Final Project

Jack Tomlon

12/3/2024

Introduction

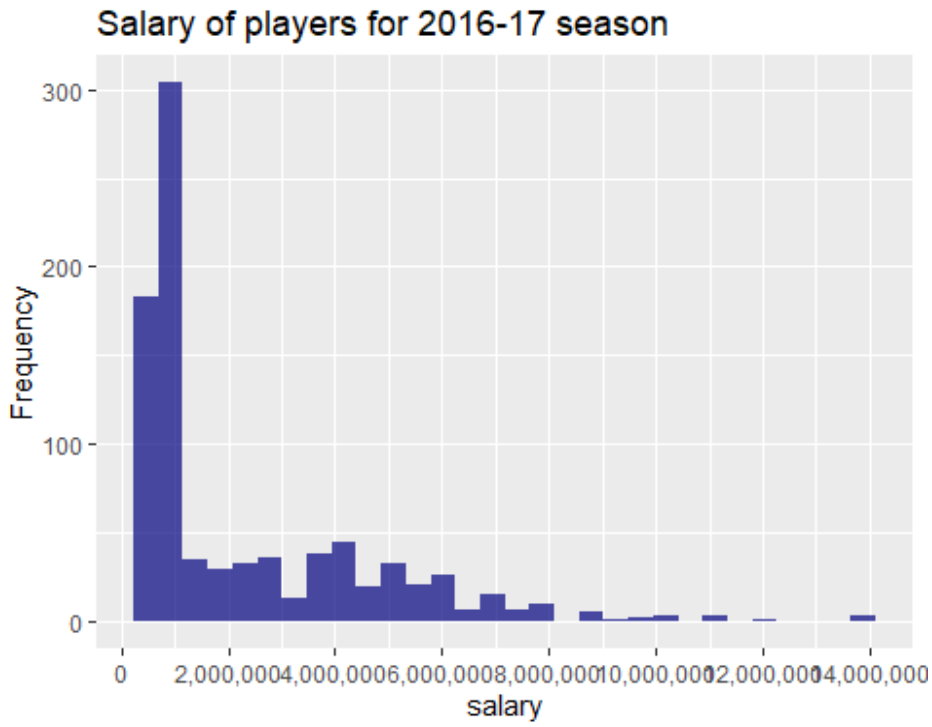
This paper will dive into the realm of NHL Hockey statistics, specifically looking at player salaries. I am a lifelong hockey player and fan which is why I chose to dive into hockey analytics. In researching this topic, I hope to gain even more experience with data science in the realm of hockey. NHL salaries are generally lower than leagues like the NBA and MLB. This is primarily due to sport popularity, less TV and advertising contracts, and salary caps for teams. Unlike baseball, NHL teams have a set amount of money that can be used to pay players. In the 2024-25 season it was 88 million dollars. Currently the highest paid player in the NHL is Leon draisaitl of the Edmonton Oilers at a eight-year 112 million dollar contract or 14 million annually, the second highest? 12.6 million for Nathan MacKinnon of the Colorado Avalanche. Comparatively the average annual NBA salary is around 12 million just shy of the NHL's top stars.

An important note: the salary cap for that season was 73.1 million dollars.

The data set chosen entitled "Predict NHL Player Salaries" Kaggle.com. It was posted 7 years ago and the link to the original raw data does not appear to be correct anymore but I cross referenced multiple observations with other sources to gauge credibility of the source and it appeared to be satisfactory for the needs of this paper. The data set contains 151 variables on ~870 observations/players. All of the stats recorded were from the 2016-17 season and the salary is for that year.

The data set gives us what we would logically think to have as predictors for salary but, it also has many other predictors that can be used for other types of predictions and models. The variable definitions can be found in the appendix. Specifically I would like to see the variables: Born, City, Pr/St, Cntry, Nat, Ht, Wt, DftYr, DftRd, Ovrl, Hand, Position, Team, GP, G, A to be kept in the model or at least considered then proven other wise. The variable mentions are all easy to track and you could find them for any player in the NHL. Below we can see distribution of players age and salaries in the data set. I found the salaries very interesting

Limitation: Since the data is based off one season with under 1000 observations and with many player missing values the data set became even smaller. Additionally, there was little to no information on the data quality and if there was any bias on the selected player that were recorded.



