

Lab_practice

Ximu Wang

2019/3/4

```
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##   filter, lag
## The following objects are masked from 'package:base':
##   intersect, setdiff, setequal, union
library(pROC)

## Type 'citation("pROC")' for a citation.

##
## Attaching package: 'pROC'
## The following objects are masked from 'package:stats':
##   cov, smooth, var
library(rsample)

## Loading required package: tidyr
library(xgboost)

##
## Attaching package: 'xgboost'
## The following object is masked from 'package:dplyr':
##   slice
library(ggplot2)
library(glmnet)

## Loading required package: Matrix

##
## Attaching package: 'Matrix'
## The following object is masked from 'package:tidyr':
##   expand
## Loading required package: foreach
## Loaded glmnet 2.0-16
##
## Attaching package: 'glmnet'
```

```

## The following object is masked from 'package:pROC':
##
##      auc

library(car)

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##      recode

library(MASS)

##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##      select

library(class)
library(rpart)
library(ROCR)

## Loading required package: gplots

##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##      lowess

library(caret)

## Loading required package: lattice

```

Import the data set and clean it.

```

labdata <- read.csv("INFSCI_2160_R_LAB_DATASET.csv")
head(labdata)

##   ESRD_FLG_COMP AGE_ON_CONTACT_DATE   BMI BP_SYSTOLIC BP_DIASTOLIC
## 1              0          69.0     NA        142           80
## 2              0          72.1 32.54       192          105
## 3              0          39.7 29.43       126           93
## 4              0          66.0     NA        NA           NA
## 5              0          76.6     NA        NA           NA
## 6              0          67.6     NA        130           67
##   EGFR_CLOSEST GLUCOSE_CLOSEST_F WBC_CLOSEST_F HGB_CLOSEST_F
## 1        47.0            102         5.4        14.7
## 2        53.0            175         6.4        10.4
## 3        55.0             87        NA           NA
## 4        51.0             84        10.7        9.3

```

## 5	52.6	117	NA	NA
## 6	53.0	90	4.4	10.6
## 1	145	4.5	85	99.0
## 2	138	4.2	36	84.6
## 3	138	4.2	61	157.0
## 4	141	4.1	NA	NA
## 5	141	4.2	47	111.0
## 6	138	4.5	41	116.0
## 1	NA	NA	NA	NA
## 2	NA	NA	NA	NA
## 3	NA	NA	NA	NA
## 4	NA	NA	NA	NA
## 5	NA	NA	NA	NA
## 6	NA	NA	NA	NA
## 1	NA	NA	NA	NA
## 2	3.0	76.0	12.0	0.0
## 3	NA	NA	NA	NA
## 4	1.8	70.4	11.4	0.6
## 5	NA	NA	NA	NA
## 6	2.0	68.0	3.0	1.0
## 1	NA	NA	NA	NA
## 2	NA	NA	NA	NA
## 3	NA	NA	NA	NA
## 4	0.731	5.7	NA	NA
## 5	NA	NA	NA	NA
## 6	0.790	NA	NA	NA
## 1	NA	NA	NA	NA
## 2	NA	NA	NA	NA
## 3	NA	NA	NA	NA
## 4	2.3	756	NA	1910
## 5	NA	NA	NA	NA
## 6	2.0	325	NA	NA
## 1	207	32	43.8	
## 2	109	NA	31.1	
## 3	NA	NA	NA	
## 4	310	42	28.2	
## 5	NA	NA	NA	
## 6	238	NA	31.6	
## 1	2	2	NA	32.4
## 2	NA	NA	8.7	29.0
## 3	NA	NA	NA	NA
## 4	4	4	8.6	28.7
## 5	NA	NA	NA	NA
## 6	NA	NA	7.3	32.6
## 1	96.3	15.2	33.6	4.55
## 2	86.4	16.6	33.6	3.60

## 3	NA	NA	NA	NA	
## 4	87.1	15.2	33.0	3.24	
## 5	NA	NA	NA	NA	
## 6	96.8	NA	33.7	3.26	
## LYMPHOCYTE_CLOSEST_F	CA125_CLOSEST_F	BILIRUBIN_CLOSEST_F	ALT_CLOSEST_F		
## 1	NA	NA	0.5	10	
## 2	9.0	NA	0.9	29	
## 3	NA	NA	NA	NA	
## 4	15.8	NA	0.6	140	
## 5	NA	NA	0.5	15	
## 6	NA	NA	0.4	19	
## AST_CLOSEST_F	FOLLOW_UP_EGFR_VALUE	FEMALE	RACE_F	URINEPROTEIN_CLOSEST_F	
## 1	15	57.0	1	1	
## 2	19	52.0	0	1	
## 3	NA	55.0	0	2	
## 4	21	54.0	1	1	
## 5	19	51.4	1	1	
## 6	20	48.0	1	1	
## OBESE_HST_F	MORBIDOBEBESE_HST_F	PH_HST_F	COPD_HST_F	CHF_HST_F	AFIB_HST_F
## 1	0	0	0	1	0
## 2	1	0	0	0	0
## 3	1	0	0	0	0
## 4	1	1	0	0	0
## 5	1	0	0	1	0
## 6	0	0	0	0	0
## DIAB_HST_F	CAD_HST_F	OSTEO_HST_F	HTN_HST_F	CHOLESTEROL_CLOSEST_F	
## 1	0	0	0	1	201
## 2	0	0	0	0	157
## 3	0	0	0	1	240
## 4	0	0	0	1	NA
## 5	0	0	0	1	210
## 6	0	0	1	0	189
## CANCER_HST_F	LUNG_CANCER_HST_F	ASTHMA_HST_F	GERD_HST_F		
## 1	1	0	0	0	
## 2	0	0	0	0	
## 3	0	0	1	0	
## 4	0	0	0	0	
## 5	0	0	1	1	
## 6	1	0	0	1	
## FIBROMYALGIA_HST_F	DEPRESSION_HST_F	PSORIATIC_ARTHRITIS_HST_F			
## 1	0	0	0	0	
## 2	0	0	0	0	
## 3	0	0	0	0	
## 4	0	0	0	0	
## 5	0	0	0	0	
## 6	0	0	0	0	
## RHEUM_ARTHRITIS_HST_F	LUPUS_HST_F	VTVF_HST_F	STROKE_HST_F		
## 1	0	0	0	0	
## 2	0	0	0	0	
## 3	0	0	0	0	
## 4	0	0	0	0	
## 5	0	0	0	0	
## 6	0	0	0	0	
## VASCULARDISEASE_HST_F	LOWBACKPAIN_HST_F	DVT_HST_F	PE_HST_F		

```

## 1          0          0          0
## 2          0          0          0
## 3          0          0          0
## 4          0          0          0
## 5          0          1          0
## 6          0          1          0
##   HYPOTHYROIDISM_HST_F ADRENAL_INSUFFICIENCY_HST_F INFERTILITY_HST_F
## 1          0          0          0
## 2          0          0          0
## 3          0          0          0
## 4          0          0          0
## 5          0          0          0
## 6          0          0          0
##   OBS_SLEEPAPNEA_HST_F CARDIAC_ARREST_HST_F HEMO_STROKE_HST_F
## 1          0          0          0
## 2          0          0          0
## 3          0          0          0
## 4          0          0          0
## 5          0          0          0
## 6          0          0          0
##   MAJOR_BLEED_HST_F MACULAR_DEGEN_HST_F ANXIETY_HST_F HYPERLIPIDEMIA_HST_F
## 1          0          0          0          1
## 2          0          0          0          0
## 3          0          0          0          0
## 4          0          0          0          0
## 5          0          0          0          1
## 6          0          0          0          0
##   HIV_HST_F METFORMIN_FLAG OPIOIDS_FLAG ALPHA_BLOCKERS CENTRAL_ANTAGONISTS
## 1          0          0          1          0          0
## 2          0          1          0          0          0
## 3          1          0          0          0          0
## 4          0          0          1          0          0
## 5          0          0          1          0          0
## 6          0          0          1          0          0
##   RENIN_BETA_BLOCKERS ACE_INHIB ARB ALDOSTERONE_BLOCKERS VASODIALATORS
## 1          0          0          0          0          0          0
## 2          0          1          1          0          0          1
## 3          0          0          0          0          0          0
## 4          0          1          0          0          0          0
## 5          0          0          0          1          0          0
## 6          0          0          0          0          0          0
##   DIURETICS CALCIUM_BLOCKERS STATINS INSULIN_MEDS ASPIRIN WARFARIN DOACS
## 1          1          1          0          0          1          0          0
## 2          1          0          1          0          1          0          0
## 3          1          0          0          0          0          0          0
## 4          1          0          1          0          0          0          1
## 5          1          0          1          0          0          0          0
## 6          0          0          0          0          0          0          0
##   MEDROL PREDNISONE INHALED_STEROID_WITH_LABA INHALED_STEROID_WITHOUT_LABA
## 1          0          0          0          0
## 2          0          0          0          0
## 3          0          0          0          0
## 4          0          0          0          0
## 5          0          0          0          0

```

```

## 6      0      0      0      0      0
## ASTHMA_BIOLOGICS SHORT_ACTING_BRONCHO_DIALATORS TNF_INHIBITORS
## 1      0      0      0      0      0
## 2      0      0      0      0      0
## 3      0      0      0      0      0
## 4      0      0      0      0      0
## 5      0      0      0      0      0
## 6      0      1      0      0      0
## IMMUNOMODULATORS AMINOSALICYLATES CORTICOSTEROIDS ARNI ALLOPURINOL
## 1      0      0      0      0      0
## 2      0      0      0      0      0
## 3      0      0      0      0      0
## 4      0      0      0      0      0
## 5      0      0      0      0      0
## 6      1      0      0      0      0
## NSAIDS ANTIDEPRESSANTS SEIZURE MUSCLERELAXANT DIGOXIN INOTROPES
## 1      0      0      0      0      0
## 2      0      0      0      0      0
## 3      0      0      0      0      0
## 4      0      0      0      0      0
## 5      1      0      0      0      0
## 6      0      1      0      0      0
## ANTI_ARRHYTHMIC ANTIPLATELET SULFONYLUREA GLP_1AGONIST
## 1      0      0      0      0      0
## 2      0      1      1      0      0
## 3      0      0      0      0      0
## 4      0      0      0      0      0
## 5      0      0      0      0      0
## 6      0      0      0      0      0
## THIAZOLIDINEDIONE SGLT2_INHIBITOR DPP4_INHIBITOR
## 1      0      0      0      0      0
## 2      0      0      0      0      0
## 3      0      0      0      0      0
## 4      0      0      0      0      0
## 5      0      0      0      0      0
## 6      0      0      0      0      0
## ALPHA_GLUCOSIDASE_INHIBITOR AMYLINOMIMETIC RAPID_ACTING_INSULIN
## 1      0      0      0      0      0
## 2      0      0      0      0      0
## 3      0      0      0      0      0
## 4      0      0      0      0      0
## 5      0      0      0      0      0
## 6      0      0      0      0      0
## SHORT_ACTING_INSULIN INTERMEDIATE_ACTING_INSULIN LONG_ACTING_INSULIN
## 1      0      0      0      0      0
## 2      0      0      0      0      0
## 3      0      0      0      0      0
## 4      0      0      0      0      0
## 5      0      0      0      0      0
## 6      0      0      0      0      0
## PHOSPHORUS_CLOSEST_F CA_CLOSEST_F MINOCYCLINE DOXYCYCLINE MELATONIN
## 1      3.9     9.0     0     0     0
## 2      NA      8.3     0     0     0
## 3      NA      8.7     0     0     0

