

Analyst Price Target Analysis using GML

Logistic regression model

```
#Drop Cols that will not be used in model
finaldf <- read.csv("AnalystRatingData/ratings.csv", check.names = TRUE)
drops <- c("Date", "Firm", "Action", "Rating", "Ticker", "", "Price.After.One.Year", "Price.Target")
finaldf <- finaldf[, !(names(finaldf) %in% drops)]
```

```
#apply factors to columns that have 0 and 1 and as.numeric to columns that are numbers
```

```
#factor_cols <- -which(names(finaldf) %in% c('Price Target', 'Price After One Year'))
#finaldf[factor_cols] <- lapply(finaldf[factor_cols], factor)
#numeric_cols <- c('Price.Target', 'Price.After.One.Year')
#finaldf[numeric_cols] <- lapply(finaldf[numeric_cols], as.numeric)

finaldf[names(finaldf)] <- lapply(finaldf[names(finaldf)], factor)
```

```
nzv <- nearZeroVar(finaldf, saveMetrics = TRUE)
sorted <- nzv[order(nzv$freqRatio), ]
print(sorted)
```

##	freqRatio	percentUnique	zeroVar
## Reiterated.Rating	1.081721	0.03792907	FALSE
## TargetAchieved	1.348775	0.03792907	FALSE
## Buy	1.465171	0.03792907	FALSE
## Lower.Price.Target	4.813671	0.03792907	FALSE
## Boost.Price.Target	4.845898	0.03792907	FALSE
## Outperform	5.383777	0.03792907	FALSE
## Neutral	8.466786	0.03792907	FALSE
## Hold	10.929864	0.03792907	FALSE
## Overweight	11.584726	0.03792907	FALSE
## Initiated.Coverage	12.731771	0.03792907	FALSE
## Royal.Bank.Of.Canada	17.566901	0.03792907	FALSE
## Deutsche.Bank.AG	18.675373	0.03792907	FALSE
## Upgrade	18.749064	0.03792907	FALSE
## Jefferies.Group.LLC	20.610656	0.03792907	FALSE
## Piper.Jaffray.Companies	21.438298	0.03792907	FALSE
## Downgrade	22.026201	0.03792907	FALSE
## Barclays.PLC	24.597087	0.03792907	FALSE
## Credit.Suisse.Group	26.041026	0.03792907	FALSE
## B..Riley	27.502703	0.03792907	FALSE
## Cowen.and.Company	27.657609	0.03792907	FALSE
## Morgan.Stanley	28.294444	0.03792907	FALSE
## X	28.960227	0.03792907	FALSE
## Goldman.Sachs.Group..Inc...The.	31.549383	0.03792907	FALSE
## Market.Perform	31.956250	0.03792907	FALSE
## J.P.Morgan.Chase...Co	32.163522	0.03792907	FALSE
## Citigroup.Inc.	37.772059	0.03792907	FALSE
## Equal.Weight	38.059259	0.03792907	FALSE
## Robert.W..Baird	39.561538	0.03792907	FALSE
## Canaccord.Genuity	42.221311	0.03792907	FALSE
## Nomura	45.663717	0.03792907	FALSE
## Stifel.Nicolaus	46.080357	0.03792907	FALSE
## Oppenheimer.Holdings..Inc.	46.504505	0.03792907	FALSE
## Bank.of.America.Corporation	46.936364	0.03792907	FALSE
## BMO.Capital.Markets	48.745283	0.03792907	FALSE
## Susquehanna.Bancshares.Inc	51.730000	0.03792907	FALSE
## Wedbush	52.806122	0.03792907	FALSE
## Mizuho	53.360825	0.03792907	FALSE
## Sell	55.095745	0.03792907	FALSE
## Leerink.Swann	56.945055	0.03792907	FALSE
## MKM.Partners	57.588889	0.03792907	FALSE
## Pacific.Crest	58.247191	0.03792907	FALSE
## Sector.Perform	58.247191	0.03792907	FALSE
## FBR...Co	58.920455	0.03792907	FALSE
## Macquarie	68.381579	0.03792907	FALSE
## Cantor.Fitzgerald	70.256757	0.03792907	FALSE
## Brean.Capital	72.236111	0.03792907	FALSE
## Sanford.C..Bernstein	73.267606	0.03792907	FALSE
## Evercore.ISI	75.420290	0.03792907	FALSE
## SunTrust.Banks..Inc.	75.420290	0.03792907	FALSE
## Sterne.Agee.CRT	82.698413	0.03792907	FALSE
## Topeka.Capital.Markets	85.442623	0.03792907	FALSE
## Positive	85.442623	0.03792907	FALSE

## Argus	91.508772	0.03792907	FALSE
## Needham...Company.LLC	91.508772	0.03792907	FALSE
## Set.Price.Target	102.392157	0.03792907	FALSE
## Raymond.James.Financial..Inc.	111.191489	0.03792907	FALSE
## Underperform	113.630435	0.03792907	FALSE
## BTIG.Research	116.177778	0.03792907	FALSE
## Pivotal.Research	124.547619	0.03792907	FALSE
## Underweight	130.825000	0.03792907	FALSE
## Strong.Buy	145.472222	0.03792907	FALSE
## Axiom.Securities	174.766667	0.03792907	FALSE
## JMP.Securities	174.766667	0.03792907	FALSE
## Telsey.Advisory.Group	174.766667	0.03792907	FALSE
## Summit.Research	201.807692	0.03792907	FALSE
## Maxim.Group	