Untitled

```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 3.5.2
```

```
library(stringr)
```

```
## Warning: package 'stringr' was built under R version 3.5.2
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.5.2
```

```
library(kknn) ## knn library
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.5.2
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 3.5.2
```

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

```
##
## Attaching package: 'caret'
```

```
## The following object is masked from 'package:kknn':
##
##     contr.dummy
```

```
#Prepare US Presidents Data set

data <- read.csv('us_presidents.csv')
data$start <- as.character(data$start)

data <- data %>% separate(col="start", into =c('day','start_year'), sep=-5, extra = 'drop', remove = FA
data <- data %>% separate(col="end", into =c('day','end_year'), sep=-5, extra = 'drop', remove = FALSE)
data <- data %>% separate(col="president", into =c("first", "middle", 'middle2', 'last'), sep=" ", extra
data$start_year <- as.numeric(data$start_year)
data$end_year <- as.numeric(data$end_year)
```

```
## Warning: NAs introduced by coercion
```

```
data$day <- NULL


data$second_term <- ((data$end_year - data$start_year) > 4)
```

```r
write.csv(data, "us_presidents_cleaned.csv")
```

```r
# Create speeches csv

files <- list.files("./state-of-the-union-corpus-1989-2017")

files_split <- read.table(text = files, sep = "_")

files_split <- files_split %>% separate(col="V2", into ="year", sep='.txt', extra = 'drop', remove = TRU

speeches = 1:length(files)

for(i in speeches) {
  fileName <- paste('./state-of-the-union-corpus-1989-2017/',files[i], sep="")
  speeches <- readChar(fileName, file.info(fileName)$size)
}

data <- data.frame("year" = files_split$year, "name" = files_split$V1, "text" = speeches)

write.csv(data, "speeches.csv")
```

```r
# Add lemmatized data to speeches csv

files <- list.files("./state-of-the-union-lemmatized")

speeches = 1:length(files)

for(i in speeches) {
  fileName <- paste('./state-of-the-union-lemmatized/',files[i], sep="")
  lemmatizedText <- readChar(fileName, file.info(fileName)$size)
}

data$lemmatizedText <- lemmatizedText

write.csv(data, "speeches.csv")
```

```r
# Add topics to speeches csv

presidents <- read.csv('us_presidents_cleaned.csv')
data <- read.csv('speeches.csv')
extra_data <- read.csv('extra_data.csv')

files <- list.files("./topics")

fileLen = 1:length(files)
topicsBlob <- vector(mode = "list", length = length(files))

for(i in fileLen) {
  fileName <- paste('./topics/',files[i], sep="")
  topicsBlob[i] <- readChar(fileName, file.info(fileName)$size)
}

data$topicsBlob <- topicsBlob
```

```r
# separate into topics column

data <- data %>%
  separate(topicsBlob, into=c("topic0","topic1", "topic2", "topic3", "topic4", "topic5", "topic6", "topi
```