```

```

## 4          3.0      8.5      0      0      0
## 5          NA       9.2      0      0      0
## 6          3.1      8.8      0      0      0
##   CKD_HST FOLLOW_UP_CKD_HST TOTAL_HOSP ID
## 1          1           1        1  1
## 2          0           0        4  2
## 3          0           0       NA  3
## 4          0           0        3  4
## 5          0           0       NA  5
## 6          0           0       NA  6

```

If the proportion of NA in some features are greater than 0.2, than we discard them, beacause they are worthless in prediction.

```

i <- 1
rownum <- nrow(labdata)
while (i <= ncol(labdata)) {
  if (sum(is.na(labdata[, c(i)]))/rownum > 0.2) labdata <- labdata[, -i] else i <- i+1
}
#delete ID
labdata <- labdata[, -(i-1)]
head(labdata)

##   ESRD_FLG_COMP AGE_ON_CONTACT_DATE   BMI BP_SYSTOLIC BP_DIASTOLIC
## 1          0           69.0      NA        142            80
## 2          0           72.1 32.54      192            105
## 3          0           39.7 29.43      126            93
## 4          0           66.0      NA        NA            NA
## 5          0           76.6      NA        NA            NA
## 6          0           67.6      NA        130            67
##   EGFR_CLOSEST GLUCOSE_CLOSEST_F SODIUM_CLOSEST_F K_CLOSEST_F
## 1        47.0          102          145        4.5
## 2        53.0          175          138        4.2
## 3        55.0          87           138        4.2
## 4        51.0          84           141        4.1
## 5        52.6          117          141        4.2
## 6        53.0          90           138        4.5
##   BILIRUBIN_CLOSEST_F ALT_CLOSEST_F AST_CLOSEST_F FOLLOW_UP_EGFR_VALUE
## 1          0.5           10          15        57.0
## 2          0.9           29          19        52.0
## 3          NA            NA         NA        55.0
## 4          0.6           140          21        54.0
## 5          0.5           15           19        51.4
## 6          0.4           19           20        48.0
##   FEMALE RACE_F OBESE_HST_F MORBIDOBSE_HST_F PH_HST_F COPD_HST_F
## 1     1     1          0            0          0          1
## 2     0     1          1            0          0          0
## 3     0     2          1            0          0          0
## 4     1     1          1            1          0          0
## 5     1     1          1            0          0          1
## 6     1     1          0            0          0          0
##   CHF_HST_F AFIB_HST_F DIAB_HST_F CAD_HST_F OSTEO_HST_F HTN_HST_F
## 1     0     0          0            0          0          1
## 2     0     0          0            0          0          0
## 3     0     0          0            0          0          1

```

## 4	0	1	0	0	0	1
## 5	0	0	0	0	0	1
## 6	0	0	0	0	1	0
## CANCER_HST_F LUNG_CANCER_HST_F ASTHMA_HST_F GERD_HST_F						
## 1	1	0	0	0	0	
## 2	0	0	0	0	0	
## 3	0	0	1	0	0	
## 4	0	0	0	0	0	
## 5	0	0	1	1	1	
## 6	1	0	0	0	1	
## FIBROMYALGIA_HST_F DEPRESSION_HST_F PSORIATIC_ARTHRITIS_HST_F						
## 1	0	0	0	0	0	
## 2	0	0	0	0	0	
## 3	0	0	0	0	0	
## 4	0	0	0	0	0	
## 5	0	0	0	0	0	
## 6	0	0	0	0	0	
## RHEUM_ARTHRITIS_HST_F LUPUS_HST_F VTVF_HST_F STROKE_HST_F						
## 1	0	0	0	0	0	
## 2	0	0	0	0	0	
## 3	0	0	0	0	0	
## 4	0	0	0	0	0	
## 5	0	0	0	0	0	
## 6	0	1	0	0	0	
## VASCULARDISEASE_HST_F LOWBACKPAIN_HST_F DVT_HST_F PE_HST_F						
## 1	0	0	0	0	0	
## 2	0	0	0	0	0	
## 3	0	0	0	0	0	
## 4	0	0	0	0	0	
## 5	0	1	0	0	0	
## 6	0	1	0	0	0	
## HYPOTHYROIDISM_HST_F ADRENAL_INSUFFICIENCY_HST_F INFERTILITY_HST_F						
## 1	0	0	0	0	0	
## 2	0	0	0	0	0	
## 3	0	0	0	0	0	
## 4	0	0	0	0	0	
## 5	0	0	0	0	0	
## 6	0	0	0	0	0	
## OBS_SLEEPAPNEA_HST_F CARDIAC_ARREST_HST_F HEMO_STROKE_HST_F						
## 1	0	0	0	0	0	
## 2	0	0	0	0	0	
## 3	0	0	0	0	0	
## 4	0	0	0	0	0	
## 5	0	0	0	0	0	
## 6	0	0	0	0	0	
## MAJOR_BLEED_HST_F MACULAR_DEGEN_HST_F ANXIETY_HST_F HYPERLIPIDEMIA_HST_F						
## 1	0	0	0	0	1	
## 2	0	0	0	0	0	
## 3	0	0	0	0	0	
## 4	0	0	0	0	0	
## 5	0	0	0	0	1	
## 6	0	0	0	0	0	
## HIV_HST_F METFORMIN_FLAG OPIOIDS_FLAG ALPHA_BLOCKERS CENTRAL_ANTAGONISTS						
## 1	0	0	1	0	0	

## 2	0	1	0	0	0
## 3	1	0	0	0	0
## 4	0	0	1	0	0
## 5	0	0	1	0	0
## 6	0	0	1	0	0
## RENIN BETA_BLOCKERS ACE_INHIB ARB ALDOSTERONE_BLOCKERS VASODIALATORS					
## 1	0	0	0	0	0
## 2	0	1	1	0	0
## 3	0	0	0	0	0
## 4	0	1	0	0	0
## 5	0	0	0	1	0
## 6	0	0	0	0	0
## DIURETICS CALCIUM_BLOCKERS STATINS INSULIN_MEDS ASPIRIN WARFARIN DOACS					
## 1	1	1	0	0	0
## 2	1	0	1	0	1
## 3	1	0	0	0	0
## 4	1	0	1	0	0
## 5	1	0	1	0	0
## 6	0	0	0	0	0
## MEDROL PREDNISONE INHALED_STEROID_WITH_LABA INHALED_STEROID_WITHOUT_LABA					
## 1	0	0	0	0	0
## 2	0	0	0	0	0
## 3	0	0	0	0	0
## 4	0	0	0	0	0
## 5	0	0	0	0	0
## 6	0	0	0	0	0
## ASTHMA_BIOLOGICS SHORT_ACTING_BRONCHO_DIALATORS TNF_INHIBITORS					
## 1	0	0	0	0	0
## 2	0	0	0	0	0
## 3	0	0	0	0	0
## 4	0	0	0	0	0
## 5	0	0	0	0	0
## 6	0	0	1	0	0
## IMMUNOMODULATORS AMINOSALICYLATES CORTICOSTEROIDS ARNI ALLOPURINOL					
## 1	0	0	0	0	0
## 2	0	0	0	0	0
## 3	0	0	0	0	0
## 4	0	0	0	0	0
## 5	0	0	0	0	0
## 6	1	0	0	0	0
## NSAIDS ANTIDEPRESSANTS SEIZURE MUSCLERELAXANT DIGOXIN INOTROPES					
## 1	0	0	0	0	0
## 2	0	0	0	0	0
## 3	0	0	0	0	0
## 4	0	0	0	0	0
## 5	1	0	0	0	0
## 6	0	1	0	0	0
## ANTI_ARRHYTHMIC ANTIPLATELET SULFONYLUREA GLP_1AGONIST					
## 1	0	0	0	0	0
## 2	0	1	1	0	0
## 3	0	0	0	0	0
## 4	0	0	0	0	0
## 5	0	0	0	0	0
## 6	0	0	0	0	0

```

## THIAZOLIDINEDIONE SGLT2_INHIBITOR DPP4_INHIBITOR
## 1          0          0          0
## 2          0          0          0
## 3          0          0          0
## 4          0          0          0
## 5          0          0          0
## 6          0          0          0
## ALPHA_GLUCOSIDASE_INHIBITOR AMYLINOMIMETIC RAPID_ACTING_INSULIN
## 1          0          0          0
## 2          0          0          0
## 3          0          0          0
## 4          0          0          0
## 5          0          0          0
## 6          0          0          0
## SHORT_ACTING_INSULIN INTERMEDIATE_ACTING_INSULIN LONG_ACTING_INSULIN
## 1          0          0          0
## 2          0          0          0
## 3          0          0          0
## 4          0          0          0
## 5          0          0          0
## 6          0          0          0
## CA_CLOSEST_F MINOCYCLINE DOXYCYCLINE MELATONIN CKD_HST FOLLOW_UP_CKD_HST
## 1    9.0      0      0      0      1      1
## 2    8.3      0      0      0      0      0
## 3    8.7      0      0      0      0      0
## 4    8.5      0      0      0      0      0
## 5    9.2      0      0      0      0      0
## 6    8.8      0      0      0      0      0

```

Delete the rows that have NA data.

```

labdata <- na.omit(labdata)
head(labdata)

```

```

## ESRD_FLG_COMP AGE_ON_CONTACT_DATE   BMI BP_SYSTOLIC BP_DIASTOLIC
## 2          0      72.1 32.54      192      105
## 7          0      76.4 25.68      138       62
## 9          0      72.2 31.19      148       78
## 10         0      77.3 20.85      140       80
## 11         0      65.8 30.17      140       70
## 12         0      75.0 16.83      130       70
## EGFR_CLOSEST GLUCOSE_CLOSEST_F SODIUM_CLOSEST_F K_CLOSEST_F
## 2        53      175      138      4.2
## 7        53      129      139      4.6
## 9        37      113      143      4.0
## 10       43      270      129      4.3
## 11       41      164      141      4.6
## 12       46      96       135      4.1
## BILIRUBIN_CLOSEST_F ALT_CLOSEST_F AST_CLOSEST_F FOLLOW_UP_EGFR_VALUE
## 2        0.9      29       19      52
## 7        0.6      15       22      45
## 9        0.4      12       13      37
## 10       0.4      34       17      37
## 11       0.5      33       15      50
## 12       0.9      8        15      44

```

```

##      FEMALE RACE_F OBESE_HST_F MORBIDOBSESE_HST_F PH_HST_F COPD_HST_F
## 2      0     1       1           0       0       0
## 7      0     1       0           0       0       0
## 9      0     1       1           0       0       1
## 10     1     1       0           0       0       0
## 11     1     1       1           0       0       0
## 12     1     1       0           0       0       1
##      CHF_HST_F AFIB_HST_F DIAB_HST_F CAD_HST_F OSTEO_HST_F HTN_HST_F
## 2      0     0       0           0       0       0
## 7      0     0       0           0       0       1
## 9      1     0       1           0       0       1
## 10     0     0       0           0       0       1
## 11     0     0       1           0       0       1
## 12     0     0       0           0       1       1
##      CANCER_HST_F LUNG_CANCER_HST_F ASTHMA_HST_F GERD_HST_F
## 2      0             0           0       0
## 7      1             0           0       0
## 9      0             0           1       0
## 10     1             0           0       0
## 11     0             0           0       1
## 12     0             0           1       0
##      FIBROMYALGIA_HST_F DEPRESSION_HST_F PSORIATIC_ARTHRITIS_HST_F
## 2      0             0           0       0
## 7      0             0           0       0
## 9      0             0           0       0
## 10     0             0           0       0
## 11     0             0           0       0
## 12     0             0           0       0
##      RHEUM_ARTHRITIS_HST_F LUPUS_HST_F VTVF_HST_F STROKE_HST_F
## 2      0             0           0       0
## 7      0             0           0       0
## 9      0             0           0       0
## 10     0             0           0       0
## 11     0             0           0       0
## 12     0             0           0       0
##      VASCULARDISEASE_HST_F LOWBACKPAIN_HST_F DVT_HST_F PE_HST_F
## 2      0             0           0       0
## 7      0             0           0       0
## 9      0             0           0       0
## 10     0             0           0       0
## 11     0             0           1       0
## 12     0             0           0       0
##      HYPOTHYROIDISM_HST_F ADRENAL_INSUFFICIENCY_HST_F INFERTILITY_HST_F
## 2      0             0           0       0
## 7      0             0           0       0
## 9      0             0           0       0
## 10     0             0           0       0
## 11     1             0           0       0
## 12     1             0           0       0
##      OBS_SLEEPAPNEA_HST_F CARDIAC_ARREST_HST_F HEMO_STROKE_HST_F
## 2      0             0           0       0
## 7      0             0           0       0
## 9      0             0           0       0
## 10     0             0           0       0

