H1N1_2

AYA

2025-03-11

CLEAR THE ENVIRONMENT

```
rm(list  =ls())
```

LOAD THE NECESSARY LIBRARIES

```
library(mlbench)
library(caret)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```
#install.packages("caTools")
library(caTools)
#install.packages("ranger")
library(ranger)
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
```

```
#install.packages("doParallel")
library(doParallel)
```

```
## Loading required package: foreach
```

```
## Loading required package: iterators
```

```
## Loading required package: parallel
```

LOAD THE DATASET

```
#load the training features
training_features <- read.csv("training_set_features.csv", row.names = 1, header = T, stringsAsFactors =
head(training_features,5)
```

```
##   h1n1_concern h1n1_knowledge behavioral_antiviral_meds behavioral_avoidance
## 0            1              0                         0                    0
## 1            3              2                         0                    1
## 2            1              1                         0                    1
## 3            1              1                         0                    1
## 4            2              1                         0                    1
##   behavioral_face_mask behavioral_wash_hands behavioral_large_gatherings
## 0                    0                     0                           0
## 1                    0                     1                           0
## 2                    0                     0                           0
## 3                    0                     1                           1
## 4                    0                     1                           1
##   behavioral_outside_home behavioral_touch_face doctor_recc_h1n1
## 0                       1                     1               0
## 1                       1                     1               0
## 2                       0                     0              NA
## 3                       0                     0               0
## 4                       0                     1               0
##   doctor_recc_seasonal chronic_med_condition child_under_6_months health_worker
## 0                    0                     0                     0            0
## 1                    0                     0                     0            0
## 2                   NA                     1                     0            0
## 3                    1                     1                     0            0
## 4                    0                     0                     0            0
##   health_insurance opinion_h1n1_vacc_effective opinion_h1n1_risk
## 0                1                           3                 1
## 1                1                           5                 4
## 2               NA                           3                 1
## 3               NA                           3                 3
## 4               NA                           3                 3
##   opinion_h1n1_sick_from_vacc opinion_seas_vacc_effective opinion_seas_risk
## 0                           2                           2                 1
## 1                           4                           4                 2
## 2                           1                           4                 1
## 3                           5                           5                 4
## 4                           2                           3                 1
##   opinion_seas_sick_from_vacc     age_group        education  race    sex
## 0                           2 55 - 64 Years       < 12 Years White Female
## 1                           4 35 - 44 Years         12 Years White   Male
## 2                           2 18 - 34 Years College Graduate White   Male
## 3                           1     65+ Years         12 Years White Female
## 4                           4 45 - 54 Years     Some College White Female
##              income_poverty marital_status rent_or_own  employment_status
## 0             Below Poverty    Not Married         Own Not in Labor Force
## 1             Below Poverty    Not Married        Rent           Employed
## 2 <= $75,000, Above Poverty    Not Married         Own           Employed
## 3             Below Poverty    Not Married        Rent Not in Labor Force
## 4 <= $75,000, Above Poverty        Married         Own           Employed
##   hhs_geo_region               census_msa household_adults household_children
## 0       oxchjgsf                  Non-MSA               0                  0
## 1       bhuqouqj MSA, Not Principle  City               0                  0
## 2       qufhixun MSA, Not Principle  City               2                  0
```

```
## 3       lrircsnp      MSA, Principle City                0                 0
## 4       qufhixun MSA, Not Principle  City                1                 0
##   employment_industry employment_occupation
## 0
## 1            pxcmvdjn             xgwztkwe
## 2            rucpziij             xtkaffoo
## 3
## 4            wxleyezf             emcorrxb
```

```
#load the validation labels
training_labels <- read.csv("training_set_labels.csv", row.names = 1, header = T, stringsAsFactors = T)
#View(training_labels)
```

```
#combining labels
h1n1 <- cbind(training_features, training_labels[1:2])
```

```
head(h1n1,5)
```

```
##   h1n1_concern h1n1_knowledge behavioral_antiviral_meds behavioral_avoidance
## 0            1              0                         0                    0
## 1            3              2                         0                    1
## 2            1              1                         0                    1
## 3            1              1                         0                    1
## 4            2              1                         0                    1
##   behavioral_face_mask behavioral_wash_hands behavioral_large_gatherings
## 0                    0                     0                           0
## 1                    0                     1                           0
## 2                    0                     0                           0
## 3                    0                     1                           1
## 4                    0                     1                           1
##   behavioral_outside_home behavioral_touch_face doctor_recc_h1n1
## 0                       1                     1               0
## 1                       1                     1               0
## 2                       0                     0              NA
## 3                       0                     0               0
## 4                       0                     1               0
##   doctor_recc_seasonal chronic_med_condition child_under_6_months health_worker
## 0                    0                     0                    0             0
## 1                    0                     0                    0             0
## 2                   NA                     1                    0             0
## 3                    1                     1                    0             0
## 4                    0                     0                    0             0
##   health_insurance opinion_h1n1_vacc_effective opinion_h1n1_risk
## 0                1                           3                 1
## 1                1                           5                 4
## 2               NA                           3                 1
## 3               NA                           3                 3
## 4               NA                           3                 3
##   opinion_h1n1_sick_from_vacc opinion_seas_vacc_effective opinion_seas_risk
## 0                           2                           2                 1
## 1                           4                           4                 2
## 2                           1                           4                 1
## 3                           5                           5                 4
## 4                           2                           3                 1
##   opinion_seas_sick_from_vacc     age_group        education  race    sex
```

```
## 0                          2 55 - 64 Years       < 12 Years White Female
## 1                          4 35 - 44 Years         12 Years White   Male
## 2                          2 18 - 34 Years College Graduate White   Male
## 3                          1     65+ Years         12 Years White Female
## 4                          4 45 - 54 Years     Some College White Female
##              income_poverty marital_status rent_or_own  employment_status
## 0             Below Poverty    Not Married         Own Not in Labor Force
## 1             Below Poverty    Not Married        Rent           Employed
## 2 <= $75,000, Above Poverty    Not Married         Own           Employed
## 3             Below Poverty    Not Married        Rent Not in Labor Force
## 4 <= $75,000, Above Poverty        Married         Own           Employed
##   hhs_geo_region               census_msa household_adults household_children
## 0       oxchjgsf                  Non-MSA                0                  0
## 1       bhuqouqj MSA, Not Principle  City                0                  0
## 2       qufhixun MSA, Not Principle  City                2                  0
## 3       lrircsnp      MSA, Principle City                0                  0
## 4       qufhixun MSA, Not Principle  City                1                  0
##   employment_industry employment_occupation h1n1_vaccine seasonal_vaccine
## 0                                                     0                0
## 1            pxcmvdjn              xgwztkwe            0                1
## 2            rucpziij              xtkaffoo            0                0
## 3                                                     0                1
## 4            wxleyezf              emcorrxb            0                0
```

DATA CLEANING

```
#sum of all NA values in each dataset
colSums(is.na(h1n1))
```

```
##                h1n1_concern             h1n1_knowledge
##                          92                        116
##   behavioral_antiviral_meds       behavioral_avoidance
##                          71                        208
##       behavioral_face_mask      behavioral_wash_hands
##                          19                         42
## behavioral_large_gatherings    behavioral_outside_home
##                          87                         82
##      behavioral_touch_face           doctor_recc_h1n1
##                         128                       2160
##       doctor_recc_seasonal      chronic_med_condition
##                       2160                        971
##       child_under_6_months             health_worker
##                        820                        804
##           health_insurance opinion_h1n1_vacc_effective
##                      12274                        391
##          opinion_h1n1_risk opinion_h1n1_sick_from_vacc
##                        388                        395
## opinion_seas_vacc_effective          opinion_seas_risk
##                        462                        514
## opinion_seas_sick_from_vacc                  age_group
##                        537                          0
##                  education                       race
##                          0                          0
```

```
##                          sex              income_poverty
##                            0                           0
##               marital_status                 rent_or_own
##                            0                           0
##            employment_status              hhs_geo_region
##                            0                           0
##                  census_msa            household_adults
##                            0                         249
##           household_children         employment_industry
##                          249                           0
##       employment_occupation                h1n1_vaccine
##                            0                           0
##             seasonal_vaccine
##                            0
```

health insurance has the largest NA values 12274. Filling up thr data might cause inaccuracies so we will drop the column "health_insurance".

```
#removing "health_insurance" column from the dataset
h1n1$health_insurance <- NULL

#also dropping "hhs_goe_region", "employment_industry" and "employment_occupation"
h1n1$hhs_geo_region <- NULL
h1n1$employment_industry <- NULL
h1n1$employment_occupation <- NULL

#View(h1n1)
```

```
#dropping all NA values in the dataset
h1n1 <- na.omit(h1n1)

#checking the dimension of NA after dropping all the values
dim(h1n1)
```

```
## [1] 22976    33
```

From 26707 entries, we now have a total of 22976(a difference of 3731).

