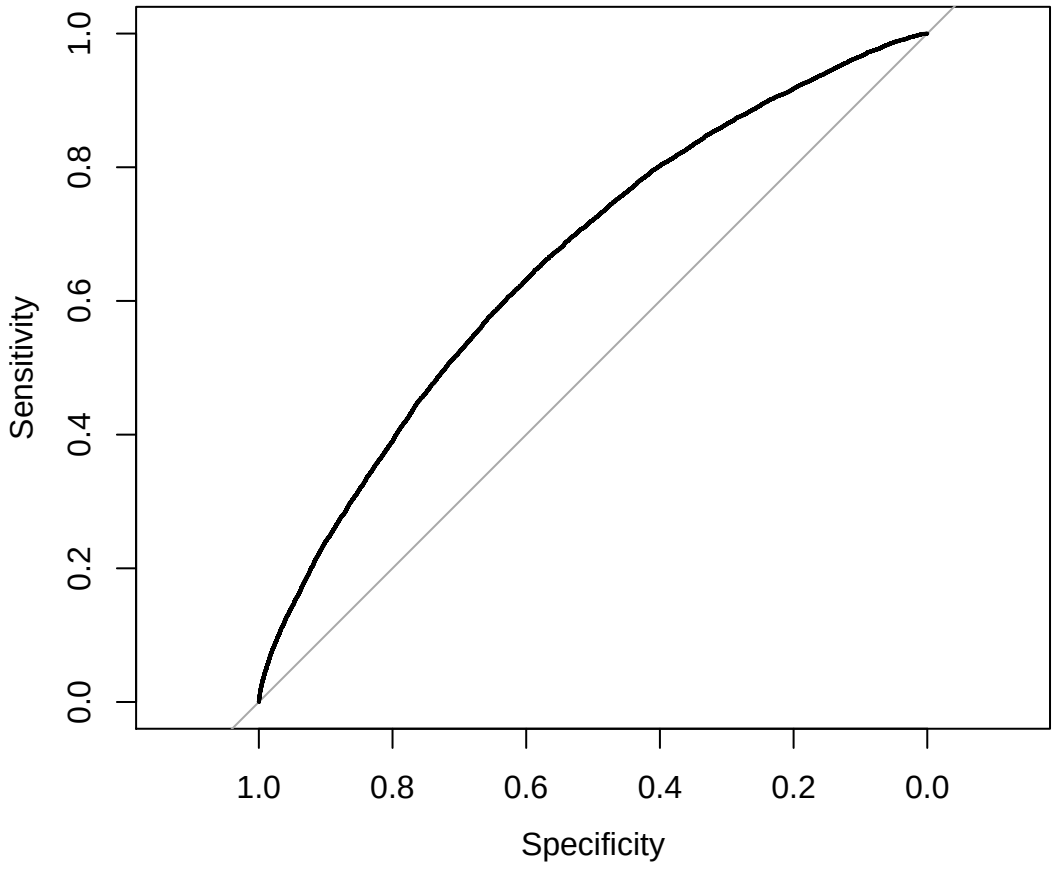


**Figure 17 - Elastic Net ANOVA with insignificant variables**

```
## Analysis of Deviance Table (Type II tests)
##
## Response: readmitted
##                      LR Chisq Df Pr(>Chisq)
## time_in_hospital        1.88  1   0.170617
## num_medications         3.48  1   0.062200 .
## number_emergency       18.41  1  1.783e-05 ***
## number_inpatient      787.74  1  < 2.2e-16 ***
## number_diagnoses        4.27  1   0.038901 *
## insulin                15.92  3   0.001175 **
## diabetesMed            24.21  1  8.639e-07 ***
## disch_disp_modified   216.44  3  < 2.2e-16 ***
## age_mod                33.65  3  2.343e-07 ***
## diag1_mod             200.38 23  < 2.2e-16 ***
## diag3_mod             111.00 20  1.292e-14 ***
## not_unique            111.72  1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Figure 18 - Elastic Net ANOVA with no insignificant variables

```
## Analysis of Deviance Table (Type II tests)
##
## Response: readmitted
##                      LR Chisq Df Pr(>Chisq)
## number_emergency       17.66  1  2.639e-05 ***
## number_inpatient      794.62  1  < 2.2e-16 ***
## number_diagnoses        7.06  1  0.0078733 **
## insulin                19.03  3  0.0002699 ***
## diabetesMed            26.45  1  2.707e-07 ***
## disch_disp_modified   243.37  3  < 2.2e-16 ***
## age_mod                34.52  3  1.538e-07 ***
## diag1_mod             201.51 23  < 2.2e-16 ***
## diag3_mod             115.39 20  2.023e-15 ***
## not_unique            110.77  1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```