```

## 11	0	0	0		
## 12	0	0	0		
## MAJOR_BLEED_HST_F	MACULAR_DEGEN_HST_F	ANXIETY_HST_F			
## 2	0	0	0		
## 7	0	0	0		
## 9	0	0	0		
## 10	0	0	1		
## 11	0	1	1		
## 12	0	0	0		
## HYPERLIPIDEMIA_HST_F	HIV_HST_F	METFORMIN_FLAG	OPIOIDS_FLAG		
## 2	0	0	1		
## 7	1	0	0		
## 9	1	0	0		
## 10	1	0	0		
## 11	1	0	1		
## 12	1	0	0		
## ALPHA_BLOCKERS	CENTRAL_ANTAGONISTS	RENIN	BETA_BLOCKERS	ACE_INHIB	ARB
## 2	0	0	1	1	0
## 7	0	0	1	1	1
## 9	0	0	1	0	0
## 10	0	0	0	1	0
## 11	1	0	1	1	0
## 12	0	0	1	1	0
## ALDOSTERONE_BLOCKERS	VASODIALATORS	DIURETICS	CALCIUM_BLOCKERS	STATINS	
## 2	0	1	1	0	1
## 7	0	0	1	1	1
## 9	0	0	1	1	1
## 10	0	0	1	0	1
## 11	0	0	1	1	1
## 12	0	0	0	1	0
## INSULIN_MEDS	ASPIRIN	WARFARIN	DOACS	MEDROL	PREDNISONE
## 2	0	1	0	0	0
## 7	0	0	0	0	0
## 9	0	1	0	0	0
## 10	0	1	0	0	0
## 11	0	0	0	0	0
## 12	0	0	0	0	0
## INHALED_STEROID_WITH_LABA	INHALED_STEROID_WITHOUT_LABA	ASTHMA_BIOLOGICS			
## 2	0	0	0	0	0
## 7	0	0	0	0	0
## 9	0	0	0	0	0
## 10	0	0	0	0	0
## 11	0	0	0	0	0
## 12	0	0	0	0	0
## SHORT_ACTING_BRONCHO_DIALATORS	TNF_INHIBITORS	IMMUNOMODULATORS			
## 2	0	0	0	0	0
## 7	0	0	0	0	0
## 9	1	0	0	0	0
## 10	0	0	0	0	0
## 11	0	0	0	0	0
## 12	1	0	0	0	0
## AMINOSALICYLATES	CORTICOSTEROIDS	ARNI	ALLOPURINOL	NSAIDS	
## 2	0	0	0	0	0
## 7	0	0	0	0	1

## 9	0	0	1	0	
## 10	0	0	0	0	
## 11	0	0	0	0	
## 12	0	0	0	0	
##	ANTIDEPRESSANTS	SEIZURE	MUSCLERELAXANT	DIGOXIN	INOTROPES
## 2	0	0	0	0	0
## 7	0	0	0	0	0
## 9	0	0	0	0	0
## 10	0	0	0	0	0
## 11	0	0	0	0	0
## 12	1	0	0	0	0
##	ANTI_ARRHYTHMIC	ANTIPLATELET	SULFONYLUREA	GLP_1AGONIST	
## 2	0	1	1	0	
## 7	0	0	0	0	
## 9	0	0	0	0	
## 10	0	0	0	0	
## 11	0	0	1	0	
## 12	0	1	0	0	
##	THIAZOLIDINEDIONE	SGLT2_INHIBITOR	DPP4_INHIBITOR		
## 2	0	0	0		
## 7	0	0	0		
## 9	0	0	0		
## 10	0	0	0		
## 11	0	0	0		
## 12	0	0	0		
##	ALPHA_GLUCOSIDASE_INHIBITOR	AMYLINOMIMETIC	RAPID_ACTING_INSULIN		
## 2	0	0	0		
## 7	0	0	0		
## 9	0	0	0		
## 10	0	0	0		
## 11	0	0	0		
## 12	0	0	0		
##	SHORT_ACTING_INSULIN	INTERMEDIATE_ACTING_INSULIN	LONG_ACTING_INSULIN		
## 2	0	0	0		
## 7	0	0	0		
## 9	0	0	0		
## 10	0	0	0		
## 11	0	0	0		
## 12	0	0	0		
##	CA_CLOSEST_F	MINOCYCLINE	DOXYCYCLINE	MELATONIN	CKD_HST
## 2	8.3	0	0	0	0
## 7	9.8	0	0	0	0
## 9	9.5	0	0	0	1
## 10	9.4	0	0	0	0
## 11	9.3	0	0	0	0
## 12	9.5	0	0	0	0
##	FOLLOW_UP_CKD_HST				
## 2	0				
## 7	1				
## 9	1				
## 10	0				
## 11	0				
## 12	0				

```
summary(labdata)
```

```
## ESRD_FLG_COMP AGE_ON_CONTACT_DATE BMI BP_SYSTOLIC
## Min. :0.00000 Min. :18.60 Min. :13.48 Min. : 68.0
## 1st Qu.:0.00000 1st Qu.:62.50 1st Qu.:26.46 1st Qu.:120.0
## Median :0.00000 Median :70.20 Median :30.42 Median :130.0
## Mean : 0.02377 Mean :68.77 Mean :31.50 Mean :130.6
## 3rd Qu.:0.00000 3rd Qu.:76.90 3rd Qu.:35.43 3rd Qu.:140.0
## Max. :1.00000 Max. :84.40 Max. :88.57 Max. :270.0
## BP_DIASTOLIC EGFR_CLOSEST GLUCOSE_CLOSEST_F SODIUM_CLOSEST_F
## Min. : 36.00 Min. :15.00 Min. : 0.0 Min. : 3.9
## 1st Qu.: 70.00 1st Qu.:43.00 1st Qu.: 93.0 1st Qu.:138.0
## Median : 76.00 Median :50.00 Median :105.0 Median :140.0
## Mean : 75.18 Mean :47.77 Mean :122.4 Mean :139.5
## 3rd Qu.: 81.00 3rd Qu.:55.00 3rd Qu.:132.0 3rd Qu.:141.0
## Max. :140.00 Max. :58.90 Max. :738.0 Max. :151.0
## K_CLOSEST_F BILIRUBIN_CLOSEST_F ALT_CLOSEST_F AST_CLOSEST_F
## Min. :2.40 Min. : 0.1000 Min. : 1.00 Min. : 3.00
## 1st Qu.:4.00 1st Qu.: 0.4000 1st Qu.: 15.00 1st Qu.: 16.00
## Median :4.30 Median : 0.5000 Median : 21.00 Median : 20.00
## Mean : 4.32 Mean : 0.5935 Mean : 25.23 Mean : 23.14
## 3rd Qu.:4.60 3rd Qu.: 0.7000 3rd Qu.: 30.00 3rd Qu.: 26.00
## Max. : 6.90 Max. :39.8000 Max. :1045.00 Max. :1132.00
## FOLLOW_UP_EGFR_VALUE FEMALE RACE_F OBESE_HST_F
## Min. :15.00 Min. :0.000 Min. :1.000 Min. :0.0000
## 1st Qu.:43.00 1st Qu.:0.000 1st Qu.:1.000 1st Qu.:0.0000
## Median :50.00 Median :1.000 Median :1.000 Median :1.0000
## Mean : 47.48 Mean : 0.598 Mean : 1.128 Mean : 0.6286
## 3rd Qu.:54.00 3rd Qu.:1.000 3rd Qu.:1.000 3rd Qu.:1.0000
## Max. : 58.90 Max. :1.000 Max. :3.000 Max. :1.0000
## MORBIDOBEBE_HST_F PH_HST_F COPD_HST_F CHF_HST_F
## Min. :0.0000 Min. :0.00000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :0.0000 Median :0.00000 Median :0.0000 Median :0.0000
## Mean : 0.1719 Mean : 0.03622 Mean : 0.1508 Mean : 0.1613
## 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.0000
## Max. : 1.0000 Max. :1.00000 Max. :1.0000 Max. :1.0000
## AFIB_HST_F DIAB_HST_F CAD_HST_F OSTEO_HST_F
## Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :0.0000 Median :0.0000 Median :0.0000 Median :0.0000
## Mean : 0.1435 Mean : 0.4441 Mean : 0.2788 Mean : 0.1203
## 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.0000
## Max. : 1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000
## HTN_HST_F CANCER_HST_F LUNG_CANCER_HST_F ASTHMA_HST_F
## Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:1.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000
## Median :1.0000 Median :0.0000 Median :0.0000 Median :0.0000
## Mean : 0.7802 Mean : 0.2742 Mean : 0.0127 Mean : 0.2056
## 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:0.0000
## Max. : 1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000
## G俞D_HST_F FIBROMYALGIA_HST_F DEPRESSION_HST_F
## Min. :0.0000 Min. :0.00000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.0000
```

```

## Median :0.0000  Median :0.00000  Median :0.0000
## Mean   :0.3099  Mean   :0.02388  Mean   :0.1846
## 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:0.0000
## Max.   :1.0000  Max.   :1.00000  Max.   :1.0000
## PSORIATIC_ARTHRITIS_HST_F RHEUM_ARTHRITIS_HST_F LUPUS_HST_F
## Min.   :0.00000      Min.   :0.00000      Min.   :0.0000
## 1st Qu.:0.00000      1st Qu.:0.00000      1st Qu.:0.0000
## Median :0.00000      Median :0.00000      Median :0.0000
## Mean   :0.02536      Mean   :0.03679      Mean   :0.0117
## 3rd Qu.:0.00000      3rd Qu.:0.00000      3rd Qu.:0.0000
## Max.   :1.00000      Max.   :1.00000      Max.   :1.0000
## VTVF_HST_F           STROKE_HST_F       VASCULARDISEASE_HST_F
## Min.   :0.00000      Min.   :0.00000      Min.   :0.00000
## 1st Qu.:0.00000      1st Qu.:0.00000      1st Qu.:0.00000
## Median :0.00000      Median :0.00000      Median :0.00000
## Mean   :0.01935      Mean   :0.1183       Mean   :0.04897
## 3rd Qu.:0.00000      3rd Qu.:0.00000      3rd Qu.:0.00000
## Max.   :1.00000      Max.   :1.00000      Max.   :1.00000
## LOWBACKPAIN_HST_F    DVT_HST_F         PE_HST_F
## Min.   :0.00000      Min.   :0.00000      Min.   :0.00000
## 1st Qu.:0.00000      1st Qu.:0.00000      1st Qu.:0.00000
## Median :0.00000      Median :0.00000      Median :0.00000
## Mean   :0.3139       Mean   :0.03305      Mean   :0.02826
## 3rd Qu.:1.00000      3rd Qu.:0.00000      3rd Qu.:0.00000
## Max.   :1.00000      Max.   :1.00000      Max.   :1.00000
## HYPOTHYROIDISM_HST_F ADRENAL_INSUFFICIENCY_HST_F INFERTILITY_HST_F
## Min.   :0.00000      Min.   :0.00000      Min.   :0.0000000
## 1st Qu.:0.00000      1st Qu.:0.00000      1st Qu.:0.0000000
## Median :0.00000      Median :0.00000      Median :0.0000000
## Mean   :0.2368       Mean   :0.01819      Mean   :0.003954
## 3rd Qu.:0.00000      3rd Qu.:0.00000      3rd Qu.:0.0000000
## Max.   :1.00000      Max.   :1.00000      Max.   :1.0000000
## OBS_SLEEPAPNEA_HST_F CARDIAC_ARREST_HST_F HEMO_STROKE_HST_F
## Min.   :0.00000      Min.   :0.00000      Min.   :0.0000000
## 1st Qu.:0.00000      1st Qu.:0.00000      1st Qu.:0.0000000
## Median :0.00000      Median :0.00000      Median :0.0000000
## Mean   :0.09562      Mean   :0.002161     Mean   :0.008487
## 3rd Qu.:0.00000      3rd Qu.:0.00000      3rd Qu.:0.0000000
## Max.   :1.00000      Max.   :1.00000      Max.   :1.0000000
## MAJOR_BLEED_HST_F   MACULAR_DEGEN_HST_F ANXIETY_HST_F
## Min.   :0.00000      Min.   :0.00000      Min.   :0.00000
## 1st Qu.:0.00000      1st Qu.:0.00000      1st Qu.:0.00000
## Median :0.00000      Median :0.00000      Median :0.00000
## Mean   :0.1671       Mean   :0.01871      Mean   :0.1622
## 3rd Qu.:0.00000      3rd Qu.:0.00000      3rd Qu.:0.00000
## Max.   :1.00000      Max.   :1.00000      Max.   :1.00000
## HYPERLIPIDEMIA_HST_F HIV_HST_F          METFORMIN_FLAG   OPIOIDS_FLAG
## Min.   :0.00000      Min.   :0.0000000  Min.   :0.0000  Min.   :0.0000
## 1st Qu.:0.00000      1st Qu.:0.0000000  1st Qu.:0.0000  1st Qu.:0.0000
## Median :1.00000      Median :0.0000000  Median :0.0000  Median :0.0000
## Mean   :0.7256       Mean   :0.007117   Mean   :0.2064  Mean   :0.2881
## 3rd Qu.:1.00000      3rd Qu.:0.0000000  3rd Qu.:0.0000  3rd Qu.:1.0000
## Max.   :1.00000      Max.   :1.0000000  Max.   :1.0000  Max.   :1.0000
## ALPHA_BLOCKERS      CENTRAL_ANTAGONISTS  RENIN        BETA_BLOCKERS