CONVERTING NOMINAL DATA INTO NUMERIC DATA

```
#duplicating dataset
h1n1_2 <- h1n1

h1n1_2[,32:33] <- NULL
#View(h1n1_2)
```

```
#binarising the nominal attributes
binary_data <- dummyVars(~., data = h1n1_2)

#View(binary_data)

#adding the conversion to the data
new_data <- predict(binary_data, newdata = h1n1_2)
```

```
#adding "h1n1_vaccine class" back to original dataset
new_data2 <- cbind(new_data, h1n1[32])
```

```r
#View(new_data2)
```

```r
#converting 0 and 1 to "yes" and "no" for the decision class
new_data2$h1n1_vaccine <- factor(new_data2$h1n1_vaccine, levels = c(0, 1), labels = c("no", "yes"))
head(new_data2,5)
```

```
##   h1n1_concern h1n1_knowledge behavioral_antiviral_meds behavioral_avoidance
## 0            1              0                         0                    0
## 1            3              2                         0                    1
## 3            1              1                         0                    1
## 4            2              1                         0                    1
## 5            3              1                         0                    1
##   behavioral_face_mask behavioral_wash_hands behavioral_large_gatherings
## 0                    0                     0                           0
## 1                    0                     1                           0
## 3                    0                     1                           1
## 4                    0                     1                           1
## 5                    0                     1                           0
##   behavioral_outside_home behavioral_touch_face doctor_recc_h1n1
## 0                       1                     1                0
## 1                       1                     1                0
## 3                       0                     0                0
## 4                       0                     1                0
## 5                       0                     1                0
##   doctor_recc_seasonal chronic_med_condition child_under_6_months health_worker
## 0                    0                     0                    0             0
## 1                    0                     0                    0             0
## 3                    1                     1                    0             0
## 4                    0                     0                    0             0
## 5                    1                     0                    0             0
##   opinion_h1n1_vacc_effective opinion_h1n1_risk opinion_h1n1_sick_from_vacc
## 0                           3                 1                           2
## 1                           5                 4                           4
## 3                           3                 3                           5
## 4                           3                 3                           2
## 5                           5                 2                           1
##   opinion_seas_vacc_effective opinion_seas_risk opinion_seas_sick_from_vacc
## 0                           2                 1                           2
## 1                           4                 2                           4
## 3                           5                 4                           1
## 4                           3                 1                           4
## 5                           5                 4                           4
##   age_group.18 - 34 Years age_group.35 - 44 Years age_group.45 - 54 Years
## 0                       0                       0                       0
## 1                       0                       1                       0
## 3                       0                       0                       0
## 4                       0                       0                       1
## 5                       0                       0                       0
##   age_group.55 - 64 Years age_group.65+ Years education. education.< 12 Years
## 0                       1                   0          0                    1
## 1                       0                   0          0                    0
## 3                       0                   1          0                    0
## 4                       0                   0          0                    0
## 5                       0                   1          0                    0
```

```
##   education.12 Years education.College Graduate education.Some College
## 0                  0                           0                      0
## 1                  1                           0                      0
## 3                  1                           0                      0
## 4                  0                           0                      1
## 5                  1                           0                      0
##   race.Black race.Hispanic race.Other or Multiple race.White sex.Female
## 0          0             0                       0          1          1
## 1          0             0                       0          1          0
## 3          0             0                       0          1          1
## 4          0             0                       0          1          1
## 5          0             0                       0          1          0
##   sex.Male income_poverty. income_poverty.<= $75,000, Above Poverty
## 0        0               0                                       0
## 1        1               0                                       0
## 3        0               0                                       0
## 4        0               0                                       1
## 5        1               0                                       1
##   income_poverty.> $75,000 income_poverty.Below Poverty marital_status.
## 0                        0                            1               0
## 1                        0                            1               0
## 3                        0                            1               0
## 4                        0                            0               0
## 5                        0                            0               0
##   marital_status.Married marital_status.Not Married rent_or_own.
## 0                      0                          1           0
## 1                      0                          1           0
## 3                      0                          1           0
## 4                      1                          0           0
## 5                      1                          0           0
##   rent_or_own.Own rent_or_own.Rent employment_status.
## 0               1                0                  0
## 1               0                1                  0
## 3               0                1                  0
## 4               1                0                  0
## 5               1                0                  0
##   employment_status.Employed employment_status.Not in Labor Force
## 0                          0                                    1
## 1                          1                                    0
## 3                          0                                    1
## 4                          1                                    0
## 5                          1                                    0
##   employment_status.Unemployed census_msa.MSA, Not Principle  City
## 0                            0                                  0
## 1                            0                                  1
## 3                            0                                  0
## 4                            0                                  1
## 5                            0                                  0
##   census_msa.MSA, Principle City census_msa.Non-MSA household_adults
## 0                             0                  1                 0
## 1                             0                  0                 0
## 3                             1                  0                 0
## 4                             0                  0                 1
## 5                             1                  0                 2
```

```
##   household_children h1n1_vaccine
## 0                  0           no
## 1                  0           no
## 3                  0           no
## 4                  0           no
## 5                  3           no
```

DATA PREPROCESSING

Not needed since we are using random forest to train the model. And random forest is sensitive to feature scaling (but can perform normalization).

TRAINING THE MODEL

```
#ensuring reproducibility
set.seed(123)

#applying training algorithms
lg_model <- glm(h1n1_vaccine~., data = new_data2, family = binomial)
```

```
#summary(lg_model)
```

TEST DATA

Applying on actual test data. ## LOADING THE DATA

```
#load the validation labels
testing_data <- read.csv("test_set_features.csv", row.names = 1, header = T, stringsAsFactors = T)
head(testing_data,5)
```

```
##       h1n1_concern h1n1_knowledge behavioral_antiviral_meds
## 26707            2              2                         0
## 26708            1              1                         0
## 26709            2              2                         0
## 26710            1              1                         0
## 26711            3              1                         1
##       behavioral_avoidance behavioral_face_mask behavioral_wash_hands
## 26707                    1                    0                     1
## 26708                    0                    0                     0
## 26709                    0                    1                     1
## 26710                    0                    0                     0
## 26711                    1                    0                     1
##       behavioral_large_gatherings behavioral_outside_home behavioral_touch_face
## 26707                           1                       0                     1
## 26708                           0                       0                     0
## 26709                           1                       1                     1
## 26710                           0                       0                     0
## 26711                           1                       1                     1
##       doctor_recc_h1n1 doctor_recc_seasonal chronic_med_condition
## 26707                0                    0                     0
## 26708                0                    0                     0
## 26709                0                    0                     0
## 26710                1                    1                     1
```

```
## 26711               0                   0                   0
##       child_under_6_months health_worker health_insurance
## 26707                    0             0                1
## 26708                    0             0                0
## 26709                    0             0               NA
## 26710                    0             0                1
## 26711                    0             1                1
##       opinion_h1n1_vacc_effective opinion_h1n1_risk opinion_h1n1_sick_from_vacc
## 26707                           5                1                           1
## 26708                           4                1                           1
## 26709                           5                4                           2
## 26710                           4                2                           2
## 26711                           5                2                           4
##       opinion_seas_vacc_effective opinion_seas_risk opinion_seas_sick_from_vacc
## 26707                           5                1                           1
## 26708                           4                1                           1
## 26709                           5                4                           4
## 26710                           4                4                           2
## 26711                           4                4                           2
##          age_group        education     race    sex            income_poverty
## 26707 35 - 44 Years College Graduate Hispanic Female                > $75,000
## 26708 18 - 34 Years         12 Years    White   Male              Below Poverty
## 26709 55 - 64 Years College Graduate    White   Male                 > $75,000
## 26710     65+ Years         12 Years    White Female <= $75,000, Above Poverty
## 26711 35 - 44 Years         12 Years    Black Female <= $75,000, Above Poverty
##       marital_status rent_or_own  employment_status hhs_geo_region
## 26707    Not Married        Rent           Employed       mlyzmhmf
## 26708    Not Married        Rent           Employed       bhuqouqj
## 26709        Married         Own           Employed       lrircsnp
## 26710        Married         Own Not in Labor Force       lrircsnp
## 26711    Not Married         Own           Employed       lzgpxyit
##                      census_msa household_adults household_children
## 26707 MSA, Not Principle  City                1                  0
## 26708                  Non-MSA                3                  0
## 26709                  Non-MSA                1                  0
## 26710 MSA, Not Principle  City                1                  0
## 26711                  Non-MSA                0                  1
##       employment_industry employment_occupation
## 26707            atmlpfrs              hfxkjkmi
## 26708            atmlpfrs              xqwwgdyp
## 26709            nduyfdeo              pvmttkik
## 26710
## 26711            fcxhlnwr              mxkfnird
```