(distribution of the salary range for the season)

Data evaluation (Materials)

To clean the data, I first disregarded team statistics as the model should be built by measure that the player has individually created. I created the Age variable from "Born" Then, I individually picked the following variables:
'Age','Cntry','Ht','Wt','Hand','Position','Team','GP','G','A','plus.minus','Pass','PIM'. A majority of the variables were heavily correlated. For example the variable 'Points' is the players goals and assists added together. Also the variable Cntry had a few observations where that player was the only person in the data set with that country of origin. This would cause problems in Cross validation when the model that was created doesn't have that observation. The same thing happened would happen with Position but I altered this variable to have the levels 'Defense' or 'Forward'. So I filtered the data for countries 'CAN','USA','SWE','RUS','SVK','FIN' (the size shrunk by 73). Also not all players were drafted leaving 'NA' values under DftYr, DftRd, and Ovrl. I chose to take the variable out rather than only have the model on drafted players. A handful of the variable were extensively categorical and when using the dummy variable method to use them it creates 100+ new variables. Unfortunately, my machine would not be able to handle that multiple times though various models so I needed narrow the selection of variables by hand to start.

Modeling introduction (Methods)

Values created on seed '478', RMSE was chosen to evaluate model to allow for the 'punishment' of outliers but also allow the value to be a managable size. MSE created values around $2.0 * 10^{12}$

Linear Model:

Age, Country, Height, Weight, Hand (R/L), Position, Games Played, Goals, Assists, Plus-minus, Pass, PIM.
RMSE 1543401
(appendix 8)

A linear Step-wise model using AIC:

Salary ~ Age + Wt + Position + GP + G + A + plus.minus, with a AIC 18137.16 RMSE: 1538034
(appendix 9)

Random Forest:

RMSE: 1500074
(appendix 10)

K- nearest Neighbors at knn = 4:

RMSE: 1724816
(appendix 11)

Elastic Net regularization:

RMSE: 2848689
(appendix 12)

Best Model:

Out of all the models on an 80% training and 20% validation set the random forest out performing the other models when compared by RMSE. With the second best being a linear model with both direction stepwise selection. It is likely the model over-fit to the training data. The selection process is unstable a very dependent on the direction and present data. In this case I found both direction process to be the best after testing between the direction over different samples of the data. The elastic-net method would have dealt with any multi-collinearity better than the linear model would have if there were any.

Limitations:

Random Forests have low interpretability and could be prone to over-fitting. Also the data has a majority of lower salaries compared to its few high paying players. This may lead to a bias toward lower salaries. Tuning the parameters of the random forest may help these.

Final models and analysis.

Comparison of Model Performance on seed 478. Models at Appendix 13 and 14,

Model	Best Hyperparameters	RMSE
Stepwise Regression	Not applicable	1538034
Random Forest	mtry = 6	1505429

(appendix 1)

Explainable Model: Stepwise Regression

Stepwise regression is an interpretable model where each feature's coefficient can be analyzed for its direct impact on the target variable (Salary).

Coefficients of both Direction Stepwise Model (appendix 2):

	Positive Coefficients
Age	154375.3
Wt	17491.3
G	57263.4
A	104046.9

	Negative Coefficients
(Intercept)	-5760774.26
PositionForward	-491605.74
GP	-12464.97
plus.minus	-27282.80

	P Values
(Intercept)	0.0000000
Age	0.0000000
Wt	0.0000363
PositionForward	0.0012824
GP	0.0003412
G	0.0000080
A	0.0000000
plus.minus	0.0000320

Metric	Value
R-squared	0.5527384
Adjusted R-squared	0.5477609

	2.5 %	97.5 %
(Intercept)	-7509005.324	-4012543.192
Age	126373.122	182377.564
Wt	9233.497	25749.104
PositionForward	-790071.067	-193140.409
GP	-19261.298	-5668.635
G	32289.595	82237.211
A	87285.278	120808.531
plus.minus	-40071.879	-14493.716

The most influential variable when looking at the the coefficients is Age. At first, this sounded counter intuitive as almost every player experiences a decline in performance with age. But, once you circle back to how contracts are given to high level players it can make sense. Great players that stay in the league (past "prime" age) will be locked in by multi-year contracts that will be paying a serious amount of money near the end of their career. Also, rookie contract for young players are substantially less that that of players in the late 20s early 30s. Something to take note of is that Assists (+ $104,046) almost doubles the impact of Goals (+ $57263.4) when looking at Salary. Logically, an average person