```
## Warning: Expected 12 pieces. Missing pieces filled with `NA` in 228
## rows [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19,
## 20, ...].
```

```r
data$topic0 = NULL
data$topic11 = NULL

# do some preliminary analysis of topics and encode variables

data$war <- if_else(grepl("war", data$topicsBlob), 1, 0)
data$warFirstTopic <- if_else(grepl("war", data$topic1), 1, 0)

data$nation <- if_else(grepl("nation|national|nationality", data$topicsBlob), 1, 0)
data$nationFirstTopic <- if_else(grepl("nation|national|nationality", data$topic1), 1, 0)

data$debt <- if_else(grepl("debt", data$topicsBlob), 1, 0)
data$debtFirstTopic <- if_else(grepl("debt", data$topic1), 1, 0)

data$economy <- if_else(grepl("debt|economy|fiscal|expenditures|appropriations|commerce|market|produce"
data$economyFirstTopic <- if_else(grepl("debt|economy|fiscal|expenditures|appropriations|commerce|marke

data$fiscal <- if_else(grepl("debt|fiscal|floating|commerce|treasury|receipts|money|loan|budget|taxes",
data$fiscalFirstTopic <- if_else(grepl("debt|fiscal|floating|commerce|treasury|receipts|money|loan|budg

data$law <- if_else(grepl("law|order|right|laws|legislation", data$topicsBlob), 1, 0)
data$lawFirstTopic <- if_else(grepl("law|order|right|laws|legislation", data$topic1), 1, 0)

data$railroads <- if_else(grepl("railroad|railroads", data$topicsBlob), 1, 0)
data$interstate <- if_else(grepl("interstate", data$topicsBlob), 1, 0)
data$medical <- if_else(grepl("medical", data$topicsBlob), 1, 0)
data$world <- if_else(grepl("world", data$topicsBlob), 1, 0)
data$taxes <- if_else(grepl("taxes", data$topicsBlob), 1, 0)
data$armenia <- if_else(grepl("armenia", data$topicsBlob), 1, 0)
data$islands <- if_else(grepl("islands|island", data$topicsBlob), 1, 0)
data$gentlemen <- if_else(grepl("gentleme|men", data$topicsBlob), 1, 0)
data$ladies <- if_else(grepl("ladies|women", data$topicsBlob), 1, 0)
data$powerful <- if_else(grepl("force|power|action|act|make", data$topicsBlob), 1, 0)
data$believe <- if_else(grepl("believe", data$topicsBlob), 1, 0)
data$public <- if_else(grepl("public", data$topicsBlob), 1, 0)
data$private <- if_else(grepl("private", data$topicsBlob), 1, 0)
data$duty <- if_else(grepl("duty", data$topicsBlob), 1, 0)
data$storage <- if_else(grepl("storage", data$topicsBlob), 1, 0)

#these are too big and their job is done; take them out

data$lemmatizedText <- NULL
data$text <- NULL

# merge in us presidents data
```

```r
data <- merge(data, presidents, by.x = "name", by.y = "last", all.x=TRUE)

# merge in other data

data <- merge(data, extra_data, by.x = "year", by.y = "Year", all.x=TRUE)


data <- apply(data,2,as.character)

write.csv(data, "speeches.csv")
```

```r
data <- read.csv('topicblob_vectorized.csv')

data <- data %>% drop_na(party)

num_observations <- nrow(data)

majority = num_observations * .8

rows <- sample(nrow(data), majority)

train <- data[rows, ]
test <- data[-rows, ]

result <- kknn(party ~ topicsBlob, train, test, k=3, kernel = "rectangular")

predictions <- predict(result, newdata=test)

errors <- 1:length(test$party)

errors$answer <- test$party
```

```
## Warning in errors$answer <- test$party: Coercing LHS to a list
```

```r
errors$prediction <- predictions

errors$match <- as.numeric(if_else(errors$answer == errors$prediction, 1, 0))

print(sum(errors$match)/length(errors))
```

```
## [1] 0.3061224
```

```r
print(predictions)
```

```
##  [1] Republican Republican Republican Republican Republican Republican
##  [7] Republican Republican Republican Republican Republican Republican
## [13] Republican Republican Republican Republican Republican Republican
## [19] Republican Republican Republican Republican Republican Republican
## [25] Republican Republican Republican Republican Republican Republican
## [31] Republican Republican Republican Republican Republican Republican
## [37] Republican Republican Republican Republican Republican Republican
## [43] Republican Republican Republican Republican
## 9 Levels: Democratic Democratic-  Republican ... Whig   April 4, 1841  -  September 13, 1841
```

```r
immigration <- read.csv('immigration.csv')
war <- read.csv('war.csv')
economy <- read.csv('economy.csv')
interesting <- read.csv('interesting.csv')

data <- merge(immigration, war, by.x = "year", by.y = "year", all.x=TRUE)
data <- merge(data, economy, by.x = "year", by.y = "year", all.x=TRUE)
data <- merge(data, interesting, by.x = "year", by.y = "year", all.x=TRUE)
```

```
## Warning in merge.data.frame(data, interesting, by.x = "year", by.y
## = "year", : column names 'president.x', 'party.x', 'speech_total.x',
## 'president.y', 'party.y', 'speech_total.y' are duplicated in the result
```

```r
data$president.x <- NULL