HANDLING MISSING VALUES

```
colSums(is.na(testing_data))
```

```
##               h1n1_concern             h1n1_knowledge
##                         85                        122
##   behavioral_antiviral_meds       behavioral_avoidance
##                         79                        213
##        behavioral_face_mask      behavioral_wash_hands
##                         19                         40
```

```
## behavioral_large_gatherings     behavioral_outside_home
##                           72                         82
##       behavioral_touch_face            doctor_recc_h1n1
##                          128                       2160
##        doctor_recc_seasonal       chronic_med_condition
##                        2160                        932
##        child_under_6_months               health_worker
##                         813                        789
##            health_insurance opinion_h1n1_vacc_effective
##                       12228                        398
##           opinion_h1n1_risk opinion_h1n1_sick_from_vacc
##                         380                        375
## opinion_seas_vacc_effective           opinion_seas_risk
##                         452                        499
## opinion_seas_sick_from_vacc                  age_group
##                         521                          0
##                   education                       race
##                           0                          0
##                         sex              income_poverty
##                           0                          0
##              marital_status                 rent_or_own
##                           0                          0
##           employment_status              hhs_geo_region
##                           0                          0
##                  census_msa            household_adults
##                           0                        225
##          household_children         employment_industry
##                         225                          0
##       employment_occupation
##                           0
```

```
test_data <- testing_data
#removing "health_insurance" column from the dataset
test_data$health_insurance <- NULL

#also dropping "hhs_goe_region", "employment_industry" and "employment_occupation"
test_data$hhs_geo_region <- NULL
test_data$employment_industry <- NULL
test_data$employment_occupation <- NULL
```

```
#dropping all NA values in the dataset
test_data <- na.omit(test_data)

#checking the dimension of NA after dropping all the values
dim(test_data)
```

```
## [1] 22971    31
```

22971 out of 26708(a difference of 3737).

CONVERTING NOMINAL VALUES TO NUMERIC VALUES

```
#binarising the nominal attributes
binary_data2 <- dummyVars(~., data = test_data)
```

```r
#View(binary_data)

#adding the conversion to the data
test_data2 <- predict(binary_data2, newdata = test_data)
head(test_data2,5)
```

```
##       h1n1_concern h1n1_knowledge behavioral_antiviral_meds
## 26707            2              2                         0
## 26708            1              1                         0
## 26709            2              2                         0
## 26710            1              1                         0
## 26711            3              1                         1
##       behavioral_avoidance behavioral_face_mask behavioral_wash_hands
## 26707                    1                    0                     1
## 26708                    0                    0                     0
## 26709                    0                    1                     1
## 26710                    0                    0                     0
## 26711                    1                    0                     1
##       behavioral_large_gatherings behavioral_outside_home behavioral_touch_face
## 26707                           1                       0                     1
## 26708                           0                       0                     0
## 26709                           1                       1                     1
## 26710                           0                       0                     0
## 26711                           1                       1                     1
##       doctor_recc_h1n1 doctor_recc_seasonal chronic_med_condition
## 26707                0                    0                     0
## 26708                0                    0                     0
## 26709                0                    0                     0
## 26710                1                    1                     1
## 26711                0                    0                     0
##       child_under_6_months health_worker opinion_h1n1_vacc_effective
## 26707                    0             0                          5
## 26708                    0             0                          4
## 26709                    0             0                          5
## 26710                    0             0                          4
## 26711                    0             1                          5
##       opinion_h1n1_risk opinion_h1n1_sick_from_vacc opinion_seas_vacc_effective
## 26707                 1                           1                          5
## 26708                 1                           1                          4
## 26709                 4                           2                          5
## 26710                 2                           2                          4
## 26711                 2                           4                          4
##       opinion_seas_risk opinion_seas_sick_from_vacc age_group.18 - 34 Years
## 26707                 1                           1                      0
## 26708                 1                           1                      1
## 26709                 4                           4                      0
## 26710                 4                           2                      0
## 26711                 4                           2                      0
##       age_group.35 - 44 Years age_group.45 - 54 Years age_group.55 - 64 Years
## 26707                       1                       0                       0
## 26708                       0                       0                       0
## 26709                       0                       0                       1
## 26710                       0                       0                       0
## 26711                       1                       0                       0
```

```
##       age_group.65+ Years education. education.< 12 Years education.12 Years
## 26707                   0          0                     0                   0
## 26708                   0          0                     0                   1
## 26709                   0          0                     0                   0
## 26710                   1          0                     0                   1
## 26711                   0          0                     0                   1
##       education.College Graduate education.Some College race.Black
## 26707                          1                     0          0
## 26708                          0                     0          0
## 26709                          1                     0          0
## 26710                          0                     0          0
## 26711                          0                     0          1
##       race.Hispanic race.Other or Multiple race.White sex.Female sex.Male
## 26707             1                     0          0          1        0
## 26708             0                     0          1          0        1
## 26709             0                     0          1          0        1
## 26710             0                     0          1          1        0
## 26711             0                     0          0          1        0
##       income_poverty. income_poverty.<= $75,000, Above Poverty
## 26707               0                                       0
## 26708               0                                       0
## 26709               0                                       0
## 26710               0                                       1
## 26711               0                                       1
##       income_poverty.> $75,000 income_poverty.Below Poverty marital_status.
## 26707                        1                            0               0
## 26708                        0                            1               0
## 26709                        1                            0               0
## 26710                        0                            0               0
## 26711                        0                            0               0
##       marital_status.Married marital_status.Not Married rent_or_own.
## 26707                      0                          1           0
## 26708                      0                          1           0
## 26709                      1                          0           0
## 26710                      1                          0           0
## 26711                      0                          1           0
##       rent_or_own.Own rent_or_own.Rent employment_status.
## 26707               0               1                  0
## 26708               0               1                  0
## 26709               1               0                  0
## 26710               1               0                  0
## 26711               1               0                  0
##       employment_status.Employed employment_status.Not in Labor Force
## 26707                          1                                     0
## 26708                          1                                     0
## 26709                          1                                     0
## 26710                          0                                     1
## 26711                          1                                     0
##       employment_status.Unemployed census_msa.MSA, Not Principle  City
## 26707                            0                                  1
## 26708                            0                                  0
## 26709                            0                                  0
## 26710                            0                                  1
## 26711                            0                                  0
```

```
##       census_msa.MSA, Principle City census_msa.Non-MSA household_adults
## 26707                             0                 0                1
## 26708                             0                 1                3
## 26709                             0                 1                1
## 26710                             0                 0                1
## 26711                             0                 1                0
##       household_children
## 26707                  0
## 26708                  0
## 26709                  0
## 26710                  0
## 26711                  1
```