might think that goals would be more influential but the model say otherwise. When you remember an assists leads to a goal, a player with 10 goals and 20 assists will be more beneficial their own salary in this case than a player with 25 goals and 5 assists. Even though they both have the same amount of points the model would state the player with more assist would be paid more. While weight was able to increase salary we cannot say the same for "Position Forward", "Games Played" and "Plus-Minus". Being a forward led to a decrease of salary by 491,605. This could be attributed to teams having a majority of forwards per team over defenseman or that team will pay more for good defenseman that keep the puck out of their net. Games played is when we see the fatigue of time, it could drop salary by 12,464 per game. Although games played was the amount that the individual played in the 2016-17 season not all time. Lastly, a drop in Plus-Minus would lead to a decrease in salary of 27,282. All the variables were significant on a basis of $\alpha = 0.01$ so we know that each variable had a meaningful impact on salary. The variance was explained 55.3% by the predictors of the model, this metric seaming to be the models weakest point. So while the model had moderate explanatory power almost half of the variance remains unexplained. Looking at the 95% confidence intervals we can confirm their significance as non contain zero. When we look at Age, we can see that it would increase salary by an amount in the range of 126,373.122- 182,377.564.

Predictive Model: Random Forest Model Take-aways

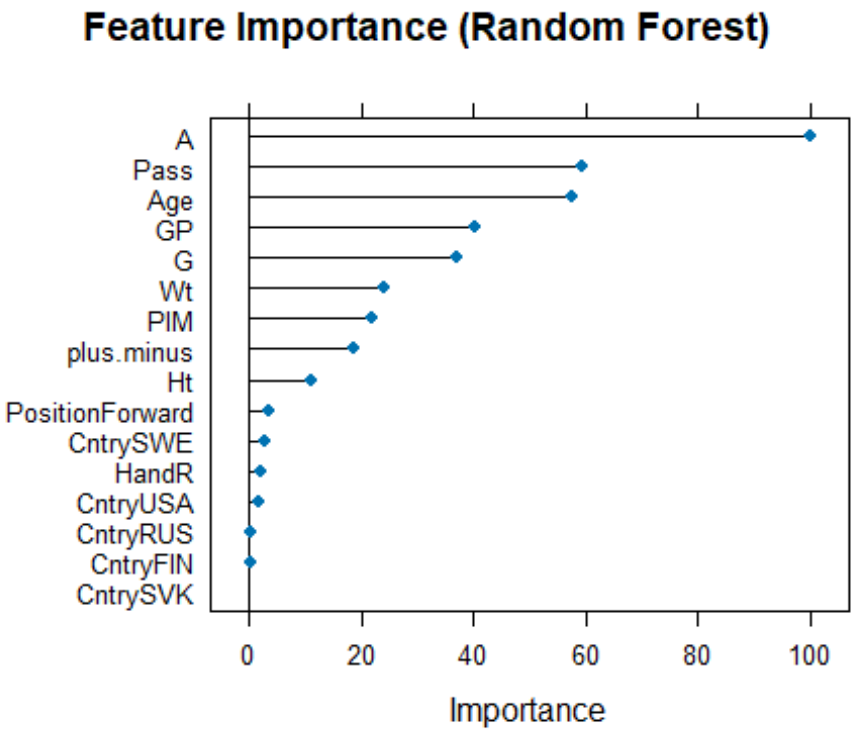
Assists Are Dominant: Assists are the most significant predictor in the random forest model, indicating that strong play-making ability directly correlates with Salary.

Age and Experience: Age is highly important, likely reflecting the combined effects of experience and performance longevity.

Physical Attributes: Weight (moderate importance) matters more than height (minimal importance), suggesting the model values physical strength over stature.

Other Metrics: Penalty minutes (PIM) and Plus/Minus have less impact on Salary, aligning with the perception that these stats are secondary to scoring and play-making.

Position's Limited Role: The low importance of being a forward suggests that other variables (e.g., goals, assists) capture performance-related information better than position alone.
(Importance graph below)


Feature Importance (Random Forest)
A
Pass
Age
GP
G
Wt
PIM
plus.minus
Ht
PositionForward
CntrySWE
HandR
CntryUSA
CntryRUS
CntryFIN
CntrySVK
0
20
40
60
80
100
Importance

(appendix 3)

Conclusions

This project provided valuable insights into hockey salary prediction and modeling approaches. Key factors like assists, goals, and age strongly influenced salaries, with physical metrics like weight playing a moderate role. Step-wise regression offered clear, interpretable results, while random forest captured complex relationships and delivered better predictive performance. By blending interpretability and prediction, we gained a comprehensive understanding of salary determinants and identified avenues for future refinement

Modeling take-aways:

Stepwise Regression: Useful for understanding key drivers but limited in modeling non-linear relationships.

Random Forest: Strong predictive accuracy but lacked transparency for stakeholder interpretation.

Pros and Cons and Future considerations

Pros: Stepwise regression was interpretable, and random forest captured complex interactions.