218.708333	0.03792907	FALSE
## Wunderlich	228.260870	0.03792907	FALSE
## Benchmark.Co.	238.681818	0.03792907	FALSE
## Drexel.Hamilton	238.681818	0.03792907	FALSE
## Rosenblatt.Securities	262.650000	0.03792907	FALSE
## Buckingham.Research	276.526316	0.03792907	FALSE
## Craig.Hallum	276.526316	0.03792907	FALSE
## S.P.Equity.Research	291.944444	0.03792907	FALSE
## Global.Equities.Research	328.562500	0.03792907	FALSE
## KeyCorp	350.533333	0.03792907	FALSE
## Roth.Capital	375.642857	0.03792907	FALSE
## Moffett.Nathanson	404.615385	0.03792907	FALSE
## Standpoint.Research	404.615385	0.03792907	FALSE
## Wells.Fargo...Co	404.615385	0.03792907	FALSE
## Market.Outperform	404.615385	0.03792907	FALSE
## Monness.Crespi...Hardt	438.416667	0.03792907	FALSE
## Top.Pick	438.416667	0.03792907	FALSE
## Chardan.Capital	478.363636	0.03792907	FALSE
## FBN.Securities	478.363636	0.03792907	FALSE
## HC.Wainwright	478.363636	0.03792907	FALSE
## Conviction.Buy	526.300000	0.03792907	FALSE
## Atlantic.Securities	584.888889	0.03792907	FALSE
## BNP.Paribas	584.888889	0.03792907	FALSE
## Stephens	584.888889	0.03792907	FALSE
## Avondale.Partners	658.125000	0.03792907	FALSE
## Gabelli	658.125000	0.03792907	FALSE
## William.Blair	658.125000	0.03792907	FALSE
## Compass.Point	752.285714	0.03792907	FALSE
## Dougherty...Co	752.285714	0.03792907	FALSE
## HSBC.Holdings.plc	752.285714	0.03792907	FALSE
## Reduce	752.285714	0.03792907	FALSE
## Barrington.Research	877.833333	0.03792907	FALSE
## Wolfe.Research	877.833333	0.03792907	FALSE
## Berenberg.Bank	1053.600000	0.03792907	FALSE
## Janney.Montgomery.Scott	1053.600000	0.03792907	FALSE
## Northland.Securities	1053.600000	0.03792907	FALSE
## Societe.Generale	1053.600000	0.03792907	FALSE
## Ascendant.Capital.Markets	1317.250000	0.03792907	FALSE
## BB.T.Corporation	1317.250000	0.03792907	FALSE
## DA.Davidson	1317.250000	0.03792907	FALSE
## Guggenheim	1317.250000	0.03792907	FALSE

## Jyske.Bank	1317.250000	0.03792907	FALSE
## Keefe..Bruyette...Woods	1317.250000	0.03792907	FALSE
## Longbow.Research	1317.250000	0.03792907	FALSE
## Scotiabank	1317.250000	0.03792907	FALSE
## Independent.Research.GmbH	1756.666667	0.03792907	FALSE
## UBS.Group.AG	1756.666667	0.03792907	FALSE
## In.Line	1756.666667	0.03792907	FALSE
## Mkt.Perform	1756.666667	0.03792907	FALSE
## Boenning.Scattergood	2635.500000	0.03792907	FALSE
## Evercore.Partners.Inc	2635.500000	0.03792907	FALSE
## Hilliard.Lyons	2635.500000	0.03792907	FALSE
## Imperial.Capital	2635.500000	0.03792907	FALSE
## T.H..Capital	2635.500000	0.03792907	FALSE
## BGC.Financial	5272.000000	0.03792907	FALSE
## CRT.Capital	5272.000000	0.03792907	FALSE
## Feltl...Co.	5272.000000	0.03792907	FALSE
## First.Analysis	5272.000000	0.03792907	FALSE
## Hovde.Group	5272.000000	0.03792907	FALSE
## Mitsubishi.UFJ.Financial.Group	5272.000000	0.03792907	FALSE
## MLV...Co.	5272.000000	0.03792907	FALSE
## National.Bank.Financial	5272.000000	0.03792907	FALSE
## Sidoti	5272.000000	0.03792907	FALSE
## Taglich.Brothers	5272.000000	0.03792907	FALSE
## Williams.Capital	5272.000000	0.03792907	FALSE
## But.Estimates.Debate.As.Orders.Slow	5272.000000	0.03792907	FALSE
## Focus.List	5272.000000	0.03792907	FALSE
## Overweight.Rating.	5272.000000	0.03792907	FALSE
## Sector.Outperform	5272.000000	0.03792907	FALSE
## Weight	5272.000000	0.03792907	FALSE
##	nzv		
## Reiterated.Rating	FALSE		
## TargetAchieved	FALSE		
## Buy	FALSE		
## Lower.Price.Target	FALSE		
## Boost.Price.Target	FALSE		
## Outperform	FALSE		
## Neutral	FALSE		
## Hold	FALSE		
## Overweight	FALSE		
## Initiated.Coverage	FALSE		
## Royal.Bank.Of.Canada	FALSE		
## Deutsche.Bank.AG	FALSE		
## Upgrade	FALSE		
## Jefferies.Group.LLC	TRUE		
## Piper.Jaffray.Companies	TRUE		
## Downgrade	TRUE		
## Barclays.PLC	TRUE		
## Credit.Suisse.Group	TRUE		
## B..Riley	TRUE		
## Cowen.and.Company	TRUE		
## Morgan.Stanley	TRUE		
## X	TRUE		
## Goldman.Sachs.Group..Inc...The.	