data$president.y <- NULL
data$party.x <- NULL
data$party.y <- NULL

data$president.x <- NULL
data$party.x <- NULL

data$speech_total.x <- NULL
data$speech_total.y <- NULL
data$speech_total.x <- NULL
data$speech_total.y <- NULL

names(data)[names(data) == 'president.y'] <- 'president'
names(data)[names(data) == 'party.y'] <- 'party'

write.csv(data, "combine_topic_words.csv")
```

```r
#data2 <- data %>% filter(year >= 1866)
#data2$party[data2$party == "National Union   [i]   ( Democratic )   [j]"] <- "Democratic"

num_observations <- nrow(data)

majority = num_observations * .8

rows <- sample(nrow(data), majority)

train <- data[rows, ]
test <- data[-rows, ]

fit <- glm(party ~ ., train, family='binomial')
```

```
## Warning: glm.fit: algorithm did not converge
```

```r
print(summary(fit))
```

```
##
## Call:
## glm(formula = party ~ ., family = "binomial", data = train)
##
## Deviance Residuals:
##    [1]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
```

```
##  [24]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##  [47]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##  [70]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##  [93]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## [116]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## [139]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
## [162]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
##
## Coefficients: (32 not defined because of singularities)
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          2.638e+03  7.248e+07       0        1
## year                -1.459e+00  4.050e+04       0        1
## asylum               7.429e+01  3.297e+06       0        1
## border               2.711e+01  9.772e+05       0        1
## borders             -2.916e+01  7.370e+05       0        1
## canada               1.787e+01  6.853e+05       0        1
## citizen             -4.843e+01  1.789e+06       0        1
## citizens             4.896e+01  1.887e+06       0        1
## citizenship          9.918e+00  3.979e+05       0        1
## cross               -1.462e+01  7.204e+05       0        1
## crossing             5.970e+01  2.200e+06       0        1
## cuba                -2.441e+01  6.727e+05       0        1
## customs             -1.298e+01  4.616e+05       0        1
## deport               2.409e+00  3.499e+06       0        1
## deportation         -3.959e+02  1.465e+07       0        1
## deportations         8.797e+02  3.956e+07       0        1
## deported             1.276e+03  3.849e+07       0        1
## detain               2.589e+02  1.036e+07       0        1
## detained            -3.394e+02  1.138e+07       0        1
## documented           2.878e+02  1.276e+07       0        1
## emigration          -4.889e+01  1.911e+06       0        1
## ethnic              -4.114e+02  2.144e+07       0        1
## foreign             -3.785e+00  1.261e+05       0        1
## foreigner            5.259e+01  1.549e+06       0        1
## hispanic             5.159e+02  2.177e+07       0        1
## illegal              5.622e+01  2.173e+06       0        1
## immigrant           -6.903e+02  2.440e+07       0        1
## immigrants           5.818e+02  2.105e+07       0        1
## immigrate                   NA         NA      NA       NA
## immigration          1.093e+02  3.826e+06       0        1
## labor                6.753e+00  2.277e+05       0        1
## laborers            -4.824e+01  1.061e+06       0        1
## latin                2.707e+01  1.080e+06       0        1
## latino               1.033e+03  4.242e+07       0        1
## mexican              1.085e+01  4.156e+05       0        1
## mexico              -6.767e+00  2.542e+05       0        1
## migrant              1.232e+02  5.296e+06       0        1
## migrants            -9.411e+01  4.772e+06       0        1
## minorities          -2.933e+02  1.126e+07       0        1
## minority             1.025e+02  3.028e+06       0        1
## naturalization       2.911e+01  1.398e+06       0        1
## naturalized         -6.999e+01  2.392e+06       0        1
## racial              -1.734e+02  6.672e+06       0        1
## racism               1.733e+02  6.353e+06       0        1
```

```