```

```

## Min. :0.00000 Min. :0.0000000 Min. :0.0000000 Min. :0.00000
## 1st Qu.:0.00000 1st Qu.:0.0000000 1st Qu.:0.0000000 1st Qu.:0.00000
## Median :0.00000 Median :0.0000000 Median :0.0000000 Median :1.00000
## Mean :0.03421 Mean :0.0006853 Mean :0.001634 Mean :0.5049
## 3rd Qu.:0.00000 3rd Qu.:0.0000000 3rd Qu.:0.0000000 3rd Qu.:1.00000
## Max. :1.00000 Max. :1.0000000 Max. :1.0000000 Max. :1.00000
## ACE_INHIB ARB ALDOSTERONE_BLOCKERS VASODIALATORS
## Min. :0.0000 Min. :0.0000 Min. :0.00000 Min. :0.000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.000
## Median :0.0000 Median :0.0000 Median :0.00000 Median :0.000
## Mean :0.3577 Mean :0.2275 Mean :0.05936 Mean :0.033
## 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.000
## Max. :1.0000 Max. :1.0000 Max. :1.00000 Max. :1.000
## DIURETICS CALCIUM_BLOCKERS STATINS INSULIN_MEDS
## Min. :0.0000 Min. :0.0000 Min. :0.000 Min. :0.0000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.0000
## Median :1.0000 Median :0.0000 Median :1.000 Median :0.0000
## Mean :0.5045 Mean :0.3095 Mean :0.605 Mean :0.1906
## 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.000 3rd Qu.:0.0000
## Max. :1.0000 Max. :1.0000 Max. :1.000 Max. :1.0000
## ASPIRIN WARFARIN DOACS MEDROL
## Min. :0.0000 Min. :0.0000 Min. :0.00000 Min. :0.00000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000
## Median :1.0000 Median :0.0000 Median :0.00000 Median :0.00000
## Mean :0.5532 Mean :0.1092 Mean :0.04222 Mean :0.01782
## 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.00000
## Max. :1.0000 Max. :1.0000 Max. :1.00000 Max. :1.00000
## PREDNISONE INHALED_STEROID_WITH_LABA INHALED_STEROID_WITHOUT_LABA
## Min. :0.00000 Min. :0.00000 Min. :0.00000
## 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000
## Median :0.00000 Median :0.00000 Median :0.00000
## Mean :0.08714 Mean :0.09837 Mean :0.03711
## 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000
## Max. :1.00000 Max. :1.00000 Max. :1.00000
## ASTHMA_BIOLOGICS SHORT_ACTING_BRONCHO_DIALATORS TNF_INHIBITORS
## Min. :0.0000000 Min. :0.00000 Min. :0.0000000
## 1st Qu.:0.0000000 1st Qu.:0.00000 1st Qu.:0.0000000
## Median :0.0000000 Median :0.00000 Median :0.0000000
## Mean :0.0004217 Mean :0.1795 Mean :0.005008
## 3rd Qu.:0.0000000 3rd Qu.:0.00000 3rd Qu.:0.0000000
## Max. :1.0000000 Max. :1.00000 Max. :1.0000000
## IMMUNOMODULATORS AMINOSALICYLATES CORTICOSTEROIDS ARNI
## Min. :0.00000 Min. :0.0000000 Min. :0.00000 Min. :0.0000000
## 1st Qu.:0.00000 1st Qu.:0.0000000 1st Qu.:0.00000 1st Qu.:0.0000000
## Median :0.00000 Median :0.0000000 Median :0.00000 Median :0.0000000
## Mean :0.02778 Mean :0.009858 Mean :0.1625 Mean :0.0006326
## 3rd Qu.:0.00000 3rd Qu.:0.0000000 3rd Qu.:0.00000 3rd Qu.:0.0000000
## Max. :1.00000 Max. :1.0000000 Max. :1.00000 Max. :1.0000000
## ALLOPURINOL NSAIDS ANTIDEPRESSANTS SEIZURE
## Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.00000
## 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000
## Median :0.00000 Median :0.00000 Median :0.00000 Median :0.00000
## Mean :0.07206 Mean :0.1188 Mean :0.05509 Mean :0.01017
## 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000

```

```

## Max.    :1.00000  Max.    :1.00000  Max.    :1.00000  Max.    :1.00000
## MUSCLERELAXANT      DIGOXIN        INOTROPES      ANTI_ARRHYTHMIC
## Min.    :0.00000  Min.    :0.00000  Min.    :0.0000000  Min.    :0.00000
## 1st Qu.:0.00000  1st Qu.:0.00000  1st Qu.:0.0000000  1st Qu.:0.00000
## Median  :0.00000  Median :0.00000  Median :0.0000000  Median :0.00000
## Mean    :0.01402  Mean   :0.03405  Mean   :0.0005799  Mean   :0.04391
## 3rd Qu.:0.00000  3rd Qu.:0.00000  3rd Qu.:0.0000000  3rd Qu.:0.00000
## Max.    :1.00000  Max.    :1.00000  Max.    :1.0000000  Max.    :1.00000
## ANTIPLATELET      SULFONYLUREA    GLP_1AGONIST   THIAZOLIDINEDIONE
## Min.    :0.00000  Min.    :0.00000  Min.    :0.00000  Min.    :0.00000
## 1st Qu.:0.00000  1st Qu.:0.00000  1st Qu.:0.00000  1st Qu.:0.00000
## Median  :0.00000  Median :0.00000  Median :0.00000  Median :0.00000
## Mean    :0.1147   Mean   :0.1517   Mean   :0.0156   Mean   :0.02098
## 3rd Qu.:0.00000  3rd Qu.:0.00000  3rd Qu.:0.00000  3rd Qu.:0.00000
## Max.    :1.00000  Max.    :1.00000  Max.    :1.00000  Max.    :1.00000
## SGLT2_INHIBITOR    DPP4_INHIBITOR  ALPHA_GLUCOSIDASE_INHIBITOR
## Min.    :0.000000  Min.    :0.000000  Min.    :0.0000000
## 1st Qu.:0.000000  1st Qu.:0.000000  1st Qu.:0.0000000
## Median  :0.000000  Median :0.000000  Median :0.0000000
## Mean    :0.006326  Mean   :0.06589  Mean   :0.002109
## 3rd Qu.:0.000000  3rd Qu.:0.000000  3rd Qu.:0.0000000
## Max.    :1.000000  Max.    :1.000000  Max.    :1.0000000
## AMYLINOMIMETIC    RAPID_ACTING_INSULIN SHORT_ACTING_INSULIN
## Min.    :0.000000  Min.    :0.00000  Min.    :0.00000
## 1st Qu.:0.000000  1st Qu.:0.00000  1st Qu.:0.00000
## Median  :0.000000  Median :0.00000  Median :0.00000
## Mean    :0.001054  Mean   :0.1166   Mean   :0.02393
## 3rd Qu.:0.000000  3rd Qu.:0.00000  3rd Qu.:0.00000
## Max.    :1.000000  Max.    :1.00000  Max.    :1.00000
## INTERMEDIATE_ACTING_INSULIN LONG_ACTING_INSULIN CA_CLOSEST_F
## Min.    :0.00000  Min.    :0.00000  Min.    :5.500
## 1st Qu.:0.00000  1st Qu.:0.00000  1st Qu.:9.000
## Median  :0.00000  Median :0.00000  Median :9.300
## Mean    :0.02741  Mean   :0.1337   Mean   :9.326
## 3rd Qu.:0.00000  3rd Qu.:0.00000  3rd Qu.:9.700
## Max.    :1.00000  Max.    :1.00000  Max.    :37.000
## MINOCYCLINE       DOXYCYCLINE     MELATONIN      CKD_HST
## Min.    :0.000000  Min.    :0.00000  Min.    :0.00000  Min.    :0.00
## 1st Qu.:0.000000  1st Qu.:0.00000  1st Qu.:0.00000  1st Qu.:0.00
## Median  :0.000000  Median :0.00000  Median :0.00000  Median :0.00
## Mean    :0.002478  Mean   :0.01492  Mean   :0.01724  Mean   :0.13
## 3rd Qu.:0.000000  3rd Qu.:0.00000  3rd Qu.:0.00000  3rd Qu.:0.00
## Max.    :1.000000  Max.    :1.00000  Max.    :1.00000  Max.    :1.00
## FOLLOW_UP_CKD_HST
## Min.    :0.0000
## 1st Qu.:0.0000
## Median :0.0000
## Mean    :0.1943
## 3rd Qu.:0.0000
## Max.    :1.0000