VALIDATING THE MODEL

```
#using dummyVars changes it to matrix so convert it back to a data.frame
test_data2 <- as.data.frame(test_data2)

#Predict on test data
testing <- predict(lg_model, test_data2, type = "response")

#testing
```

```
#adding the probability to my test_set_features
test_data$h1n1_vaccine <- testing
head(test_data,5)
```

```
##       h1n1_concern h1n1_knowledge behavioral_antiviral_meds
## 26707            2              2                         0
## 26708            1              1                         0
## 26709            2              2                         0
## 26710            1              1                         0
## 26711            3              1                         1
##       behavioral_avoidance behavioral_face_mask behavioral_wash_hands
## 26707                    1                    0                    1
## 26708                    0                    0                    0
## 26709                    0                    1                    1
## 26710                    0                    0                    0
## 26711                    1                    0                    1
##       behavioral_large_gatherings behavioral_outside_home behavioral_touch_face
## 26707                           1                       0                     1
## 26708                           0                       0                     0
## 26709                           1                       1                     1
## 26710                           0                       0                     0
## 26711                           1                       1                     1
##       doctor_recc_h1n1 doctor_recc_seasonal chronic_med_condition
## 26707                0                    0                    0
## 26708                0                    0                    0
## 26709                0                    0                    0
## 26710                1                    1                    1
## 26711                0                    0                    0
##       child_under_6_months health_worker opinion_h1n1_vacc_effective
## 26707                    0             0                          5
## 26708                    0             0                          4
```

```
## 26709                   0            0                         5
## 26710                   0            0                         4
## 26711                   0            1                         5
##       opinion_h1n1_risk opinion_h1n1_sick_from_vacc opinion_seas_vacc_effective
## 26707                 1                           1                          5
## 26708                 1                           1                          4
## 26709                 4                           2                          5
## 26710                 2                           2                          4
## 26711                 2                           4                          4
##       opinion_seas_risk opinion_seas_sick_from_vacc     age_group
## 26707                 1                           1 35 - 44 Years
## 26708                 1                           1 18 - 34 Years
## 26709                 4                           4 55 - 64 Years
## 26710                 4                           2     65+ Years
## 26711                 4                           2 35 - 44 Years
##              education     race    sex            income_poverty marital_status
## 26707 College Graduate Hispanic Female                 > $75,000    Not Married
## 26708         12 Years    White   Male              Below Poverty    Not Married
## 26709 College Graduate    White   Male                 > $75,000        Married
## 26710         12 Years    White Female <= $75,000, Above Poverty        Married
## 26711         12 Years    Black Female <= $75,000, Above Poverty    Not Married
##       rent_or_own  employment_status               census_msa household_adults
## 26707        Rent           Employed MSA, Not Principle  City                1
## 26708        Rent           Employed                   Non-MSA                3
## 26709         Own           Employed                   Non-MSA                1
## 26710         Own Not in Labor Force MSA, Not Principle  City                1
## 26711         Own           Employed                   Non-MSA                0
##       household_children h1n1_vaccine
## 26707                  0   0.07864108
## 26708                  0   0.05170301
## 26709                  0   0.48742835
## 26710                  0   0.48769171
## 26711                  1   0.19054854
```

TRAINING SEASONAL_VACCINE ALONE

```
#adding "seasonal_vaccine class" back to original dataset
new_data3 <- cbind(new_data, h1n1[33])
```

```
head(new_data3,5)
```

```
##   h1n1_concern h1n1_knowledge behavioral_antiviral_meds behavioral_avoidance
## 0            1              0                         0                    0
## 1            3              2                         0                    1
## 3            1              1                         0                    1
## 4            2              1                         0                    1
## 5            3              1                         0                    1
##   behavioral_face_mask behavioral_wash_hands behavioral_large_gatherings
## 0                    0                     0                           0
## 1                    0                     1                           0
## 3                    0                     1                           1
## 4                    0                     1                           1
## 5                    0                     1                           0
##   behavioral_outside_home behavioral_touch_face doctor_recc_h1n1
```

```
## 0                      1                    1               0
## 1                      1                    1               0
## 3                      0                    0               0
## 4                      0                    1               0
## 5                      0                    1               0
##   doctor_recc_seasonal chronic_med_condition child_under_6_months health_worker
## 0                    0                    0                     0             0
## 1                    0                    0                     0             0
## 3                    1                    1                     0             0
## 4                    0                    0                     0             0
## 5                    1                    0                     0             0
##   opinion_h1n1_vacc_effective opinion_h1n1_risk opinion_h1n1_sick_from_vacc
## 0                           3                1                           2
## 1                           5                4                           4
## 3                           3                3                           5
## 4                           3                3                           2
## 5                           5                2                           1
##   opinion_seas_vacc_effective opinion_seas_risk opinion_seas_sick_from_vacc
## 0                           2                1                           2
## 1                           4                2                           4
## 3                           5                4                           1
## 4                           3                1                           4
## 5                           5                4                           4
##   age_group.18 - 34 Years age_group.35 - 44 Years age_group.45 - 54 Years
## 0                       0                       0                       0
## 1                       0                       1                       0
## 3                       0                       0                       0
## 4                       0                       0                       1
## 5                       0                       0                       0
##   age_group.55 - 64 Years age_group.65+ Years education. education.< 12 Years
## 0                       1                   0          0                    1
## 1                       0                   0          0                    0
## 3                       0                   1          0                    0
## 4                       0                   0          0                    0
## 5                       0                   1          0                    0
##   education.12 Years education.College Graduate education.Some College
## 0                  0                          0                      0
## 1                  1                          0                      0
## 3                  1                          0                      0
## 4                  0                          0                      1
## 5                  1                          0                      0
##   race.Black race.Hispanic race.Other or Multiple race.White sex.Female
## 0          0             0                       0          1           1
## 1          0             0                       0          1           0
## 3          0             0                       0          1           1
## 4          0             0                       0          1           1
## 5          0             0                       0          1           0
##   sex.Male income_poverty. income_poverty.<= $75,000, Above Poverty
## 0        0                0                                      0
## 1        1                0                                      0
## 3        0                0                                      0
## 4        0                0                                      1
## 5        1                0                                      1
##   income_poverty.> $75,000 income_poverty.Below Poverty marital_status.
```

```
## 0                       0                           1              0
## 1                       0                           1              0
## 3                       0                           1              0
## 4                       0                           0              0
## 5                       0                           0              0
##   marital_status.Married marital_status.Not Married rent_or_own.
## 0                      0                          1           0
## 1                      0                          1           0
## 3                      0                          1           0
## 4                      1                          0           0
## 5                      1                          0           0
##   rent_or_own.Own rent_or_own.Rent employment_status.
## 0               1               0                  0
## 1               0               1                  0
## 3               0               1                  0
## 4               1               0                  0
## 5               1               0                  0
##   employment_status.Employed employment_status.Not in Labor Force
## 0                          0                                    1
## 1                          1                                    0
## 3                          0                                    1
## 4                          1                                    0
## 5                          1                                    0
##   employment_status.Unemployed census_msa.MSA, Not Principle  City
## 0                            0                                  0
## 1                            0                                  1
## 3                            0                                  0
## 4                            0                                  1
## 5                            0                                  0
##   census_msa.MSA, Principle City census_msa.Non-MSA household_adults
## 0                              0                  1                0
## 1                              0                  0                0
## 3                              1                  0                0
## 4                              0                  0                1
## 5                              1                  0                2
##   household_children seasonal_vaccine
## 0                  0                0
## 1                  0                1
## 3                  0                1
## 4                  0                0
## 5                  3                0
```

```
#converting 0 and 1 to "yes" and "no" for the decision class
new_data3$seasonal_vaccine <- factor(new_data3$seasonal_vaccine, levels = c(0, 1), labels = c("no", "yes
head(new_data3,5)
```

```
##   h1n1_concern h1n1_knowledge behavioral_antiviral_meds behavioral_avoidance
## 0            1              0                         0                    0
## 1            3              2                         0                    1
## 3            1              1                         0                    1
## 4            2              1                         0                    1
## 5            3              1                         0                    1
##   behavioral_face_mask behavioral_wash_hands behavioral_large_gatherings
## 0                    0                    0                           0
## 1                    0                    1                           0
```

```
## 3                    0                    1                           1
## 4                    0                    1                           1
## 5                    0                    1                           0
##   behavioral_outside_home behavioral_touch_face doctor_recc_h1n1
## 0                       1                     1                0
## 1                       1                     1                0
## 3                       0                     0                0
## 4                       0                     1                0
## 5                       0                     1                0
##   doctor_recc_seasonal chronic_med_condition child_under_6_months health_worker
## 0                    0                     0                     0             0
## 1                    0                     0                     0             0
## 3                    1                     1                     0             0
## 4                    0                     0                     0             0
## 5                    1                     0                     0             0
##   opinion_h1n1_vacc_effective opinion_h1n1_risk opinion_h1n1_sick_from_vacc
## 0                           3                1                           2
## 1                           5                4                           4
## 3                           3                3                           5
## 4                           3                3                           2
## 5                           5                2                           1
##   opinion_seas_vacc_effective opinion_seas_risk opinion_seas_sick_from_vacc
## 0                           2                1                           2
## 1                           4                2                           4
## 3                           5                4                           1
## 4                           3                1                           4
## 5                           5                4                           4
##   age_group.18 - 34 Years age_group.35 - 44 Years age_group.45 - 54 Years
## 0                       0                       0                       0
## 1                       0                       1                       0
## 3                       0                       0                       0
## 4                       0                       0                       1
## 5                       0                       0                       0
##   age_group.55 - 64 Years age_group.65+ Years education. education.< 12 Years
## 0                       1                   0          0                    1
## 1                       0                   0          0                    0
## 3                       0                   1          0                    0
## 4                       0                   0          0                    0
## 5                       0                   1          0                    0
##   education.12 Years education.College Graduate education.Some College
## 0                  0                         0                      0
## 1                  1                         0                      0
## 3                  1                         0                      0
## 4                  0                         0                      1
## 5                  1                         0                      0
##   race.Black race.Hispanic race.Other or Multiple race.White sex.Female
## 0          0             0                       0          1          1
## 1          0             0                       0          1          0
## 3          0             0                       0          1          1
## 4          0             0                       0          1          1
## 5          0             0                       0          1          0
##   sex.Male income_poverty. income_poverty.<= $75,000, Above Poverty
## 0        0               0                                       0
## 1        1               0                                       0
```

```
## 3        0               0                                        0
## 4        0               0                                        1
## 5        1               0                                        1
##   income_poverty.> $75,000 income_poverty.Below Poverty marital_status.
## 0                        0                           1               0
## 1                        0                           1               0
## 3                        0                           1               0
## 4                        0                           0               0
## 5                        0                           0               0
##   marital_status.Married marital_status.Not Married rent_or_own.
## 0                      0                          1           0
## 1                      0                          1           0
## 3                      0                          1           0
## 4                      1                          0           0
## 5                      1                          0           0
##   rent_or_own.Own rent_or_own.Rent employment_status.
## 0               1                0                  0
## 1               0                1                  0
## 3               0                1                  0
## 4               1                0                  0
## 5               1                0                  0
##   employment_status.Employed employment_status.Not in Labor Force
## 0                          0                                    1
## 1                          1                                    0
## 3                          0                                    1
## 4                          1                                    0
## 5                          1                                    0
##   employment_status.Unemployed census_msa.MSA, Not Principle  City
## 0                            0                                  0
## 1                            0                                  1
## 3                            0                                  0
## 4                            0                                  1
## 5                            0                                  0
##   census_msa.MSA, Principle City census_msa.Non-MSA household_adults
## 0                              0                  1                 0
## 1                              0                  0                 0
## 3                              1                  0                 0
## 4                              0                  0                 1
## 5                              1                  0                 2
##   household_children seasonal_vaccine
## 0                  0              no
## 1                  0             yes
## 3                  0             yes
## 4                  0              no
## 5                  3              no
```