## racist                    NA         NA      NA       NA
## refugee             -2.135e+01  7.411e+06       0        1
## refugees            -2.384e+02  2.102e+07       0        1
## separate             3.017e+00  4.327e+05       0        1
## separated            6.417e+01  3.157e+06       0        1
## undocumented         5.928e+02  1.580e+07       0        1
## immigration_total           NA         NA      NA       NA
## percent_immigration  2.320e+00  2.587e+06       0        1
## afghanistan          4.306e+01  1.888e+06       0        1
## afraid              -9.471e+01  5.889e+06       0        1
## army                -4.293e+00  3.692e+05       0        1
## battle               1.790e+01  6.183e+05       0        1
## combat              -5.713e+00  2.866e+06       0        1
## conflict             3.567e+01  1.284e+06       0        1
## cyber               -1.690e+02  1.419e+07       0        1
## cybersecurity               NA         NA      NA       NA
## enemies             -3.246e+01  9.929e+05       0        1
## enemy               -1.212e+01  4.930e+05       0        1
## fear                 1.811e+01  9.351e+05       0        1
## fears               -6.022e+00  2.177e+06       0        1
## fight                4.625e+01  1.226e+06       0        1
## fights              -2.767e+02  7.912e+06       0        1
## force                7.442e+00  2.302e+05       0        1
## fought               6.114e+00  1.218e+06       0        1
## genocide            -1.047e+03  4.478e+07       0        1
## iran                -1.906e+02  9.069e+06       0        1
## iraq                -1.291e+02  5.015e+06       0        1
## isis                -8.196e+01  2.955e+06       0        1
## israel               1.960e+02  1.115e+07       0        1
## jihad               -6.699e+03  2.357e+08       0        1
## korea                1.942e+01  8.730e+05       0        1
## marine              -6.960e+01  2.393e+06       0        1
## marines             -3.892e+01  1.990e+06       0        1
## militant                    NA         NA      NA       NA
## military             1.261e+01  5.465e+05       0        1
## missile              2.534e+01  1.066e+06       0        1
## missiles            -2.033e+02  1.059e+07       0        1
## navy                 1.718e+00  2.802e+05       0        1
## nuclear              6.088e+01  2.194e+06       0        1
## proxy               -8.811e+02  3.223e+07       0        1
## qaeda                2.802e+02  1.110e+07       0        1
## russia               5.863e+00  2.452e+05       0        1
## soldier              1.790e+01  1.064e+06       0        1
## soldiers             3.892e+01  1.427e+06       0        1
## soviet               8.835e+00  3.155e+05       0        1
## struggle             2.447e+01  8.510e+05       0        1
## taliban              8.127e+02  3.135e+07       0        1
## terror               6.089e+01  2.625e+06       0        1
## terrorism           -8.642e+01  3.455e+06       0        1
## terrorist            5.328e+02  1.750e+07       0        1
## terrorists          -7.561e+02  2.462e+07       0        1
## threat               2.140e+01  1.750e+06       0        1
## threaten            -2.586e+01  2.021e+06       0        1
## threatens            9.381e-01  1.258e+06       0        1
```

```
## threats              2.073e+00  1.160e+06       0        1
## trench               1.926e+02  7.811e+06       0        1
## trenches                    NA         NA      NA       NA
## troop                7.199e+02  2.982e+07       0        1
## troops              -7.074e+02  2.940e+07       0        1
## ussr                -2.606e+02  3.322e+07       0        1
## vietnam             -1.604e+01  5.456e+05       0        1
## war                 -4.999e+00  2.506e+05       0        1
## warfare              3.501e+01  1.269e+06       0        1
## wars                -2.204e+00  2.316e+05       0        1
## wartime             -5.475e+00  2.430e+06       0        1
## war_total                   NA         NA      NA       NA
## percent_war         -7.533e+01  2.089e+06       0        1
## billionaire         -1.739e+03  8.290e+07       0        1
## billionaires                NA         NA      NA       NA
## cash                -4.293e+01  2.133e+06       0        1
## class               -1.753e+01  5.326e+05       0        1
## consumer             2.278e+01  6.081e+05       0        1
## consumers            1.851e+01  1.194e+06       0        1
## crash                7.515e+01  2.983e+06       0        1
## crashes             -7.769e+01  1.931e+07       0        1
## debt                -1.315e+01  4.932e+05       0        1
## demand              -8.196e+00  3.633e+05       0        1
## downturn             3.762e+02  1.955e+07       0        1
## economic            -9.980e+00  2.873e+05       0        1
## economics           -3.420e+01  2.312e+06       0        1
## economy             -7.796e-01  4.074e+05       0        1
## employment          -3.759e+00  2.963e+05       0        1
## exchange            -2.116e+01  7.911e+05       0        1
## finance              2.520e+01  8.014e+05       