```

Split the data into train set and test set.

```
set.seed(1)
labdata_train_test_split <- initial_split(labdata, prop = 0.75)
train_tbl <- training(labdata_train_test_split)
test_tbl <- testing(labdata_train_test_split)
```

Use logistic regression model to exclude the features that not associated with the target($P > 0.1$).

```
model_glm <- glm(ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl)

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
summary(model_glm)

##
## Call:
## glm(formula = ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl)
##
## Deviance Residuals:
##      Min        1Q     Median        3Q       Max
## -2.5436  -0.1196  -0.0633  -0.0370   4.0239
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)           1.128e+01  3.556e+00  3.173  0.001507 **
## AGE_ON_CONTACT_DATE -5.073e-02  6.590e-03 -7.698 1.38e-14 ***
## BMI                  -3.446e-02  1.695e-02 -2.033  0.042033 *
## BP_SYSTOLIC          7.138e-03  4.358e-03  1.638  0.101434
## BP_DIASTOLIC         -5.351e-03  7.609e-03 -0.703  0.481892
## EGFR_CLOSEST          -5.180e-02  1.447e-02 -3.580  0.000343 ***
## GLUCOSE_CLOSEST_F    -5.794e-04  1.142e-03 -0.508  0.611768
## SODIUM_CLOSEST_F     1.676e-02  2.370e-02  0.707  0.479329
## K_CLOSEST_F           -4.796e-02  1.304e-01 -0.368  0.712937
## BILIRUBIN_CLOSEST_F  8.540e-02  7.442e-02  1.148  0.251161
## ALT_CLOSEST_F         -1.186e-02  5.528e-03 -2.146  0.031909 *
## AST_CLOSEST_F          1.445e-02  5.050e-03  2.861  0.004221 **
## FOLLOW_UP_EGFR_VALUE -1.156e-01  1.430e-02 -8.088 6.07e-16 ***
## FEMALE                -7.693e-01  1.567e-01 -4.910 9.12e-07 ***
## RACE_F                 3.716e-01  1.623e-01  2.289  0.022093 *
## OBESE_HST_F            2.759e-01  1.936e-01  1.425  0.154210
## MORBIDOBESE_HST_F     6.662e-02  2.727e-01  0.244  0.806987
## PH_HST_F               -3.909e-01  3.737e-01 -1.046  0.295511
## COPD_HST_F             2.813e-02  2.318e-01  0.121  0.903421
## CHF_HST_F              9.309e-02  1.959e-01  0.475  0.634694
## AFIB_HST_F              -4.270e-02  2.491e-01 -0.171  0.863894
## DIAB_HST_F              3.086e-01  2.039e-01  1.513  0.130255
## CAD_HST_F               1.649e-03  1.801e-01  0.009  0.992695
## OSTEO_HST_F             5.771e-01  2.461e-01  2.345  0.019021 *
## HTN_HST_F                1.359e-01  1.885e-01  0.721  0.470957
## CANCER_HST_F             -6.265e-02  1.639e-01 -0.382  0.702359
## LUNG_CANCER_HST_F        -5.103e-01  8.410e-01 -0.607  0.543980
```

## ASTHMA_HST_F	-2.553e-01	2.212e-01	-1.154	0.248412
## GERD_HST_F	-2.045e-01	1.725e-01	-1.185	0.235842
## FIBROMYALGIA_HST_F	-7.805e-01	1.049e+00	-0.744	0.456668
## DEPRESSION_HST_F	-8.935e-02	2.107e-01	-0.424	0.671435
## PSORIATIC_ARTHRITIS_HST_F	7.662e-01	4.322e-01	1.773	0.076263 .
## RHEUM_ARTHRITIS_HST_F	1.783e-01	4.963e-01	0.359	0.719464
## LUPUS_HST_F	1.558e-01	5.479e-01	0.284	0.776154
## VTVF_HST_F	5.597e-01	4.252e-01	1.316	0.188112
## STROKE_HST_F	-1.101e-02	2.232e-01	-0.049	0.960670
## VASCULARDISEASE_HST_F	-2.147e-01	3.485e-01	-0.616	0.537847
## LOWBACKPAIN_HST_F	-2.879e-01	1.900e-01	-1.515	0.129708
## DVT_HST_F	2.873e-01	3.432e-01	0.837	0.402559
## PE_HST_F	-2.685e-01	4.719e-01	-0.569	0.569316
## HYPOTHYROIDISM_HST_F	2.666e-01	1.810e-01	1.473	0.140675
## ADRENAL_INSUFFICIENCY_HST_F	-4.520e-01	6.726e-01	-0.672	0.501501
## INFERTILITY_HST_F	-1.459e+01	6.875e+02	-0.021	0.983065
## OBS_SLEEPAPNEA_HST_F	4.943e-01	2.422e-01	2.041	0.041238 *
## CARDIAC_ARREST_HST_F	-1.985e+00	1.691e+00	-1.174	0.240374
## HEMO_STROKE_HST_F	-5.685e-01	8.140e-01	-0.698	0.484935
## MAJOR_BLEED_HST_F	-3.382e-01	2.191e-01	-1.543	0.122818
## MACULAR_DEGEN_HST_F	5.467e-01	4.845e-01	1.128	0.259136
## ANXIETY_HST_F	4.055e-01	2.256e-01	1.798	0.072181 .
## HYPERLIPIDEMIA_HST_F	6.559e-02	1.701e-01	0.386	0.699846
## HIV_HST_F	9.424e-01	5.766e-01	1.634	0.102206
## METFORMIN_FLAG	-8.430e-02	2.620e-01	-0.322	0.747664
## OPIOIDS_FLAG	1.616e-01	1.558e-01	1.037	0.299617
## ALPHA_BLOCKERS	-2.190e-01	3.405e-01	-0.643	0.520059
## CENTRAL_ANTAGONISTS	-1.488e+01	2.122e+03	-0.007	0.994404
## RENIN	-1.396e+01	1.106e+03	-0.013	0.989931
## BETA_BLOCKERS	-5.575e-02	1.471e-01	-0.379	0.704627
## ACE_INHIB	2.442e-02	1.591e-01	0.154	0.878000
## ARB	-3.409e-01	1.901e-01	-1.793	0.072906 .
## ALDOSTERONE_BLOCKERS	3.195e-01	2.677e-01	1.193	0.232677
## VASODIALATORS	1.501e-01	2.252e-01	0.667	0.505034
## DIURETICS	2.220e-01	1.536e-01	1.445	0.148354
## CALCIUM_BLOCKERS	1.434e-01	1.451e-01	0.988	0.323198
## STATINS	-9.853e-03	1.605e-01	-0.061	0.951055
## INSULIN_MEDS	5.742e-01	3.576e-01	1.606	0.108272
## ASPIRIN	-2.005e-01	1.507e-01	-1.330	0.183405
## WARFARIN	2.063e-01	2.335e-01	0.883	0.377118
## DOACS	2.855e-01	4.102e-01	0.696	0.486348
## MEDROL	-9.223e-01	8.272e-01	-1.115	0.264852
## PREDNISONE	-2.783e-01	3.843e-01	-0.724	0.468905
## INHALED_STEROID_WITH_LABA	-4.465e-01	3.628e-01	-1.231	0.218453
## INHALED_STEROID_WITHOUT_LABA	-1.799e-01	5.977e-01	-0.301	0.763361
## ASTHMA_BIOLOGICS	-1.178e+01	2.872e+03	-0.004	0.996726
## SHORT_ACTING_BRONCHO_DIALATORS	-2.951e-01	2.294e-01	-1.287	0.198204
## TNF_INHIBITORS	-3.518e-01	1.165e+00	-0.302	0.762584
## IMMUNOMODULATORS	-4.004e-01	5.210e-01	-0.769	0.442174
## AMINOSALICYLATES	-2.229e-01	8.825e-01	-0.253	0.800610
## CORTICOSTEROIDS	3.083e-01	3.436e-01	0.897	0.369579
## ARNI	-1.394e+01	1.667e+03	-0.008	0.993326
## ALLOPURINOL	5.788e-02	2.063e-01	0.281	0.779089
## NSAIDS	-4.358e-01	3.691e-01	-1.181	0.237743

```

## ANTIDEPRESSANTS          -4.098e-01  3.577e-01  -1.145  0.252003
## SEIZURE                  -1.427e+00  9.687e-01  -1.473  0.140838
## MUSCLERELAXANT           1.062e-01  6.878e-01   0.154  0.877277
## DIGOXIN                  -3.063e-01  4.265e-01  -0.718  0.472663
## INOTROPES                 -1.403e+01  3.047e+03  -0.005  0.996326
## ANTI_ARRHYTHMIC           -5.309e-01  3.752e-01  -1.415  0.157068
## ANTIPLATELET              1.034e-01  2.169e-01   0.477  0.633542
## SULFONYLUREA               2.999e-01  2.087e-01   1.437  0.150711
## GLP_1AGONIST               -2.150e-01  6.608e-01  -0.325  0.744923
## THIAZOLIDINEDIONE          -1.730e-01  5.080e-01  -0.340  0.733500
## SGLT2_INHIBITOR             1.309e+00  7.672e-01   1.707  0.087887 .
## DPP4_INHIBITOR              -4.008e-01  3.198e-01  -1.253  0.210130
## ALPHA_GLUCOSIDASE_INHIBITOR -7.932e-01  1.465e+00  -0.542  0.588115
## AMYLINOMIMETIC              -1.431e+01  1.783e+03  -0.008  0.993598
## RAPID_ACTING_INSULIN        1.770e-01  2.592e-01   0.683  0.494656
## SHORT_ACTING_INSULIN        4.715e-01  3.903e-01   1.208  0.227041
## INTERMEDIATE_ACTING_INSULIN -3.062e-01  3.868e-01  -0.791  0.428694
## LONG_ACTING_INSULIN         -9.421e-02  2.756e-01  -0.342  0.732449
## CA_CLOSEST_F                -8.033e-01  1.180e-01  -6.806  1.00e-11 ***
## MINOCYCLINE                 -1.562e+01  7.747e+02  -0.020  0.983910
## DOXYCYCLINE                  -6.660e-01  6.662e-01  -1.000  0.317444
## MELATONIN                     -3.099e-01  6.627e-01  -0.468  0.640074
## CKD_HST                      -4.866e-02  2.387e-01  -0.204  0.838484
## FOLLOW_UP_CKD_HST            -1.750e-01  2.236e-01  -0.782  0.433996
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 3218.4  on 14227  degrees of freedom
## Residual deviance: 1761.8  on 14123  degrees of freedom
## AIC: 1971.8
##
## Number of Fisher Scoring iterations: 17

```

So, we keep features “AGE_ON_CONTACT_DATE”, “BMI”, “EGFR CLOSEST”, “ALT CLOSEST_F”, “AST CLOSEST_F”, “FOLLOW_UP_EGFR_VALUE”, “FEMALE”, “RACE_F”, “OSTEO_HST_F”, “PSORIATIC_ARTHRITIS_HST_F”, “OBS_SLEEPAPNEA_HST_F”, “ANXIETY_HST_F”, “ARB”, “SGLT2_INHIBITOR” and “CA CLOSEST_F”.

```

train_tbl <- train_tbl[, c("ESRD_FLG_COMP", "AGE_ON_CONTACT_DATE", "BMI", "EGFR_CLOSEST", "ALT_CLOSEST_F",
                           "AST_CLOSEST_F", "FOLLOW_UP_EGFR_VALUE", "FEMALE", "RACE_F", "OSTEO_HST_F",
                           "PSORIATIC_ARTHRITIS_HST_F", "OBS_SLEEPAPNEA_HST_F", "ANXIETY_HST_F", "ARB",
                           "SGLT2_INHIBITOR")]

test_tbl <- test_tbl[, c("ESRD_FLG_COMP", "AGE_ON_CONTACT_DATE", "BMI", "EGFR_CLOSEST", "ALT_CLOSEST_F",
                        "AST_CLOSEST_F", "CA_CLOSEST_F", "FOLLOW_UP_EGFR_VALUE", "FEMALE", "RACE_F", "OSTEO_HST_F",
                        "PSORIATIC_ARTHRITIS_HST_F", "OBS_SLEEPAPNEA_HST_F", "ANXIETY_HST_F", "ARB",
                        "SGLT2_INHIBITOR")]

table(train_tbl$ESRD_FLG_COMP)

##
##      0      1
## 13887    341