DATA PREPROCESSING

Not needed since we are using random forest to train the model. And random forest is sensitive to feature scaling (but can perform normalization).

TRAINING THE MODEL

```
#ensuring reproducibility
set.seed(123)

#applying training algorithms
seasonal_lg_model <- glm(seasonal_vaccine~., data = new_data3, family = binomial)
```

```
summary(seasonal_lg_model)
```

```
##
## Call:
## glm(formula = seasonal_vaccine ~ ., family = binomial, data = new_data3)
##
## Coefficients: (9 not defined because of singularities)
##                                          Estimate Std. Error z value
## (Intercept)                             -4.285777   0.153040 -28.004
## h1n1_concern                             0.019093   0.023137   0.825
## h1n1_knowledge                           0.192893   0.030653   6.293
## behavioral_antiviral_meds                0.082091   0.079360   1.034
## behavioral_avoidance                    -0.032536   0.042828  -0.760
## behavioral_face_mask                     0.011098   0.070150   0.158
## behavioral_wash_hands                    0.042377   0.051593   0.821
## behavioral_large_gatherings             -0.015794   0.044680  -0.353
## behavioral_outside_home                 -0.047249   0.045570  -1.037
## behavioral_touch_face                    0.204654   0.041174   4.970
## doctor_recc_h1n1                        -0.303603   0.053325  -5.693
## doctor_recc_seasonal                     1.427577   0.046719  30.556
## chronic_med_condition                    0.228647   0.039552   5.781
## child_under_6_months                     0.070984   0.062121   1.143
## health_worker                            0.832553   0.056631  14.701
## opinion_h1n1_vacc_effective              0.026099   0.020340   1.283
## opinion_h1n1_risk                        0.030993   0.017180   1.804
## opinion_h1n1_sick_from_vacc             -0.050605   0.015565  -3.251
## opinion_seas_vacc_effective              0.560633   0.020911  26.811
## opinion_seas_risk                        0.570878   0.015768  36.205
## opinion_seas_sick_from_vacc             -0.194909   0.015326 -12.718
## `age_group.18 - 34 Years`               -1.564474   0.065710 -23.809
## `age_group.35 - 44 Years`               -1.350439   0.070602 -19.128
## `age_group.45 - 54 Years`               -1.143455   0.060158 -19.008
## `age_group.55 - 64 Years`               -0.824732   0.054155 -15.229
## `age_group.65+ Years`                          NA         NA      NA
## education.                               0.044767   0.197979   0.226
## `education.< 12 Years`                  -0.316108   0.070336  -4.494
## `education.12 Years`                    -0.076768   0.049273  -1.558
## `education.College Graduate`             0.127639   0.043716   2.920
## `education.Some College`                       NA         NA      NA
## race.Black                              -0.297005   0.068051  -4.364
## race.Hispanic                           -0.150713   0.072612  -2.076
## `race.Other or Multiple`                 0.136957   0.072791   1.882
## race.White                                     NA         NA      NA
## sex.Female                              -0.006582   0.036069  -0.182
## sex.Male                                       NA         NA      NA
## income_poverty.                          0.217753   0.079909   2.725
```

```
## `income_poverty.<= $75,000, Above Poverty`  0.132977   0.065244   2.038
## `income_poverty.> $75,000`                  0.289906   0.074966   3.867
## `income_poverty.Below Poverty`                    NA         NA      NA
## marital_status.                            -0.010726   0.209019  -0.051
## marital_status.Married                      0.118163   0.041300   2.861
## `marital_status.Not Married`                      NA         NA      NA
## rent_or_own.                                0.224544   0.114016   1.969
## rent_or_own.Own                             0.173586   0.046733   3.714
## rent_or_own.Rent                                  NA         NA      NA
## employment_status.                          0.423719   0.195420   2.168
## employment_status.Employed                  0.221546   0.079056   2.802
## `employment_status.Not in Labor Force`      0.314434   0.081280   3.868
## employment_status.Unemployed                      NA         NA      NA
## `census_msa.MSA, Not Principle  City`       0.125698   0.041856   3.003
## `census_msa.MSA, Principle City`            0.107707   0.046512   2.316
## `census_msa.Non-MSA`                              NA         NA      NA
## household_adults                           -0.054365   0.025826  -2.105
## household_children                         -0.038971   0.022294  -1.748
##                                            Pr(>|z|)
## (Intercept)                                 < 2e-16 ***
## h1n1_concern                               0.409241
## h1n1_knowledge                             3.12e-10 ***
## behavioral_antiviral_meds                  0.300943
## behavioral_avoidance                       0.447436
## behavioral_face_mask                       0.874299
## behavioral_wash_hands                      0.411437
## behavioral_large_gatherings                0.723722
## behavioral_outside_home                    0.299806
## behavioral_touch_face                      6.68e-07 ***
## doctor_recc_h1n1                           1.25e-08 ***
## doctor_recc_seasonal                        < 2e-16 ***
## chronic_med_condition                      7.43e-09 ***
## child_under_6_months                       0.253175
## health_worker                               < 2e-16 ***
## opinion_h1n1_vacc_effective                0.199452
## opinion_h1n1_risk                          0.071235 .
## opinion_h1n1_sick_from_vacc                0.001149 **
## opinion_seas_vacc_effective                 < 2e-16 ***
## opinion_seas_risk                           < 2e-16 ***
## opinion_seas_sick_from_vacc                 < 2e-16 ***
## `age_group.18 - 34 Years`                   < 2e-16 ***
## `age_group.35 - 44 Years`                   < 2e-16 ***
## `age_group.45 - 54 Years`                   < 2e-16 ***
## `age_group.55 - 64 Years`                   < 2e-16 ***
## `age_group.65+ Years`                            NA
## education.                                 0.821107
## `education.< 12 Years`                     6.98e-06 ***
## `education.12 Years`                       0.119235
## `education.College Graduate`               0.003503 **
## `education.Some College`                         NA
## race.Black                                 1.27e-05 ***
## race.Hispanic                              0.037932 *
## `race.Other or Multiple`                   0.059902 .
## race.White                                       NA
```

```
## sex.Female                                    0.855207
## sex.Male                                           NA
## income_poverty.                               0.006430 **
## `income_poverty.<= $75,000, Above Poverty`    0.041536 *
## `income_poverty.> $75,000`                    0.000110 ***
## `income_poverty.Below Poverty`                     NA
## marital_status.                               0.959073
## marital_status.Married                        0.004222 **
## `marital_status.Not Married`                       NA
## rent_or_own.                                  0.048905 *
## rent_or_own.Own                               0.000204 ***
## rent_or_own.Rent                                   NA
## employment_status.                            0.030140 *
## employment_status.Employed                    0.005072 **
## `employment_status.Not in Labor Force`        0.000110 ***
## employment_status.Unemployed                       NA
## `census_msa.MSA, Not Principle  City`         0.002672 **
## `census_msa.MSA, Principle City`              0.020574 *
## `census_msa.Non-MSA`                               NA
## household_adults                              0.035285 *
## household_children                            0.080451 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 31807  on 22975  degrees of freedom
## Residual deviance: 21732  on 22929  degrees of freedom
## AIC: 21826
##
## Number of Fisher Scoring iterations: 5
```