0        1
## financial           -4.070e+01  1.436e+06       0        1
## gdp                         NA         NA      NA       NA
## goods               -3.079e+00  2.679e+05       0        1
## growth               3.005e+01  1.052e+06       0        1
## inflation            2.380e+01  1.061e+06       0        1
## job                 -9.874e+00  2.588e+05       0        1
## jobs                -3.832e+01  1.536e+06       0        1
## loan                 1.616e+01  8.590e+05       0        1
## loans               -3.136e+01  1.097e+06       0        1
## market               2.318e+01  8.438e+05       0        1
## markets              1.364e+01  6.981e+05       0        1
## millionaire         -1.003e+02  1.330e+07       0        1
## millionaires                NA         NA      NA       NA
## paycheck             2.529e+03  8.315e+07       0        1
## paychecks           -2.511e+03  8.177e+07       0        1
## pension              1.879e+01  7.934e+05       0        1
## pensions            -6.386e+01  2.586e+06       0        1
## pocketbook           1.789e+02  5.568e+06       0        1
## poor                -3.208e+01  1.229e+06       0        1
## poverty              1.189e+01  4.172e+05       0        1
## price                1.778e+00  2.266e+05       0        1
## prices               3.680e+00  5.745e+05       0        1
## recession           -4.966e+01  1.721e+06       0        1
## retire              -3.634e+01  1.484e+06       0        1
```

```
## retirement           4.164e+01  2.252e+06       0        1
## rich                 2.375e+00  4.280e+05       0        1
## saving               3.339e+01  9.249e+05       0        1
## stock               -7.602e+00  7.256e+05       0        1
## stocks              -3.642e+01  2.508e+06       0        1
## tax                  1.169e+01  3.564e+05       0        1
## taxes               -5.105e+01  1.657e+06       0        1
## trade                1.534e-01  6.262e+04       0        1
## unemployed          -2.630e+01  1.371e+06       0        1
## unemployment         6.783e+01  2.276e+06       0        1
## economy_total               NA         NA      NA       NA
## percent_economy      1.461e+02  5.773e+06       0        1
## presidentArthur      1.626e+02  3.695e+06       0        1
## presidentBuchanan    2.701e+02  9.878e+06       0        1
## presidentBuren       2.154e+01  4.425e+06       0        1
## presidentBush        5.655e+00  4.674e+06       0        1
## presidentCarter     -1.020e+03  4.145e+07       0        1
## presidentCleveland  -2.835e+02  8.710e+06       0        1
## presidentClinton     1.145e+03  3.868e+07       0        1
## presidentCoolidge    2.698e+02  8.603e+06       0        1
## presidentEisenhower -1.636e+02  8.107e+06       0        1
## presidentFillmore   -2.404e+02  1.097e+07       0        1
## presidentFord       -5.277e+01  9.379e+06       0        1
## presidentGrant       1.236e+02  3.207e+06       0        1
## presidentHarding     3.678e+02  1.500e+07       0        1
## presidentHarrison   -1.908e+02  1.121e+07       0        1
## presidentHayes       2.489e+02  9.189e+06       0        1
## presidentHoover      3.701e+02  8.484e+06       0        1
## presidentJackson     2.591e+02  1.520e+07       0        1
## presidentJefferson   7.124e+01  3.260e+06       0        1
## presidentJohnson    -1.921e+02  7.629e+06       0        1
## presidentKennedy    -1.048e+03  4.282e+07       0        1
## presidentLincoln     2.886e+02  9.992e+06       0        1
## presidentMadison     4.595e+01  1.697e+06       0        1
## presidentMcKinley    1.492e+03  4.553e+07       0        1
## presidentMonroe      3.139e+01  2.287e+06       0        1
## presidentNixon       2.723e+02  4.838e+06       0        1
## presidentObama      -5.592e+02  4.087e+07       0        1
## presidentPierce      1.207e+02  5.026e+06       0        1
## presidentPolk        2.155e+02  1.280e+07       0        1
## presidentReagan     -5.407e+02  1.957e+07       0        1
## presidentRoosevelt          NA         NA      NA       NA
## presidentTaft               NA         NA      NA       NA
## presidentTaylor             NA         NA      NA       NA
## presidentTruman             NA         NA      NA       NA
## presidentTrump              NA         NA      NA       NA
## presidentTyler              NA         NA      NA       NA
## presidentWashington         NA         NA      NA       NA
## presidentWilson             NA         NA      NA       NA
## against                     NA         NA      NA       NA
## american                    NA         NA      NA       NA
## change                      NA         NA      NA       NA
## child                       NA         NA      NA       NA
## crisis                      NA         NA      NA       NA
```