```

We can separate the features into several category: (“EGFR CLOSEST”, “FOLLOW_UP_EGFR_VALUE”), (“AGE_ON_CONTACT_DATE”, “FEMALE”, “RACE_F”), (“BMI”), (“ALT CLOSEST_F”, “AST CLOSEST_F”, “CA CLOSEST_F”), (“OSTEO_HST_F”, “PSORIATIC_ARTHRITIS_HST_F”, “OBS_SLEEPAPNEA_HST_F”, “ANXIETY_HST_F”), (“ARB”, “SGLT2_INHIBITOR”). Now we can build 7 models for prediction. We use xgboost model to build.

```
model_table <- data.frame("model1" = c(0,0,NA), "model2" = c(0,0,NA), "model3" = c(0,0,NA), "model4" = c(0,0,NA))
row.names(model_table) <- c("AUC", "AIC", "P_value")
```

Build logistic regression model

Model 1: (“AGE_ON_CONTACT_DATE”, “FEMALE”, “RACE_F”)

Create dataset for Model 1

```
train_tbl_1 <- train_tbl[, c("ESRD_FLG_COMP", "AGE_ON_CONTACT_DATE", "FEMALE", "RACE_F")]
test_tbl_1 <- test_tbl[, c("ESRD_FLG_COMP", "AGE_ON_CONTACT_DATE", "FEMALE", "RACE_F")]
```

Create Model 1

```
model1 <- glm(ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl_1, x = TRUE)
```

Predict test data.

```
test_preds_1 <- test_tbl %>%
  mutate(probs = predict(model1, newdata = test_tbl_1, type = "response"))
```

Calculate AUC.

```
roc_1 <- roc(test_preds_1$ESRD_FLG_COMP, test_preds_1$probs)
model1_auc <- pROC::auc(roc_1)
ci.auc(roc_1)
```

95% CI: 0.5995–0.706 (DeLong)

```
model_table["AUC", "model1"] <- model1_auc
```

Calculate AIC.

```
model1_aic <- AIC(model1)
model_table["AIC", "model1"] <- model1_aic
```

Model 2: (“EGFR_CLOSEST”, “FOLLOW_UP_EGFR_VALUE”), (“AGE_ON_CONTACT_DATE”, “FEMALE”, “RACE_F”)

Create dataset for Model 2

```
train_tbl_2 <- train_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
test_tbl_2 <- test_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
```

Create Model 2

```
model2 <- glm(ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl_2, x = TRUE)
```

Predict test data.

```
test_preds_2 <- test_tbl %>%
  mutate(xgb_probs = predict(model2, newdata = test_tbl_2, type = "response"))
```

Calculate AUC.

```
roc_2 <- roc(test_preds_2$ESRD_FLG_COMP, test_preds_2$xgb_probs)
model2_auc <- pROC::auc(roc_2)
ci.auc(roc_2)
```

```
## 95% CI: 0.8862–0.951 (DeLong)
model_table["AUC", "model2"] <- model2_auc
```

Calculate AIC

```
model2_aic <- AIC(model2)
model_table["AIC", "model2"] <- model2_aic
```

Calculate p-value compared with model1

```
comp_2vs1 <- anova(model1,model2,test = "Chisq")
P_2vs1 <- comp_2vs1$`Pr(>Chi)`^
model_table["P_value", "model2"] <- P_2vs1[2]
```

Model 3: (“EGFR_CLOSEST”, “FOLLOW_UP_EGFR_VALUE”), (“AGE_ON_CONTACT_DATE”, “FEMALE”, “RACE_F”), (“BMI”)

Create dataset for model 3.

```
train_tbl_3 <- train_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
test_tbl_3 <- test_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
```

Create model 3.

```
model3 <- glm(ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl_3, x = TRUE)
```

Predict test data

```
test_preds_3 <- test_tbl %>%
  mutate(xgb_probs = predict(model3, newdata = test_tbl_3, type = "response"))
```

Calculate AUC

```
roc_3 <- roc(test_preds_3$ESRD_FLG_COMP, test_preds_3$xgb_probs)
model3_auc <- pROC::auc(roc_3)
ci.auc(roc_3)
```

```
## 95% CI: 0.8864–0.951 (DeLong)
```

```
model_table["AUC", "model3"] <- model3_auc
```

Calculate AIC

```
model3_aic <- AIC(model3)
model_table["AIC", "model3"] <- model3_aic
```

Calculate p-value compared with model2

```
comp_3vs2 <- anova(model3,model2,test = "Chisq")
P_3vs2 <- comp_3vs2$`Pr(>Chi)`^
model_table["P_value", "model3"] <- P_3vs2[2]
```

Model 4: (“EGFR_CLOSEST”, “FOLLOW_UP_EGFR_VALUE”), (“AGE_ON_CONTACT_DATE”, “FEMALE”, “RACE_F”), (“BMI”), (“ALT_CLOSEST_F”), (“AST_CLOSEST_F”), (“CA_CLOSEST_F”)

Create data set for model 4.

```

train_tbl_4 <- train_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
test_tbl_4 <- test_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]

Create model 4.

model4 <- glm(ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl_4, x = TRUE)

Predict test data.

test_preds_4 <- test_tbl %>%
  mutate(xgb_probs = predict(model4, newdata = test_tbl_4, type = "response"))

Calculate AUC

roc_4 <- roc(test_preds_4$ESRD_FLG_COMP, test_preds_4$xgb_probs)
model4_auc <- pROC::auc(roc_4)
ci.auc(roc_4)

## 95% CI: 0.9008-0.9551 (DeLong)
model_table["AUC", "model4"] <- model4_auc

Calculate AIC

model4_aic <- AIC(model4)
model_table["AIC", "model4"] <- model4_aic

Calculate p-value compared with model2

comp_4vs3 <- anova(model4, model3, test = "Chisq")
P_4vs3 <- comp_4vs3$`Pr(>Chi)```
model_table["P_value", "model4"] <- P_4vs3[2]

Model 5: ("EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE"), ("AGE_ON_CONTACT_DATE",
"SEX", "RACE_F"), ("BMI") ("OSTEO_HST_F", "PSORIATIC_ARTHRITIS_HST_F",
"OBS_SLEEPAPNEA_HST_F", "ANXIETY_HST_F")

Create data set for model 6.

train_tbl_5 <- train_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
test_tbl_5 <- test_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]

Create model 5.

model5 <- glm(ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl_5, x = TRUE)

Predict test data.

test_preds_5 <- test_tbl %>%
  mutate(xgb_probs = predict(model5, newdata = test_tbl_5, type = "response"))

Calculate AUC

roc_5 <- roc(test_preds_5$ESRD_FLG_COMP, test_preds_5$xgb_probs)
model5_auc <- pROC::auc(roc_5)
ci.auc(roc_5)

## 95% CI: 0.8809-0.9491 (DeLong)

```

```
model_table["AUC", "model5"] <- model5_auc
```

Calculate AIC

```
model5_aic <- AIC(model5)
model_table["AIC", "model5"] <- model5_aic
```

Calculate p-value compared with model2

```
comp_5vs3 <- anova(model5, model3, test = "Chisq")
P_5vs3 <- comp_5vs3$`Pr(>Chi)`^
model_table["P_value", "model5"] <- P_5vs3[2]
```

Model 6: (“EGFR_CLOSEST”, “FOLLOW_UP_EGFR_VALUE”), (“AGE_ON_CONTACT_DATE”, “FEMALE”, “RACE_F”), (“BMI”), (“ARB”, “SGLT2_INHIBITOR”)

Create data set for model 6.

```
train_tbl_6 <- train_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
test_tbl_6 <- test_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
```

Create model 6.

```
model6 <- glm(ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl_6, x = TRUE)
```

Predict test data.

```
test_preds_6 <- test_tbl %>%
  mutate(xgb_probs = predict(model6, newdata = test_tbl_6, type = "response"))
```

Calculate AUC

```
roc_6 <- roc(test_preds_6$ESRD_FLG_COMP, test_preds_6$xgb_probs)
model6_auc <- pROC::auc(roc_6)
ci.auc(roc_6)
```

95% CI: 0.8877-0.9509 (DeLong)

```
model_table["AUC", "model6"] <- model6_auc
```

Calculate AIC

```
model6_aic <- AIC(model6)
model_table["AIC", "model6"] <- model6_aic
```

Calculate p-value compared with model2

```
comp_6vs3 <- anova(model6, model3, test = "Chisq")
P_6vs3 <- comp_6vs3$`Pr(>Chi)`^
model_table["P_value", "model6"] <- P_6vs3[2]
```

Model 7: (“EGFR_CLOSEST”, “FOLLOW_UP_EGFR_VALUE”), (“AGE_ON_CONTACT_DATE”, “FEMALE”, “RACE_F”), (“BMI”), (“ALT_CLOSEST_F”, “AST_CLOSEST_F”, “CA_CLOSEST_F”), (“OSTEO_HST_F”, “PSORIATIC_ARTHRITIS_HST_F”, “OBS_SLEEPAPNEA_HST_F”, “ANXIETY_HST_F”), (“ARB”, “SGLT2_INHIBITOR”)

Create data set for model 7.

```

train_tbl_7 <- train_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]
test_tbl_7 <- test_tbl[, c("ESRD_FLG_COMP", "EGFR_CLOSEST", "FOLLOW_UP_EGFR_VALUE", "AGE_ON_CONTACT_DATE")]

Create model 7.

model7 <- glm(ESRD_FLG_COMP ~ ., family = binomial, data = train_tbl_7, x = TRUE)

Predict test data.

test_preds_7 <- test_tbl %>%
  mutate(xgb_probs = predict(model7, newdata = test_tbl_7, type = "response"))

Calculate AUC

roc_7 <- roc(test_preds_7$ESRD_FLG_COMP, test_preds_7$xgb_probs)
model7_auc <- pROC::auc(roc_7)
ci.auc(roc_7)

## 95% CI: 0.8953-0.9528 (DeLong)
model_table["AUC", "model7"] <- model7_auc

Calculate AIC

model7_aic <- AIC(model7)
model_table["AIC", "model7"] <- model7_aic

Calculate p-value compared with model2

comp_7vs3 <- anova(model7, model3, test = "Chisq")
P_7vs3 <- comp_7vs3$`Pr(>Chi)```
model_table["P_value", "model7"] <- P_7vs3[2]

```

Model performance

Validation table.

```

model_table

##           model1      model2      model3      model4      model5
## AUC      0.6527918 9.185626e-01 0.9187196 9.279891e-01 0.9149847
## AIC     3042.3378465 1.970376e+03 1972.2492384 1.913141e+03 1972.8452323
## P_value    NA 2.280095e-234 0.7213328 4.754894e-14 0.1160175
##           model6      model7
## AUC      0.9193044 9.240442e-01
## AIC     1971.7014721 1.911874e+03
## P_value   0.1029118 3.395295e-13

```

From the table we can tell Model 2, Model 4 and Model 7 are good model. Now use NRI to compare these three models.