TRUE		
## Market.Perform	TRUE		

## J.P.Morgan.Chase...Co	TRUE
## Citigroup.Inc.	TRUE
## Equal.Weight	TRUE
## Robert.W..Baird	TRUE
## Canaccord.Genuity	TRUE
## Nomura	TRUE
## Stifel.Nicolaus	TRUE
## Oppenheimer.Holdings..Inc.	TRUE
## Bank.of.America.Corporation	TRUE
## BMO.Capital.Markets	TRUE
## Susquehanna.Bancshares.Inc	TRUE
## Wedbush	TRUE
## Mizuho	TRUE
## Sell	TRUE
## Leerink.Swann	TRUE
## MKM.Partners	TRUE
## Pacific.Crest	TRUE
## Sector.Perform	TRUE
## FBR...Co	TRUE
## Macquarie	TRUE
## Cantor.Fitzgerald	TRUE
## Brean.Capital	TRUE
## Sanford.C..Bernstein	TRUE
## Evercore.ISI	TRUE
## SunTrust.Banks..Inc.	TRUE
## Sterne.Agee.CRT	TRUE
## Topeka.Capital.Markets	TRUE
## Positive	TRUE
## Argus	TRUE
## Needham...Company.LLC	TRUE
## Set.Price.Target	TRUE
## Raymond.James.Financial..Inc.	TRUE
## Underperform	TRUE
## BTIG.Research	TRUE
## Pivotal.Research	TRUE
## Underweight	TRUE
## Strong.Buy	TRUE
## Axiom.Securities	TRUE
## JMP.Securities	TRUE
## Telsey.Advisory.Group	TRUE
## Summit.Research	TRUE
## Maxim.Group	TRUE
## Wunderlich	TRUE
## Benchmark.Co.	TRUE
## Drexel.Hamilton	TRUE
## Rosenblatt.Securities	TRUE
## Buckingham.Research	TRUE
## Craig.Hallum	TRUE
## S.P.Equity.Research	TRUE
## Global.Equities.Research	TRUE
## KeyCorp	TRUE
## Roth.Capital	TRUE
## Moffett.Nathanson	TRUE
## Standpoint.Research	TRUE

## Wells.Fargo...Co	TRUE
## Market.Outperform	TRUE
## Monness.Crespi...Hardt	TRUE
## Top.Pick	TRUE
## Chardan.Capital	TRUE
## FBN.Securities	TRUE
## HC.Wainwright	TRUE
## Conviction.Buy	TRUE
## Atlantic.Securities	TRUE
## BNP.Paribas	TRUE
## Stephens	TRUE
## Avondale.Partners	TRUE
## Gabelli	TRUE
## William.Blair	TRUE
## Compass.Point	TRUE
## Dougherty...Co	TRUE
## HSBC.Holdings.plc	TRUE
## Reduce	TRUE
## Barrington.Research	TRUE
## Wolfe.Research	TRUE
## Berenberg.Bank	TRUE
## Janney.Montgomery.Scott	TRUE
## Northland.Securities	TRUE
## Societe.Generale	TRUE
## Ascendant.Capital.Markets	TRUE
## BB.T.Corporation	TRUE
## DA.Davidson	TRUE
## Guggenheim	TRUE
## Jyske.Bank	TRUE
## Keefe..Bruyette...Woods	TRUE
## Longbow.Research	TRUE
## Scotiabank	TRUE
## Independent.Research.GmbH	TRUE
## UBS.Group.AG	TRUE
## In.Line	TRUE
## Mkt.Perform	TRUE
## Boenning.Scattergood	TRUE
## Evercore.Partners.Inc	TRUE
## Hilliard.Lyons	TRUE
## Imperial.Capital	TRUE
## T.H..Capital	TRUE
## BGC.Financial	TRUE
## CRT.Capital	TRUE
## Feltl...Co.	TRUE
## First.Analysis	TRUE
## Hovde.Group	TRUE
## Mitsubishi.UFJ.Financial.Group	TRUE
## MLV...Co.	TRUE
## National.Bank.Financial	TRUE
## Sidoti	TRUE
## Taglich.Brothers	TRUE
## Williams.Capital	TRUE
## But.Estimates.Debate.As.Orders.Slow	TRUE
## Focus.List	TRUE

```
## Overweight.Rating.          TRUE
## Sector.Outperform          TRUE
## Weight                     TRUE
```

```
#remove features that with high freq ratio (which appears less frequently). freqRatio o
f 700 means it appears once in 700 times. This
#helped us to remove some of analyst/banks that are less known and
remove_features <- rownames(nzv[nzv$freqRatio > 700, ])
#this reduced number of columns from 137 to 94
finaldf <- finaldf[,!names(finaldf)%in% remove_features]
```

```
set.seed(417)
#make.name will replace space with dot in column names
names(finaldf)<-make.names(names(finaldf),unique=TRUE)
finaldf<-x<-na.omit(finaldf)
inTrain <- createDataPartition(y=finaldf$"TargetAchieved",p = 0.7, list=FALSE)
training <- finaldf[inTrain,]
testing <- finaldf[-inTrain,]
```

```
system.time(
  model <-glm(TargetAchieved ~ ., data=testing, family = "binomial")
)
```

```
##      user  system elapsed
##    0.280    0.012    0.296
```

```
summary(model)
```

```
##
## Call:
## glm(formula = TargetAchieved ~ ., family = "binomial", data = testing)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1210  -0.9694  -0.6347   1.1134   2.2042
##
## Coefficients: (1 not defined because of singularities)
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.074e-01  9.903e-01   0.108  0.91366