VALIDATING THE MODEL (ASSITED WITH AI)

```r
#using dummyVars changes it to matrix so convert it back to a data.frame
test_data2 <- as.data.frame(test_data2)

#Predict on test data
testing2 <- predict(seasonal_lg_model, test_data2, type = "response")

#testing2
```

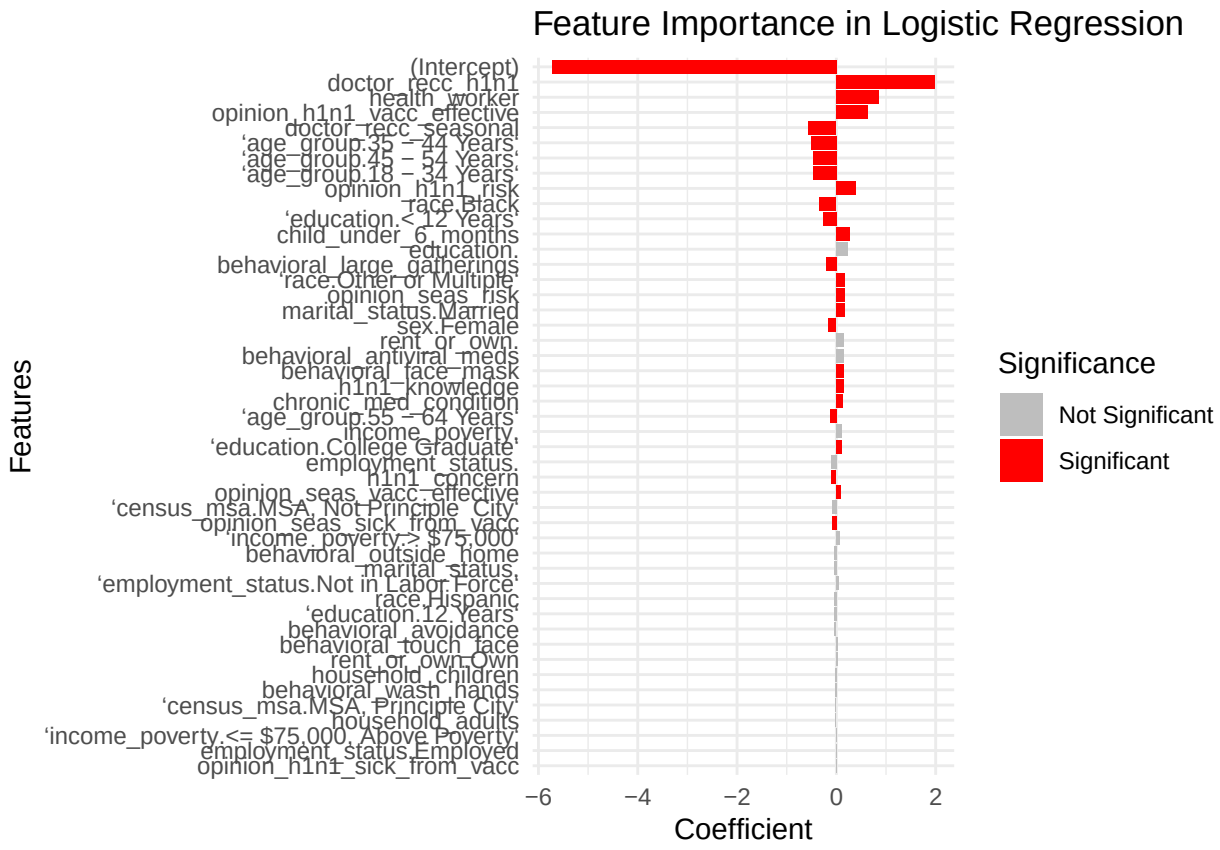
```r
#adding the probability to my test_set_features
test_data$seasonal_vaccine <- testing2
head(test_data,5)
```

```
##       h1n1_concern h1n1_knowledge behavioral_antiviral_meds
## 26707            2              2                         0
## 26708            1              1                         0
## 26709            2              2                         0
## 26710            1              1                         0
## 26711            3              1                         1
##       behavioral_avoidance behavioral_face_mask behavioral_wash_hands
## 26707                    1                    0                     1
```