```
## faith                        NA         NA      NA       NA
## family                       NA         NA      NA       NA
## great                        NA         NA      NA       NA
## hope                         NA         NA      NA       NA
## together                     NA         NA      NA       NA
## unite                        NA         NA      NA       NA
## interesting_total            NA         NA      NA       NA
## percent_interesting          NA         NA      NA       NA
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2.4156e+02  on 181  degrees of freedom
## Residual deviance: 1.0559e-09  on   0  degrees of freedom
## AIC: 364
##
## Number of Fisher Scoring iterations: 25
```

```
num_observations <- nrow(data)

majority = num_observations * .8

rows <- sample(nrow(data), majority)

train <- data[rows, ]
test <- data[-rows, ]


fit <- train(party ~ ., train, method = 'svmLinear3')

print(summary(fit))
```

```
##             Length Class      Mode
## TypeDetail     1   -none-     character
## Type           1   -none-     numeric
## W           1926   -none-     numeric
## Bias           1   -none-     numeric
## ClassNames     9   factor     numeric
## NbClass        1   -none-     numeric
## xNames       213   -none-     character
## problemType    1   -none-     character
## tuneValue      2   data.frame list
## obsLevels      9   -none-     character
## param          0   -none-     list
```

```
predictions = predict(fit, test)
errors$answer <- test$party

errors$prediction <- predictions

errors$match <- as.numeric(if_else(errors$answer == errors$prediction, 1, 0))

print("Percent Correct: \n")
```

```
## [1] "Percent Correct: \n"
```

```
print(sum(errors$match)/length(errors))
```

```
## [1] 0.5510204
```

```
print(predictions)
```

```
##  [1] Democratic-  Republican Democratic-  Republican
##  [3] Democratic-  Republican Democratic-  Republican
##  [5] Democratic-  Republican Democratic-  Republican
##  [7] Democratic               Democratic
##  [9] Republican               Democratic
## [11] Democratic               Democratic
## [13] Democratic               Republican
## [15] Democratic               Whig
## [17] Democratic               Democratic
## [19] Republican               Republican
## [21] Democratic               Democratic
## [23] Democratic               Democratic
## [25] Democratic               Republican
## [27] Republican               Democratic-  Republican
## [29] Democratic-  Republican Republican
## [31] Republican               Republican
## [33] Republican               Democratic
## [35] Democratic               Republican
## [37] Democratic               Republican
## [39] Republican               Republican
## [41] Republican               Democratic
## [43] Democratic               Republican
## [45] Democratic               Republican
## 9 Levels: Democratic Democratic-  Republican ... Whig   April 4, 1841  â€"  September 13, 1841
```

```
result <- kknn(party ~ ., train, test, k=3, kernel = "rectangular")

predictions <- predict(result, newdata=test)

errors$answer <- test$party

errors$prediction <- predictions

errors$match <- as.numeric(if_else(errors$answer == errors$prediction, 1, 0))

print("Percent Correct: \n")
```

```
## [1] "Percent Correct: \n"
```

```
print(sum(errors$match)/length(errors))
```

```
## [1] 0.5714286
```

```
print(predictions)
```

```
##  [1] Nonpartisan   [13]      Democratic-  Republican
##  [3] Democratic-  Republican Democratic-  Republican
##  [5] Democratic-  Republican Democratic-  Republican
##  [7] Democratic-  Republican Democratic
##  [9] Democratic               Democratic-  Republican
## [11] Democratic               Republican
```

```
## [13] Democratic              Republican
## [15] Republican              Republican
## [17] Republican              Republican
## [19] Democratic-  Republican Republican
## [21] Republican              Democratic
## [23] Republican              Republican
## [25] Republican              Republican
## [27] Democratic-  Republican Democratic
## [29] Democratic              Democratic-  Republican
## [31] Republican              Democratic
## [33] Democratic-  Republican Republican
## [35] Democratic              Republican
## [37] Democratic              Republican
## [39] Republican              Republican
## [41] Republican              Democratic
## [43] Republican              Republican
## [45] Democratic              Republican
## 9 Levels: Democratic Democratic-  Republican ... Whig   April 4, 1841  â€"  September 13, 1841
```