At stage 3.

```

library(nricens)

## Loading required package: survival

```

```

##  

## Attaching package: 'survival'  

## The following object is masked from 'package:caret':  

##  

##     cluster  

## The following object is masked from 'package:rpart':  

##  

##     solder  

NRI3_4vs2 <- nribin(mdl.std = model2, mdl.new = model4, updown = 'category',cut = c(0.05,0.15), niter =  

##  

## UP and DOWN calculation:  

##    #of total, case, and control subjects at t0: 14228 341 13887  

##  

##    Reclassification Table for all subjects:  

##  

##          New  

## Standard < 0.05 < 0.15 >= 0.15  

##    < 0.05    12822    139      2  

##    < 0.15     161    472      59  

##    >= 0.15      1     76     496  

##  

##    Reclassification Table for case:  

##  

##          New  

## Standard < 0.05 < 0.15 >= 0.15  

##    < 0.05      77      6      1  

##    < 0.15       3     31      6  

##    >= 0.15      0      9    208  

##  

##    Reclassification Table for control:  

##  

##          New  

## Standard < 0.05 < 0.15 >= 0.15  

##    < 0.05    12745    133      1  

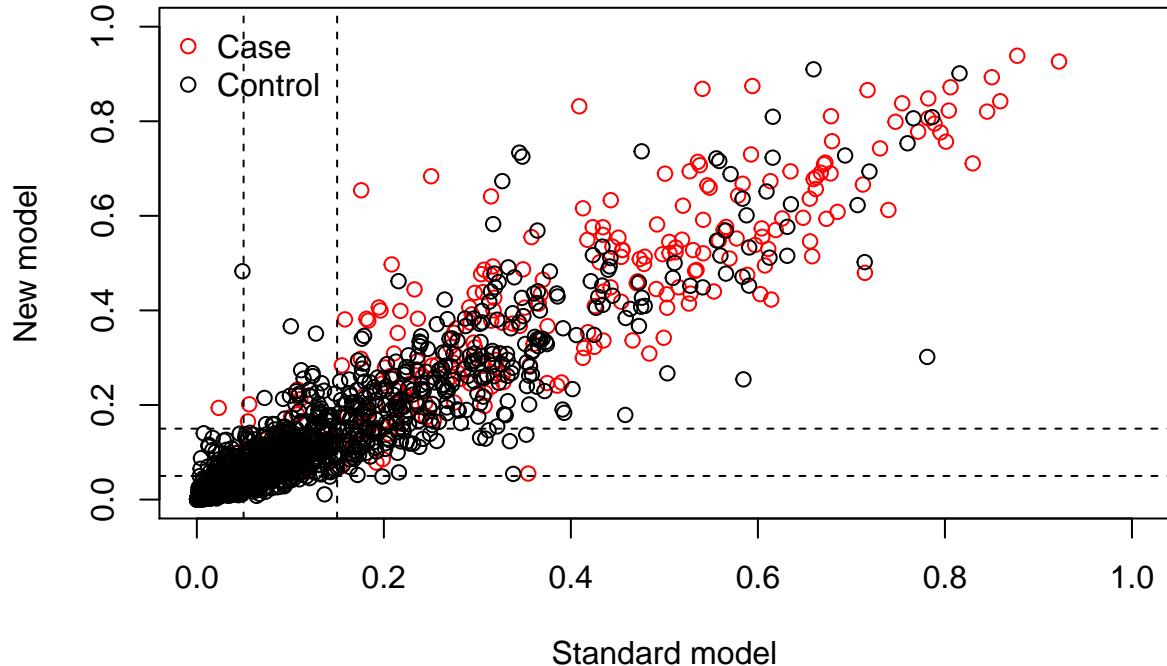
##    < 0.15     158    441      53  

##    >= 0.15      1     67     288  

##  

##    NRI estimation:  

## Point estimates:
```



```

##                               Estimate
## NRI                  0.005740933
## NRI+
## NRI-
## Pr(Up|Case)   0.038123167
## Pr(Down|Case) 0.035190616
## Pr(Down|Ctrl) 0.016274213
## Pr(Up|Ctrl)   0.013465831

NRI3_7vs2 <- nribin(mdl.std = model2, mdl.new = model7, updown = 'category',cut = c(0.05,0.15), niter = 1000)

##
## UP and DOWN calculation:
##    #of total, case, and control subjects at t0:  14228 341 13887

##
## Reclassification Table for all subjects:
##                               New
## Standard < 0.05 < 0.15 >= 0.15
##    < 0.05     12798     162      3
##    < 0.15      178     450     64
##    >= 0.15       3      79    491

##
## Reclassification Table for case:
##                               New
## Standard < 0.05 < 0.15 >= 0.15
##    < 0.05       78       5      1
##    < 0.15        4      28      8
##    >= 0.15       2       8    207

##
## Reclassification Table for control:

```

```

##          New
## Standard < 0.05 < 0.15 >= 0.15
##   < 0.05    12720     157      2
##   < 0.15      174     422      56
##   >= 0.15       1      71     284

```

```

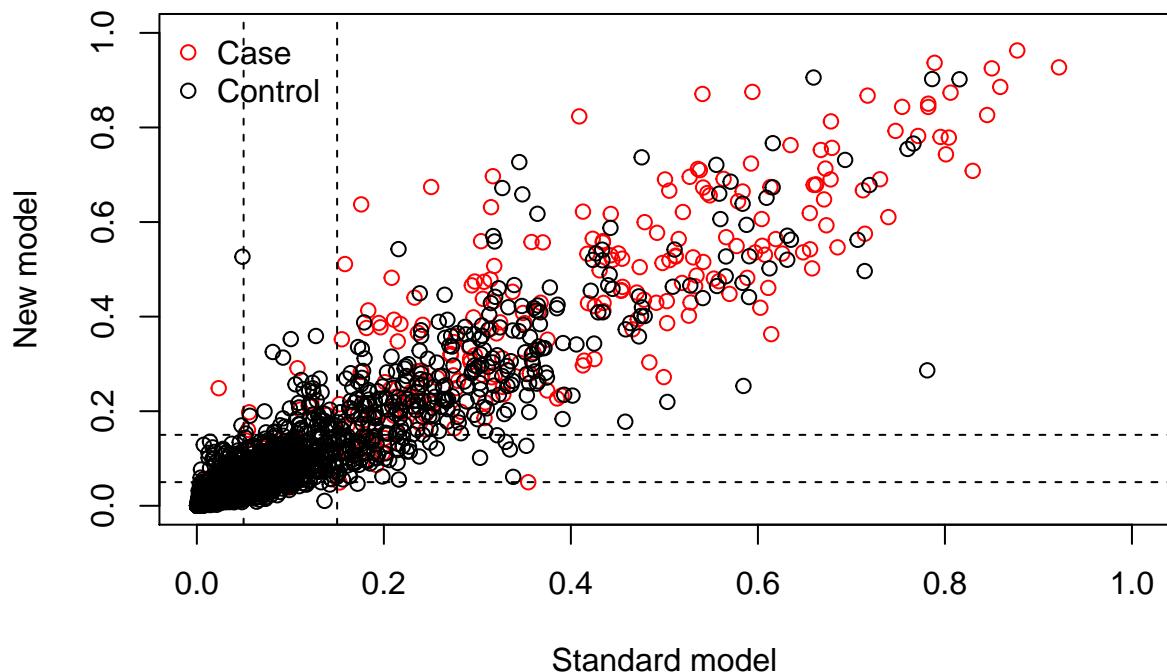
## 
## NRI estimation:

```

```

## Point estimates:

```



```

##          Estimate
## NRI           0.002232304
## NRI+          0.000000000
## NRI-          0.002232304
## Pr(Up|Case)  0.041055718
## Pr(Down|Case) 0.041055718
## Pr(Down|Ctrl) 0.017714409
## Pr(Up|Ctrl)   0.015482106

```

```
NRI3_7vs4 <- nribin(mdl.std = model4, mdl.new = model7, updown = 'category', cut = c(0.05,0.15), niter =
```

```

## 
## UP and DOWN calculation:

```

```

##   #of total, case, and control subjects at t0: 14228 341 13887

```

```

## 
## Reclassification Table for all subjects:

```

```

##          New
## Standard < 0.05 < 0.15 >= 0.15
##   < 0.05    12909     75      0
##   < 0.15      70     590     27
##   >= 0.15       0     26     531

```

```

##

```

```

##    Reclassification Table for case:

##          New
## Standard < 0.05 < 0.15 >= 0.15
##   < 0.05      77      3      0
##   < 0.15      7     34      5
##   >= 0.15      0      4    211

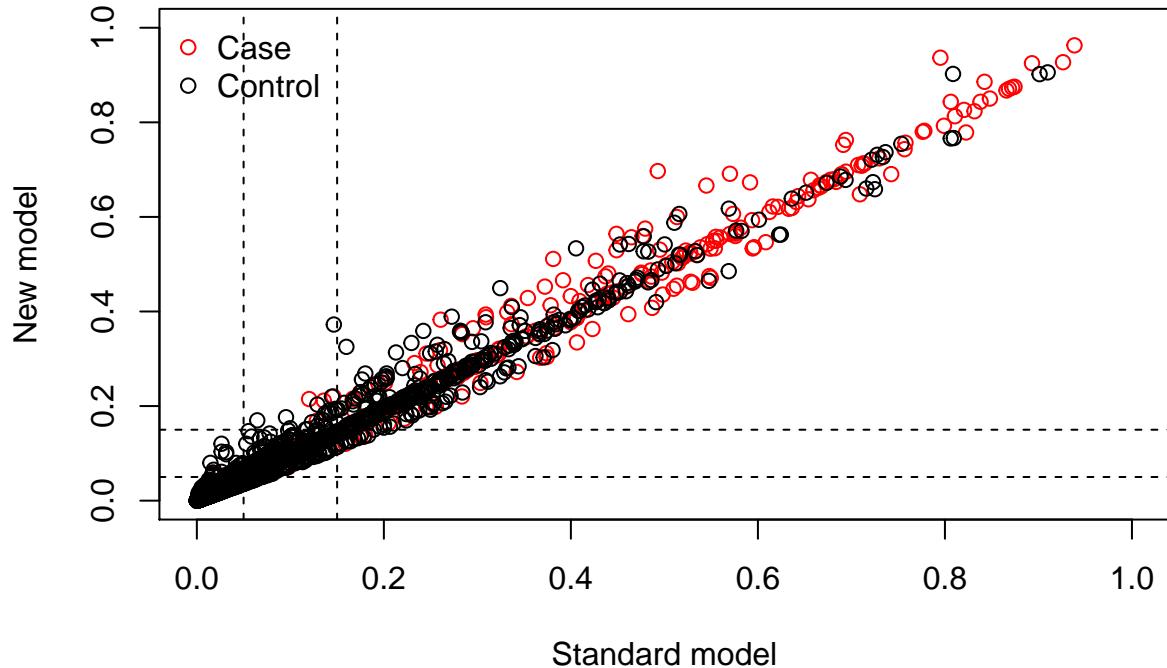
##    Reclassification Table for control:

##          New
## Standard < 0.05 < 0.15 >= 0.15
##   < 0.05    12832     72      0
##   < 0.15      63     556     22
##   >= 0.15      0     22    320

##    NRI estimation:

## Point estimates:

```



```

##                  Estimate
## NRI           -0.0094457421
## NRI+
## NRI-
## Pr(Up|Case)  0.0234604106
## Pr(Down|Case) 0.0322580645
## Pr(Down|Ctrl) 0.0061208324
## Pr(Up|Ctrl)   0.0067689206

```

At stage 4.

```
NRI3_4vs2 <- nribin(mdl.std = model2, mdl.new = model4, updown = 'category', cut = c(0.10, 0.20), niter =
```

```

##  

## UP and DOWN calculation:  

##   #of total, case, and control subjects at t0: 14228 341 13887  

##  

##   Reclassification Table for all subjects:  

##  

##      New  

## Standard < 0.1 < 0.2 >= 0.2  

##   < 0.1    13366     83      5  

##   < 0.2      104    159     58  

##  >= 0.2       4     64   385  

##  

##   Reclassification Table for case:  

##  

##      New  

## Standard < 0.1 < 0.2 >= 0.2  

##   < 0.1      98      9      1  

##   < 0.2      10      17     13  

##  >= 0.2      1      9   183  

##  

##   Reclassification Table for control:  

##  

##      New  

## Standard < 0.1 < 0.2 >= 0.2  

##   < 0.1    13268     74      4  

##   < 0.2      94    142     45  

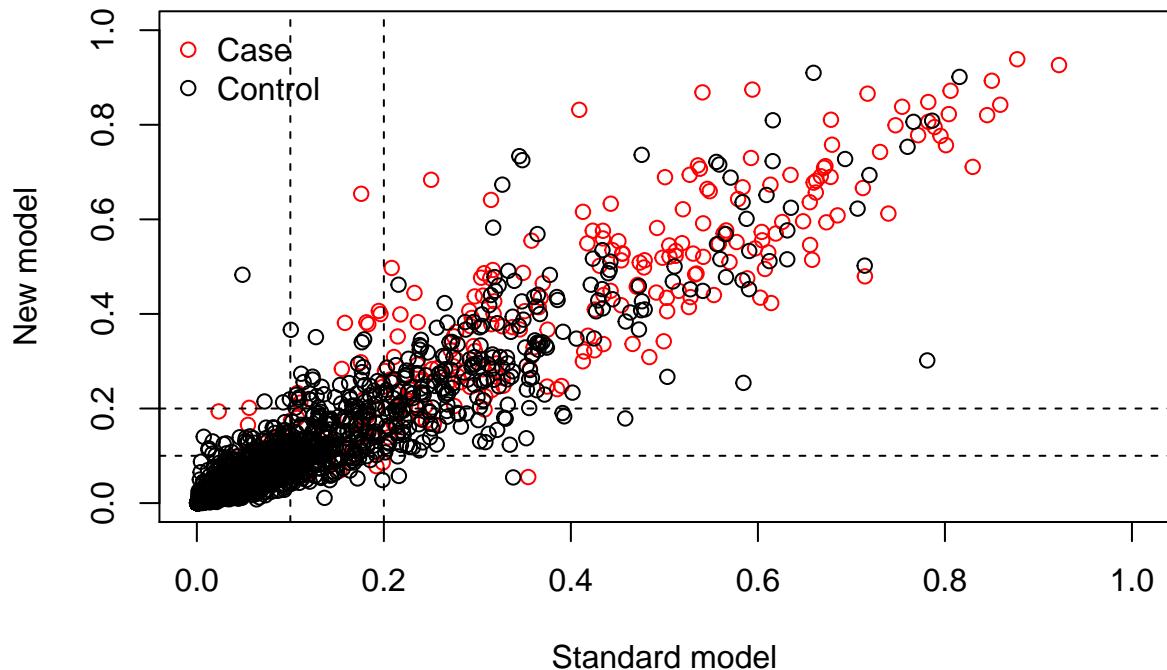
##  >= 0.2       3     55   202  

##  

## NRI estimation:  

## Point estimates:

```



```

##           Estimate
##  NRI          0.010885938
##  NRI+
##  NRI-
##  Pr(Up|Case)  0.067448680
##  Pr(Down|Case) 0.058651026
##  Pr(Down|Ctrl) 0.010945489
##  Pr(Up|Ctrl)   0.008857205

NRI3_7vs2 <- nribin(mdl.std = model2, mdl.new = model7, updown = 'category',cut = c(0.10,0.20), niter = 100)

##
## UP and DOWN calculation:
##   #of total, case, and control subjects at t0: 14228 341 13887

##
## Reclassification Table for all subjects:

##           New
## Standard < 0.1 < 0.2 >= 0.2
##   < 0.1    13344    105      5
##   < 0.2     109    154      58
##   >= 0.2      5     62    386

##
## Reclassification Table for case:

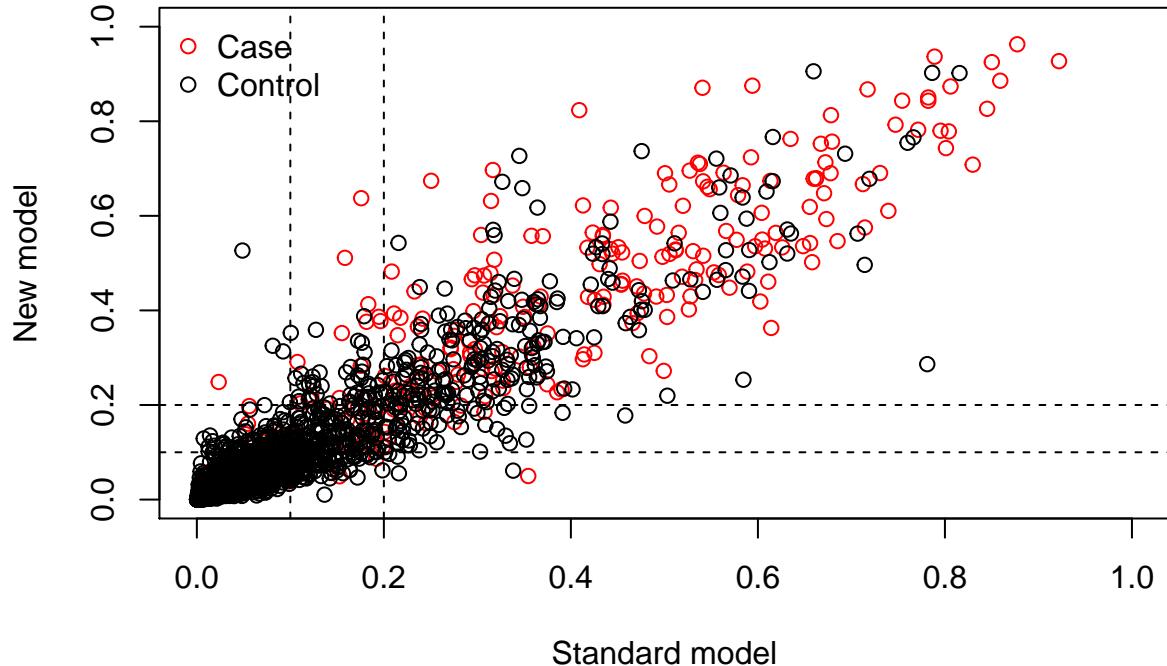
##           New
## Standard < 0.1 < 0.2 >= 0.2
##   < 0.1      98      9      1
##   < 0.2       8     18     14
##   >= 0.2      1     12   180

##
## Reclassification Table for control:

##           New
## Standard < 0.1 < 0.2 >= 0.2
##   < 0.1    13246    96      4
##   < 0.2    101    136     44
##   >= 0.2      4     50    206

##
## NRI estimation:
## Point estimates:

```



```

##           Estimate
## NRI          0.0095897617
## NRI+         0.0087976540
## NRI-         0.0007921077
## Pr(Up|Case)  0.0703812317
## Pr(Down|Case) 0.0615835777
## Pr(Down|Ctrl) 0.0111616180
## Pr(Up|Ctrl)   0.0103694102

NRI3_7vs4 <- nribin(mdl.std = model4, mdl.new = model7, updown = 'category', cut = c(0.10, 0.20), niter = 1000)

##
## UP and DOWN calculation:
##   #of total, case, and control subjects at t0: 14228 341 13887

##
## Reclassification Table for all subjects:
##           New
## Standard < 0.1 < 0.2 >= 0.2
##   < 0.1    13424    50      0
##   < 0.2       34    243     29
##   >= 0.2       0     28   420

##
## Reclassification Table for case:
##           New
## Standard < 0.1 < 0.2 >= 0.2
##   < 0.1     104      5      0
##   < 0.2       3     26      6
##   >= 0.2       0      8   189

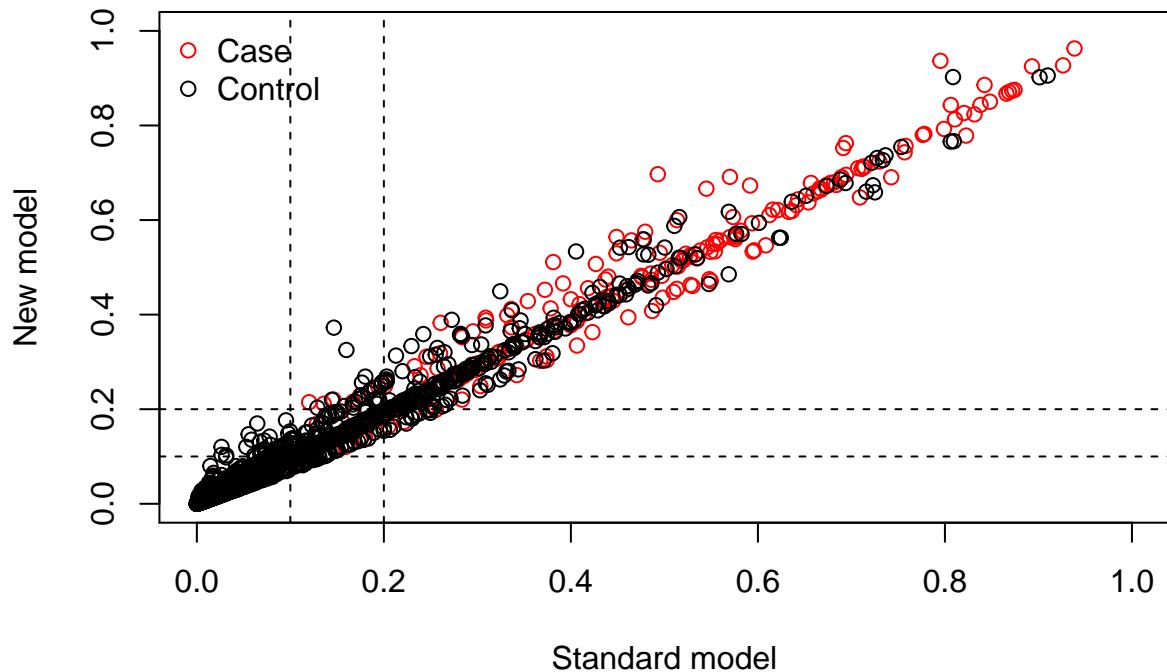
##
## Reclassification Table for control:

```

```

##          New
## Standard < 0.1 < 0.2 >= 0.2
##   < 0.1    13320     45      0
##   < 0.2      31    217     23
##   >= 0.2      0     20   231
##
## NRI estimation:
## Point estimates:

```



```

##           Estimate
## NRI        -0.001224166
## NRI+       0.000000000
## NRI-       -0.001224166
## Pr(Up|Case) 0.032258065
## Pr(Down|Case) 0.032258065
## Pr(Down|Ctrl) 0.003672499
## Pr(Up|Ctrl)  0.004896666

```

So we can see model 4 has an NRI of 0.0057 and 0.01, model 4 is the best model.