## Argus1          -3.460e-01  7.747e-01  -0.447  0.65517
## Atlantic.Securities1 -1.550e+01  2.400e+03  -0.006  0.99485
## Avondale.Partners1   -1.655e+01  2.400e+03  -0.007  0.99450
## Axiom.Securities1     4.202e-01  7.829e-01   0.537  0.59146
## B..Riley1           1.180e+00  5.082e-01   2.322  0.02022 *
## Bank.of.America.Corporation1 1.222e+00  5.776e-01   2.116  0.03434 *
## Barclays.PLC1        5.785e-01  5.826e-01   0.993  0.32067
## Benchmark.Co.1       2.046e+00  1.204e+00   1.699  0.08933 .
## BMO.Capital.Markets1  7.241e-01  5.667e-01   1.278  0.20130
## BNP.Paribas1         7.255e-01  1.540e+00   0.471  0.63752
## Brean.Capital1      -2.457e-01  6.554e-01  -0.375  0.70778
## BTIG.Research1       1.809e-01  7.891e-01   0.229  0.81872
## Buckingham.Research1  1.456e+00  8.749e-01   1.664  0.09604 .
## Canaccord.Genuity1   7.337e-01  5.611e-01   1.308  0.19101
## Cantor.Fitzgerald1   1.185e+00  5.986e-01   1.980  0.04769 *
## Chardan.Capital1     3.040e+00  1.408e+00   2.160  0.03080 *
## Citigroup.Inc.1      5.656e-01  5.268e-01   1.074  0.28303
## Cowen.and.Company1   4.290e-01  5.206e-01   0.824  0.40990
## Craig.Hallum1        1.059e+00  8.262e-01   1.282  0.19994
## Credit.Suisse.Group1  5.070e-01  5.009e-01   1.012  0.31145
## Deutsche.Bank.AG1    5.931e-01  4.775e-01   1.242  0.21420
## Drexel.Hamilton1     1.470e+00  8.936e-01   1.645  0.10001
## Evercore.ISI1       -2.339e-01  6.787e-01  -0.345  0.73034
## FBN.Securities1     -1.649e+01  1.687e+03  -0.010  0.99220
## FBR...Col           1.016e+00  6.330e-01   1.605  0.10851
## Gabelli1           1.158e+00  1.482e+00   0.782  0.43430
## Global.Equities.Research1 -1.709e-01  1.204e+00  -0.142  0.88707
## Goldman.Sachs.Group..Inc...The.1 1.356e+00  5.190e-01   2.613  0.00896 **
## HC.Wainwright1       1.284e+00  1.337e+00   0.960  0.33691
## J.P.Morgan.Chase...Col  7.249e-02  5.606e-01   0.129  0.89712
## Jefferies.Group.LLC1  1.151e-01  4.913e-01   0.234  0.81484
## JMP.Securities1      1.474e+00  1.008e+00   1.463  0.14359
## KeyCorp1            4.050e-01  1.028e+00   0.394  0.69350
## Leerink.Swann1      -1.137e+00  7.097e-01  -1.602  0.10923
## Macquarie1          6.211e-01  6.120e-01   1.015  0.31016
## Maxim.Group1         4.033e-01  8.917e-01   0.452  0.65106
## Mizuho1             9.492e-01  5.426e-01   1.749  0.08023 .
## MKM.Partners1       1.123e+00  5.972e-01   1.881  0.05996 .
## Moffett.Nathanson1    1.003e+00  8.943e-01   1.122  0.26205
## Monness.Crespi...Hardt1 2.061e+00  1.238e+00   1.666  0.09577 .
## Morgan.Stanley1      2.800e-01  5.787e-01   0.484  0.62855
## Needham...Company.LLC1 1.633e+00  7.334e-01   2.227  0.02597 *
```