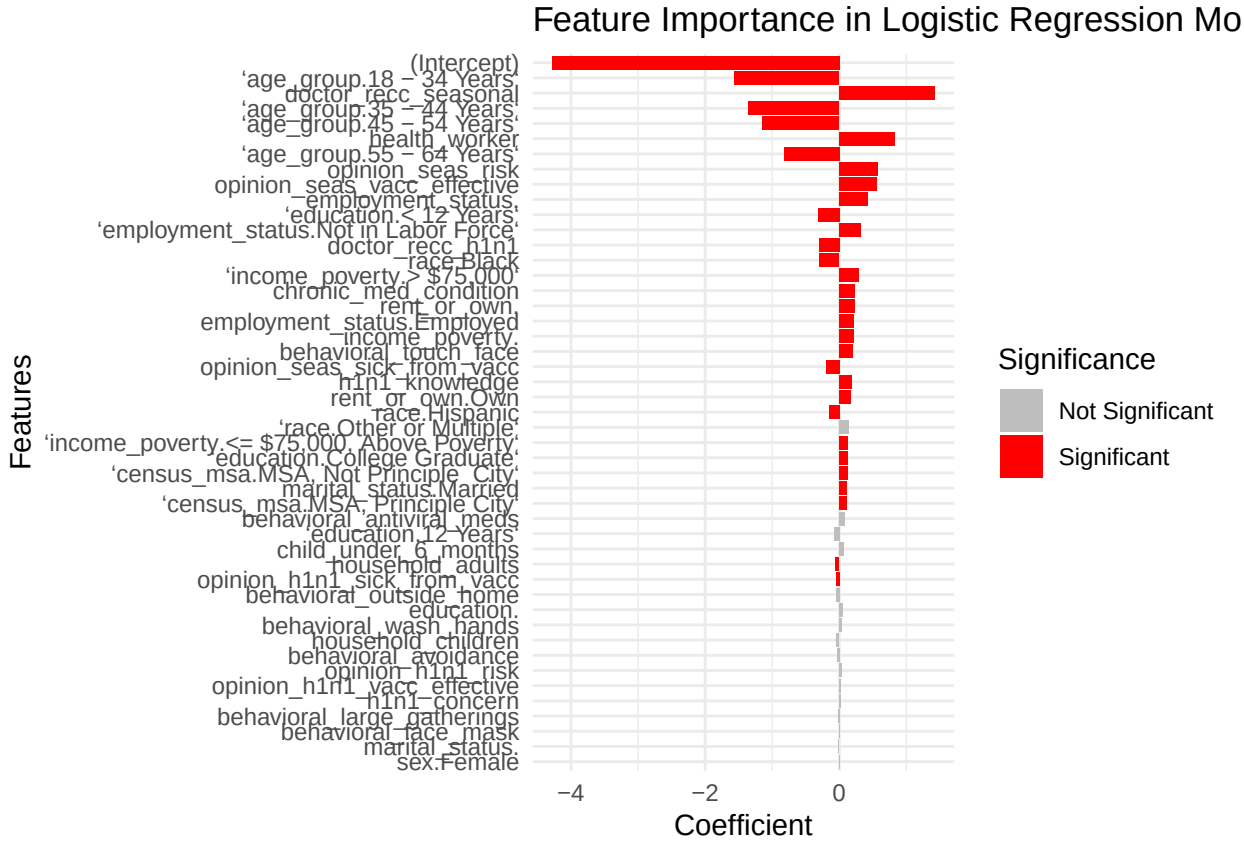
```
## 26708                    0                    0                    0
## 26709                    0                    1                    1
## 26710                    0                    0                    0
## 26711                    1                    0                    1
##       behavioral_large_gatherings behavioral_outside_home behavioral_touch_face
## 26707                           1                       0                     1
## 26708                           0                       0                     0
## 26709                           1                       1                     1
## 26710                           0                       0                     0
## 26711                           1                       1                     1
##       doctor_recc_h1n1 doctor_recc_seasonal chronic_med_condition
## 26707                0                    0                     0
## 26708                0                    0                     0
## 26709                0                    0                     0
## 26710                1                    1                     1
## 26711                0                    0                     0
##       child_under_6_months health_worker opinion_h1n1_vacc_effective
## 26707                    0             0                          5
## 26708                    0             0                          4
## 26709                    0             0                          5
## 26710                    0             0                          4
## 26711                    0             1                          5
##       opinion_h1n1_risk opinion_h1n1_sick_from_vacc opinion_seas_vacc_effective
## 26707                 1                           1                          5
## 26708                 1                           1                          4
## 26709                 4                           2                          5
## 26710                 2                           2                          4
## 26711                 2                           4                          4
##       opinion_seas_risk opinion_seas_sick_from_vacc     age_group
## 26707                 1                           1 35 - 44 Years
## 26708                 1                           1 18 - 34 Years
## 26709                 4                           4 55 - 64 Years
## 26710                 4                           2     65+ Years
## 26711                 4                           2 35 - 44 Years
##              education     race    sex            income_poverty marital_status
## 26707 College Graduate Hispanic Female                 > $75,000    Not Married
## 26708         12 Years    White   Male              Below Poverty    Not Married
## 26709 College Graduate    White   Male                 > $75,000        Married
## 26710         12 Years    White Female <= $75,000, Above Poverty        Married
## 26711         12 Years    Black Female <= $75,000, Above Poverty    Not Married
##       rent_or_own  employment_status              census_msa household_adults
## 26707        Rent           Employed MSA, Not Principle  City               1
## 26708        Rent           Employed                 Non-MSA               3
## 26709         Own           Employed                 Non-MSA               1
## 26710         Own Not in Labor Force MSA, Not Principle  City               1
## 26711         Own           Employed                 Non-MSA               0
##       household_children h1n1_vaccine seasonal_vaccine
## 26707                  0   0.07864108       0.23678376
## 26708                  0   0.05170301       0.04958104
## 26709                  0   0.48742835       0.69958555
## 26710                  0   0.48769171       0.90090987
## 26711                  1   0.19054854       0.47879411
```

```r
write.csv(test_data, "test_data.csv", row.names = TRUE)
```

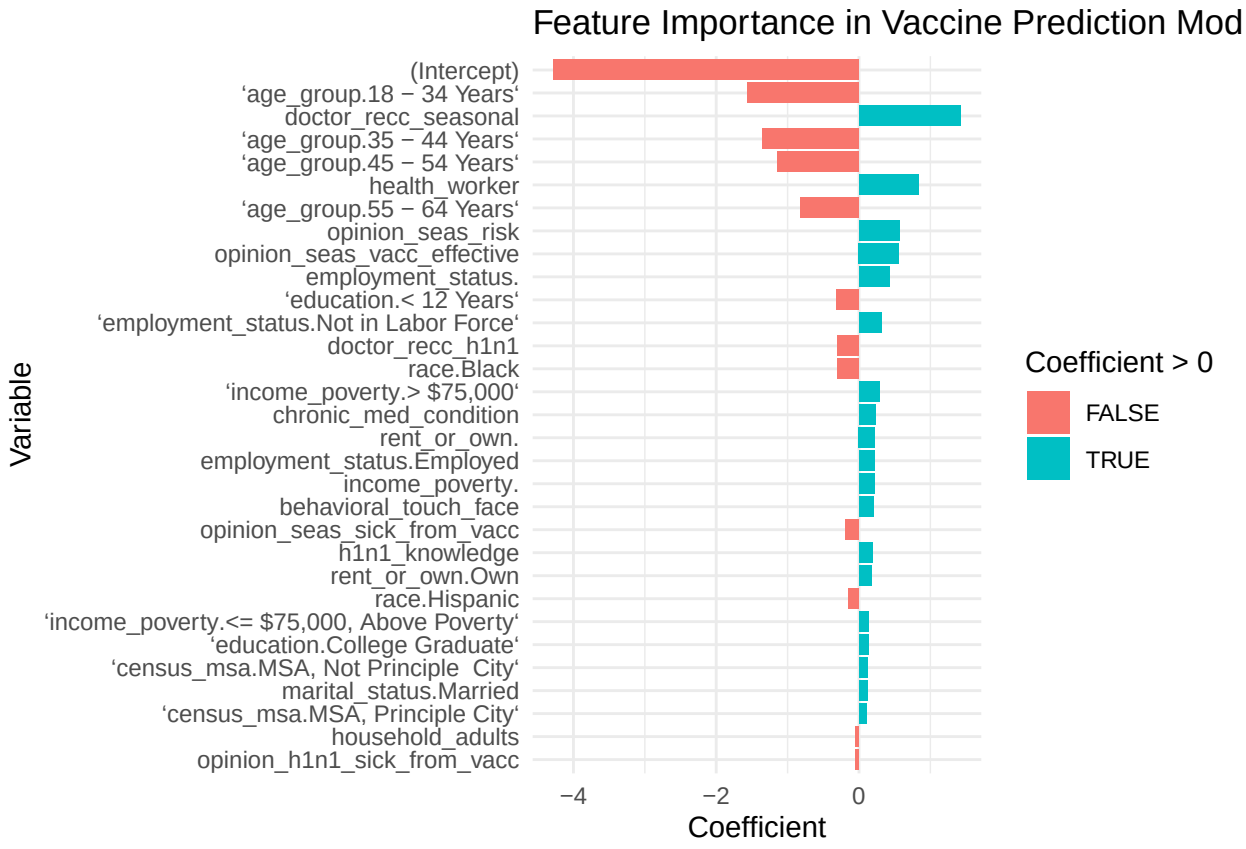
```r
# Load required library
library(ggplot2)

# Extract model summary
model_summary <- summary(lg_model)

# Create a data frame with coefficients & p-values
importance_df <- as.data.frame(model_summary$coefficients[, c("Estimate", "Pr(>|z|)")])
colnames(importance_df) <- c("Coefficient", "P_value")

# Add significance label
importance_df$Significance <- ifelse(importance_df$P_value < 0.05, "Significant", "Not Significant")

# Sort by absolute coefficient size
importance_df <- importance_df[order(abs(importance_df$Coefficient), decreasing = TRUE), ]

# Plot feature importance
ggplot(importance_df, aes(x = reorder(rownames(importance_df), abs(Coefficient)), y = Coefficient, fill
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_manual(values = c("Significant" = "red", "Not Significant" = "gray")) +
  labs(title = "Feature Importance in Logistic Regression",
       x = "Features", y = "Coefficient") +
  theme_minimal()
```