## Nomural	1.212e+00	5.147e-01	2.355	0.01854 *
## Oppenheimer.Holdings..Inc.1	-1.602e-01	6.421e-01	-0.249	0.80299
## Pacific.Crest1	7.847e-01	6.045e-01	1.298	0.19428
## Piper.Jaffray.Companies1	2.743e-01	5.056e-01	0.542	0.58751
## Pivotal.Research1	-2.439e-02	8.216e-01	-0.030	0.97632
## Raymond.James.Financial..Inc.1	3.318e+00	1.199e+00	2.766	0.00567 **
## Robert.W..Baird1	6.127e-01	5.280e-01	1.160	0.24590
## Rosenblatt.Securities1	2.849e-01	9.712e-01	0.293	0.76929
## Roth.Capital1	-5.006e-01	1.254e+00	-0.399	0.68969
## Royal.Bank.Of.Canada1	4.202e-01	5.249e-01	0.800	0.42345
## S.P.Equity.Research1	-1.469e+01	1.606e+03	-0.009	0.99271
## Sanford.C..Bernstein1	5.952e-01	6.475e-01	0.919	0.35796
## Standpoint.Research1	-5.399e-01	1.182e+00	-0.457	0.64781
## Stephens1	-1.594e+01	1.324e+03	-0.012	0.99039
## Sterne.Agee.CRT1	9.770e-01	6.941e-01	1.408	0.15927
## Stifel.Nicolaus1	4.573e-01	6.100e-01	0.750	0.45350
## Summit.Research1	-1.017e+00	1.167e+00	-0.871	0.38359
## SunTrust.Banks..Inc.1	9.503e-01	6.203e-01	1.532	0.12550
## Susquehanna.Bancshares.Inc1	2.368e-01	6.504e-01	0.364	0.71576
## Telsey.Advisory.Group1	-5.712e-04	7.653e-01	-0.001	0.99940
## Topeka.Capital.Markets1	4.081e-01	6.125e-01	0.666	0.50522
## Wedbush1	4.289e-01	6.436e-01	0.666	0.50512
## Wells.Fargo...Col	-1.537e+01	2.400e+03	-0.006	0.99489
## William.Blair1	1.131e+00	1.093e+00	1.035	0.30049
## Wunderlich1	6.684e-01	9.790e-01	0.683	0.49482
## Boost.Price.Target1	2.401e-01	2.954e-01	0.813	0.41637
## Downgrade1	-5.532e-01	3.881e-01	-1.425	0.15402
## Initiated.Coverage1	-2.088e-01	3.346e-01	-0.624	0.53263
## Lower.Price.Target1	4.321e-02	2.975e-01	0.145	0.88451
## Reiterated.Rating1	-2.290e-01	2.764e-01	-0.828	0.40748
## Set.Price.Target1	-7.840e-02	6.349e-01	-0.123	0.90172
## Upgrade1	NA	NA	NA	NA
## X1	-2.969e-01	9.993e-01	-0.297	0.76635
## Buy1	-1.183e+00	9.451e-01	-1.252	0.21067
## Conviction.Buy1	-1.794e+01	1.071e+03	-0.017	0.98663
## Equal.Weight1	-3.231e-01	1.053e+00	-0.307	0.75905
## Hold1	-1.047e-01	9.617e-01	-0.109	0.91333
## Market.Outperform1	-2.053e+00	1.740e+00	-1.180	0.23795
## Market.Perform1	8.536e-02	1.037e+00	0.082	0.93439
## Neutral1	6.220e-03	9.532e-01	0.007	0.99479
## Outperform1	-1.079e+00	9.673e-01	-1.115	0.26477
## Overweight1	-9.682e-01	9.882e-01	-0.980	0.32718
## Positive1	-1.146e+00	1.143e+00	-1.003	0.31588
## Sector.Perform1	1.781e-01	1.064e+00	0.167	0.86706
## Sell1	-2.225e+00	1.086e+00	-2.050	0.04039 *
## Strong.Buy1	-3.819e+00	1.569e+00	-2.434	0.01493 *
## Top.Pick1	1.619e+01	1.382e+03	0.012	0.99065
## Underperform1	-3.274e+00	1.430e+00	-2.290	0.02202 *
## Underweight1	-1.683e+01	6.476e+02	-0.026	0.97926
## ---				
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
##				
## (Dispersion parameter for binomial family taken to be 1)				
##				

```
##      Null deviance: 2156.7  on 1580  degrees of freedom
## Residual deviance: 1913.8  on 1490  degrees of freedom
## AIC: 2095.8
##
## Number of Fisher Scoring iterations: 15
```

```
glm.probs <-predict(model,testing,type = "response")
```

```
## Warning in predict.lm(object, newdata, se.fit, scale = 1, type =
## ifelse(type == : prediction from a rank-deficient fit may be misleading
```

```
#this creates a list with all 0 with same size as testing
glm.predict <- rep( 0 ,nrow(testing))
```

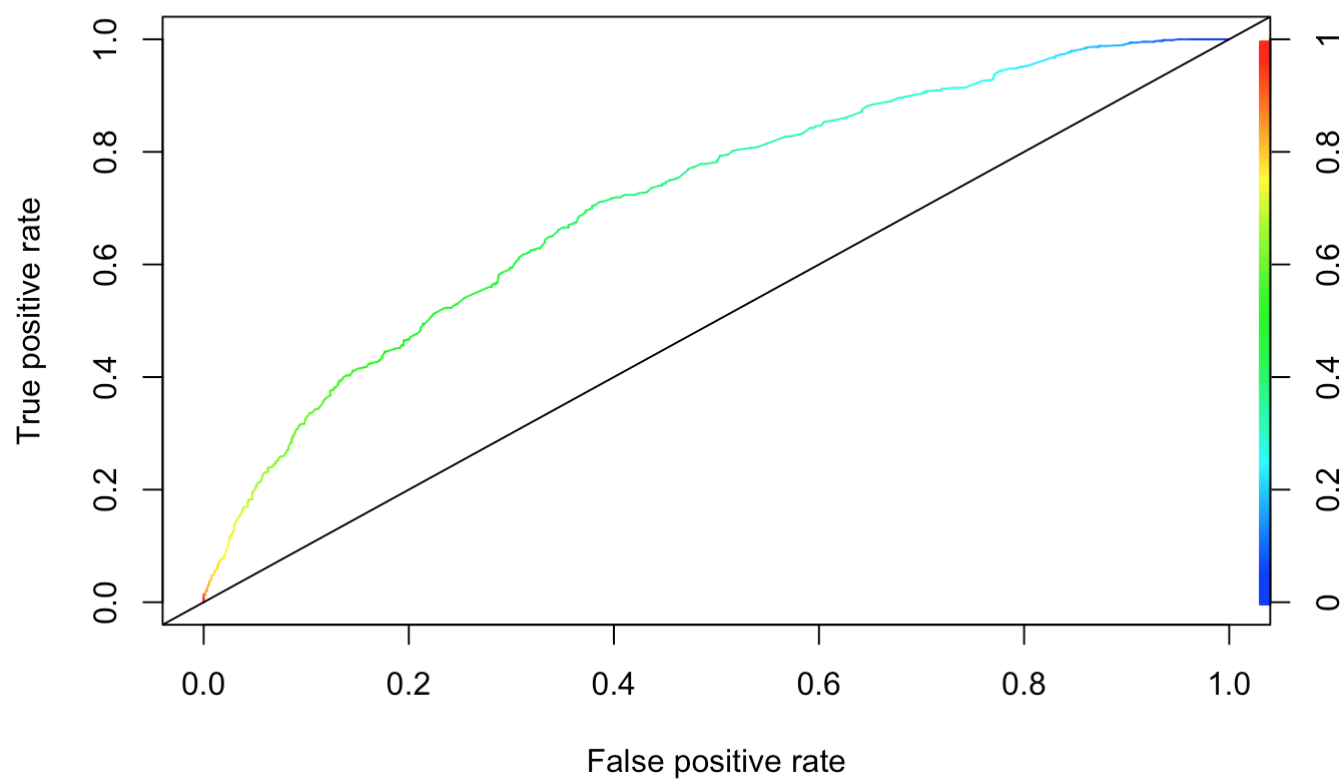
```
#In our case having higher precission is more important as getting False Positive rate i
s costly
 #(When you predict True but its actually false thus you loose money! or cant get your gu
 ranteed
#return by following analyst) and False Negative is not important because what you predi
 cted false
#and it is actually true. (ignored that rating and that rating was correct)
```

```
#Check the threshold graph to see relation between precission and recall in our case
```

```
glm.predict[glm.probs > .65] <- 1 #cutoff set to 0.5
confusionMatrix(glm.predict, testing$TargetAchieved, positive ="1", mode="prec_recall")
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##           0 851 512
##           1  57 161
##
##           Accuracy : 0.6401
##           95% CI : (0.6159, 0.6638)
##       No Information Rate : 0.5743
##       P-Value [Acc > NIR] : 5.587e-08
##
##           Kappa : 0.1934
##  Mcnemar's Test P-Value : < 2.2e-16
##
##           Precision : 0.7385
##           Recall : 0.2392
##           F1 : 0.3614
##           Prevalence : 0.4257
##       Detection Rate : 0.1018
##  Detection Prevalence : 0.1379
##       Balanced Accuracy : 0.5882
##
##       'Positive' Class : 1
##
```

```
pred <- prediction(glm.probs,testing$TargetAchieved)
perf <- performance(pred,"tpr","fpr")
plot(perf, colorize=TRUE)
abline(a=0, b=1)
```



```
performance(pred,measure = "auc")@y.values
```

```
## [[1]]  
## [1] 0.7128938
```

```

dfLength <- length(seq(0.00,0.999,0.001))
thresholdDF <- data.frame(percent = numeric(dfLength),threshold = numeric(dfLength),
precision = numeric(dfLength), recall = numeric(dfLength), f1 = numeric(dfLength))
dfListIndex <-0
for (i in seq(0.00, 0.999, 0.001)){
dfListIndex = dfListIndex +1
#get minimum value from the sorted list which will be our cutoff value
threshold = min(head(sort(glm.probs,decreasing=TRUE), n = (length(glm.probs)*(1-i))
))
thresholdDF$percent[dfListIndex] = (1-i)*100
thresholdDF$threshold[dfListIndex] = threshold
# set to 1 if probability is above threshold otehr wise 0
glm.pred = rep(0, nrow(testing))
glm.pred[glm.probs >= threshold] = 1
#create confustion matrix and get pos pred value from the confusion matrix
cm = confusionMatrix(glm.pred, testing$TargetAchieved, positive="1", mode="prec_recall")
byClass = cm$byClass
precision = byClass[5]
recall = byClass[6]
f1 = byClass[7]
thresholdDF$precision[dfListIndex] = precision
thresholdDF$recall[dfListIndex] = recall
thresholdDF$f1[dfListIndex] = f1
}

```

```