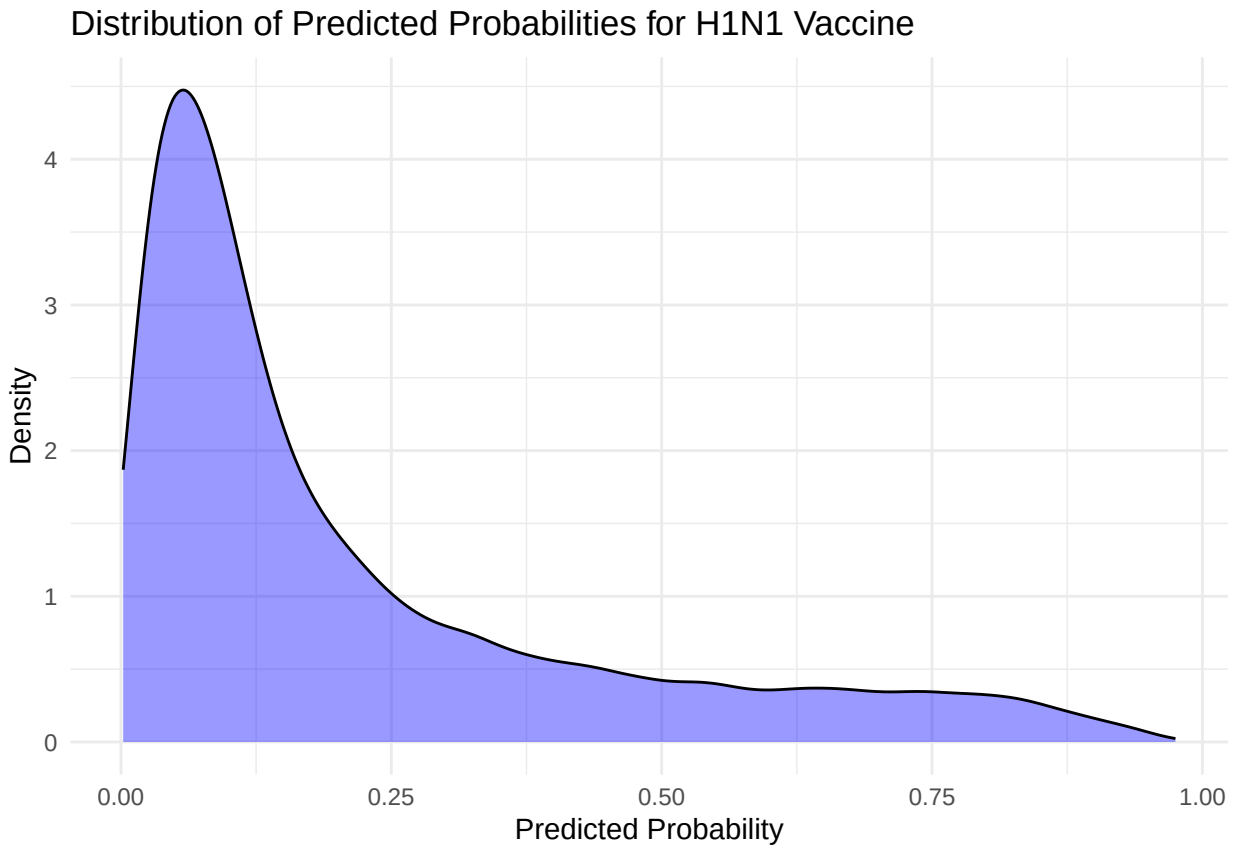
```r
# Extract coefficients and p-values
model_summary <- summary(seasonal_lg_model)
coefficients <- model_summary$coefficients[, "Estimate"]
p_values <- model_summary$coefficients[, "Pr(>|z|)"]

# Create a data frame
importance_df <- data.frame(
  Variable = names(coefficients),
  Coefficient = coefficients,
  P_value = p_values,
  Significance = ifelse(p_values < 0.05, "Significant", "Not Significant")
)

# Sort by absolute coefficient size
importance_df_sorted <- importance_df[order(abs(importance_df$Coefficient), decreasing = TRUE), ]

# Load ggplot2
library(ggplot2)

# Plot bar chart of feature importance
ggplot(importance_df_sorted, aes(x = reorder(Variable, abs(Coefficient)), y = Coefficient, fill = Signi
  geom_bar(stat = "identity", show.legend = TRUE) +
  coord_flip() +
  scale_fill_manual(values = c("Significant" = "red", "Not Significant" = "gray")) +
  labs(title = "Feature Importance in Logistic Regression Model",
       x = "Features", y = "Coefficient") +
  theme_minimal()
```



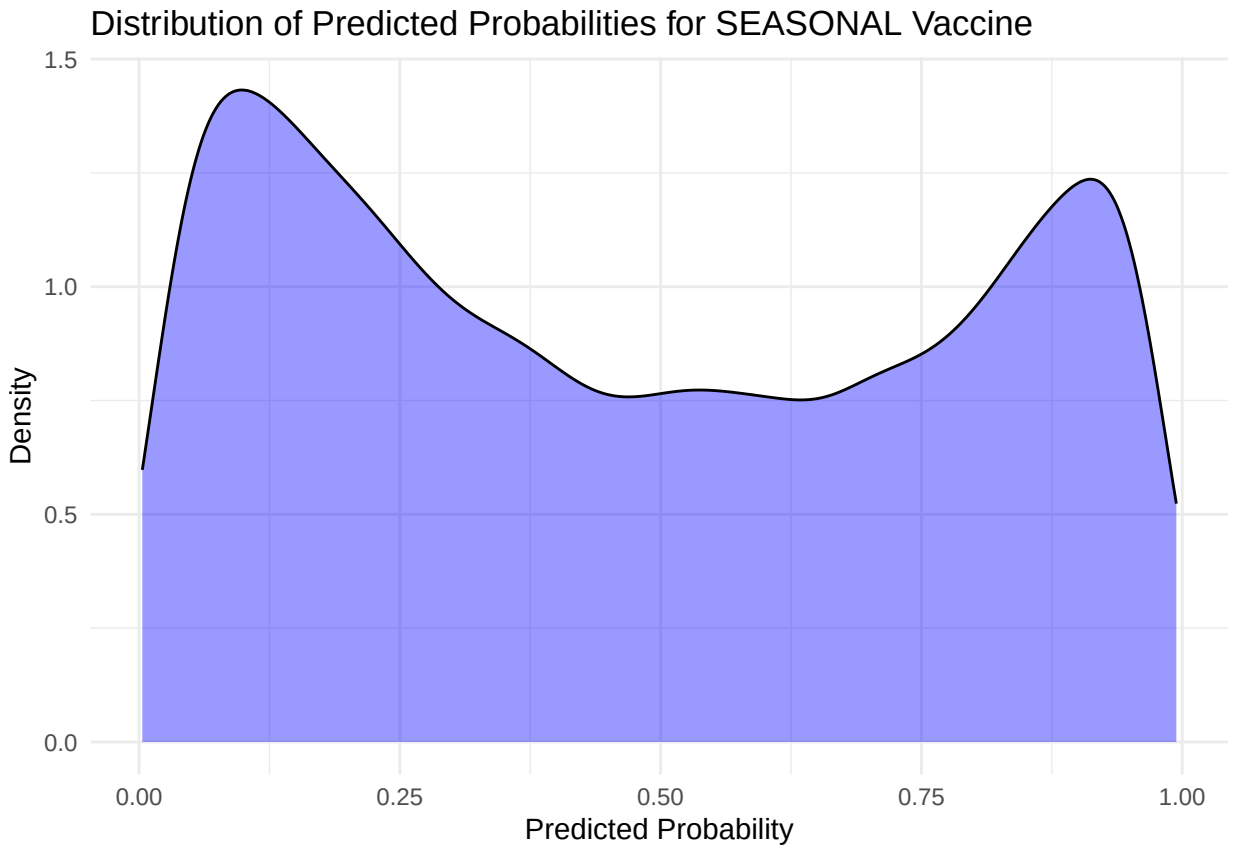
```r
library(ggplot2)

# Extract significant variables
importance_df <- data.frame(
  Variable = names(coefficients),
  Coefficient = coefficients,
  P_value = p_values
)

# Filter only significant variables (p-value < 0.05)
importance_df <- importance_df[importance_df$P_value < 0.05, ]

# Plot feature importance
ggplot(importance_df, aes(x = reorder(Variable, abs(Coefficient)), y = Coefficient, fill = Coefficient >
  geom_bar(stat = "identity") +
  coord_flip() +
  labs(title = "Feature Importance in Vaccine Prediction Model", x = "Variable", y = "Coefficient") +
  theme_minimal()
```




```r
ggplot(test_data, aes(x = h1n1_vaccine)) +
  geom_density(fill = "blue", alpha = 0.4) +
  labs(title = "Distribution of Predicted Probabilities for H1N1 Vaccine", x = "Predicted Probability",
  theme_minimal()
```




```
ggplot(test_data, aes(x = seasonal_vaccine)) +
  geom_density(fill = "blue", alpha = 0.4) +
  labs(title = "Distribution of Predicted Probabilities for SEASONAL Vaccine", x = "Predicted Probabili
  theme_minimal()
```


Distribution of Predicted Probabilities for SEASONAL Vaccine
Density
1.5
1.0
0.5
0.0
0.00
0.25
0.50
0.75
1.00
Predicted Probability