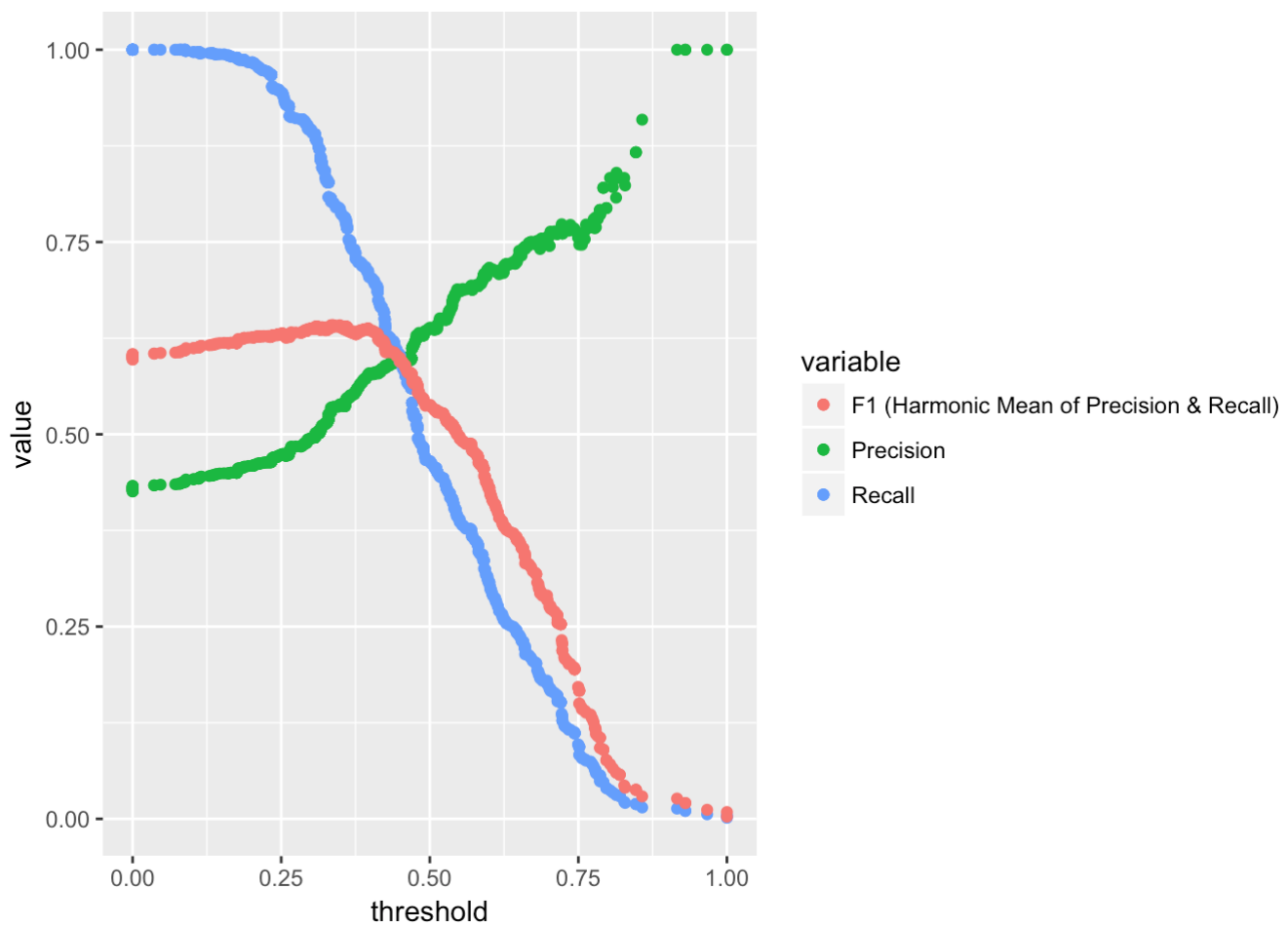
## Warning in confusionMatrix.default(glm.pred, testing$TargetAchieved,
## positive = "1", : Levels are not in the same order for reference and data.
## Refactoring data to match.

```

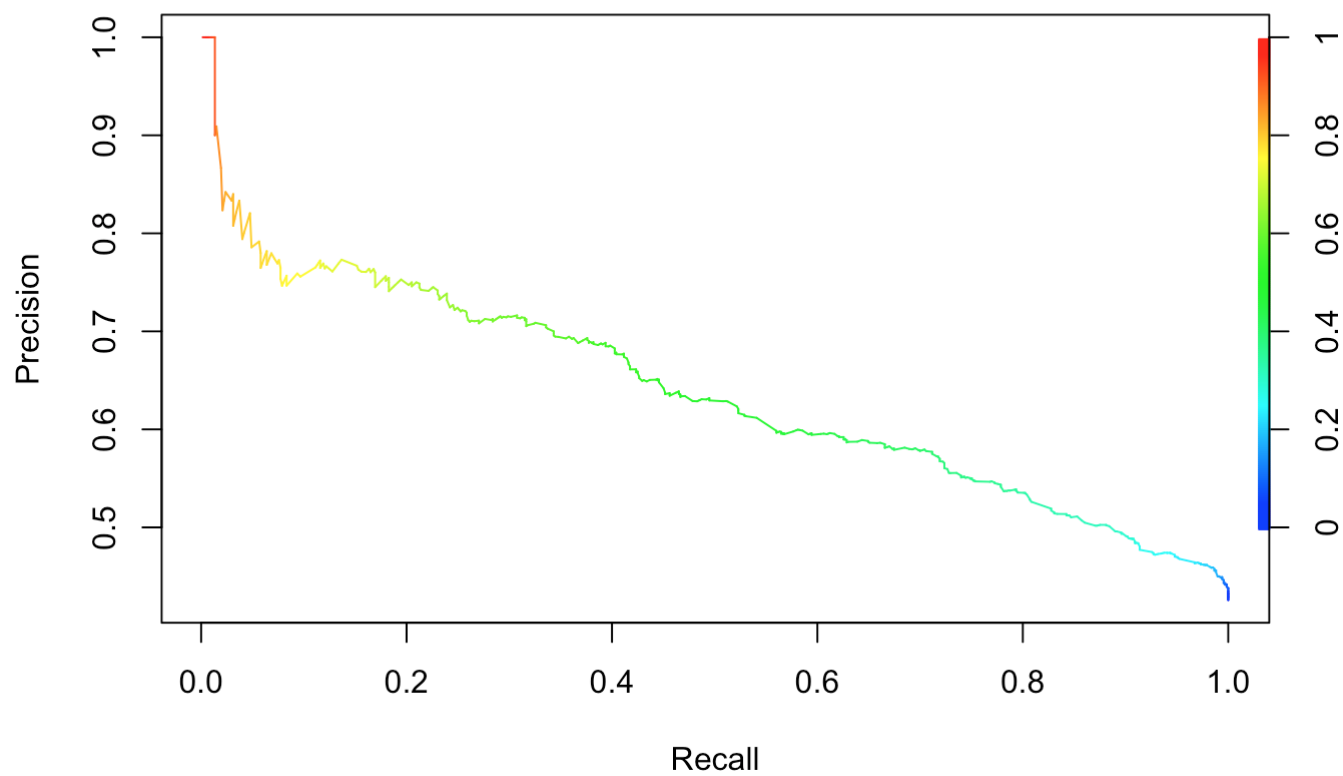
```

# Plot
ggplot(data=thresholdDF,aes(threshold, y=value,color =variable))+ geom_point(aes(y=precision , col = "Precision"))+ geom_point(aes(y=recall , col = "Recall"))+ geom_point
(aes(y=f1 , col = "F1 (Harmonic Mean of Precision & Recall)"))

```



```
perf <- performance(pred,"prec","rec")  
plot(perf, colorize=TRUE)
```



```
library(tibble)
imp <- varImp(model, scale=FALSE)
imp <- rownames_to_column(imp)
# find top N most important variables
print(imp[order(imp$Overall, decreasing=TRUE)[1:25], ])
```

##	rowname	Overall
## 48	Raymond.James.Financial..Inc.1	2.766410
## 28	Goldman.Sachs.Group..Inc...The.1	2.613405
## 87	Strong.Buy1	2.434124
## 43	Nomura1	2.354588
## 5	B..Riley1	2.322303
## 89	Underperform1	2.290000
## 42	Needham...Company.LLC1	2.226641
## 16	Chardan.Capital1	2.159619
## 6	Bank.of.America.Corporation1	2.116048
## 86	Sell1	2.049785
## 15	Cantor.Fitzgerald1	1.980087
## 38	MKM.Partners1	1.881063
## 37	Mizuho1	1.749374
## 8	Benchmark.Co.1	1.698962
## 40	Monness.Crespi...Hardt1	1.665717
## 13	Buckingham.Research1	1.664385
## 22	Drexel.Hamilton1	1.644796
## 25	FBR...Co1	1.604935
## 34	Leerink.Swann1	1.601642
## 60	SunTrust.Banks..Inc.1	1.532099
## 32	JMP.Securities1	1.462539
## 69	Downgrade1	1.425469
## 57	Sterne.Agee.CRT1	1.407545
## 14	Canaccord.Genuity1	1.307602
## 45	Pacific.Crest1	1.298022