Cons: Step-wise missed interactions; random forest had low interpretability.

Future Changes: Combining models for better balance, exploring advanced metrics (e.g., xG), and fine-tuning models for robustness.

Citations

Data set at Kaggle
https://www.kaggle.com/datasets/camnugent/predict-nhl-player-salaries

Salary cap explained
https://www.sportingnews.com/us/nhl/news/nhl-salary-cap-rules-explained/x1wwiew656afzelhsx4tnecz

2016-17 salary cap tracker
https://www.spotrac.com/nhl/cap/_/year/2016/sort/cap_maximum_space2

Appendix

1:

```
library(knitr)

# table to show RMSE in knit
results <- data.frame(
  Model = c("Stepwise Regression", "Random Forest"),
  `Best Hyperparameters` = c("Not applicable", "mtry = 6"),
  RMSE = c(step.rmse, rf.rmse)
)

kable(results,
              caption = "Comparison of Model Performance on seed 478",
              col.names = c("Model", "Best Hyperparameters", "RMSE"),
              format = "markdown")
```

2:

```
library(dplyr)

# tables for knit to show values of model
model <- both

  coef <- model$coefficients
  model.summary <- summary(model)

  pos <- data.frame(
    "Value" = coef[coef > 0]
    )
  kable(pos, caption= "Positive Coefficients")

  neg <- data.frame(
    Value = coef[coef < 0]
  )
  kable(neg, caption= "Negative Coeffcients")

  pval <- data.frame(
    Value = model.summary$coefficients[, "Pr(>|t|)"]
  )
  kable(pval, caption= "P-Values")

  r2 <- model.summary$r.squared
  adjr <- model.summary$adj.r.squared

  r.squared.df <- data.frame(
  Metric = c("R-squared", "Adjusted R-squared"),
  Value = c(r2, adjr))
```

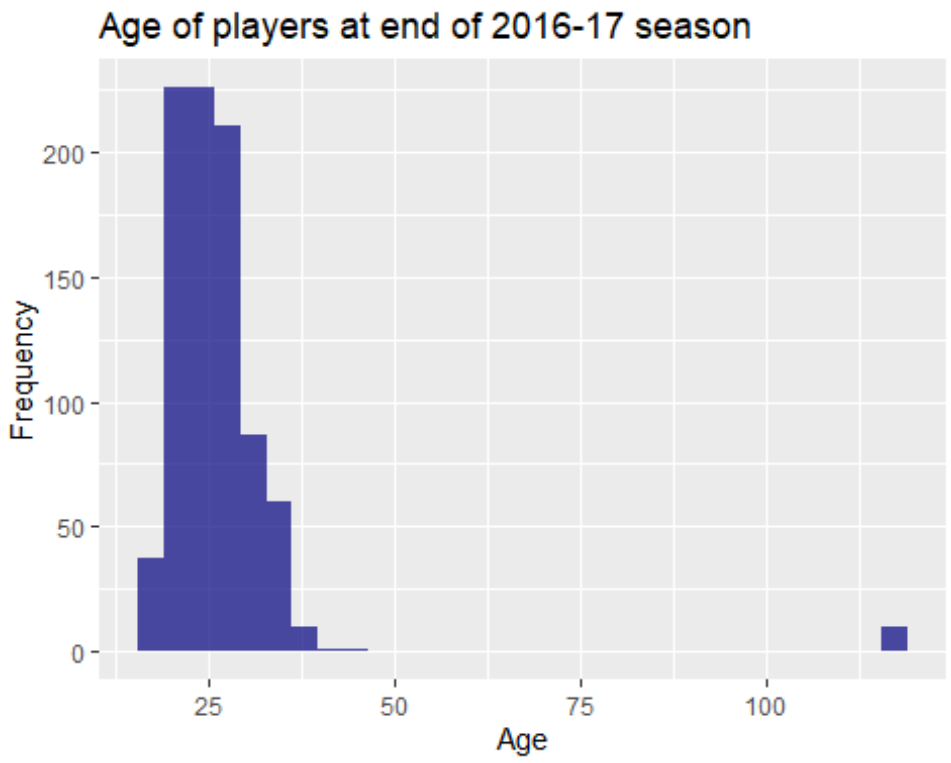
```
  kable(r.squared.df, caption= 'Adjusted R-Squared')

  conf_intervals <- confint(both)
  kable(conf_intervals, Caption= "Confidence Intervals")
```

3:

```
#random forest importance graph
importance <- varImp(tuned.rf, scale = TRUE)

plot.importance <- plot(importance, main = "Feature Importance (Random
Forest)")
plot.importance
```

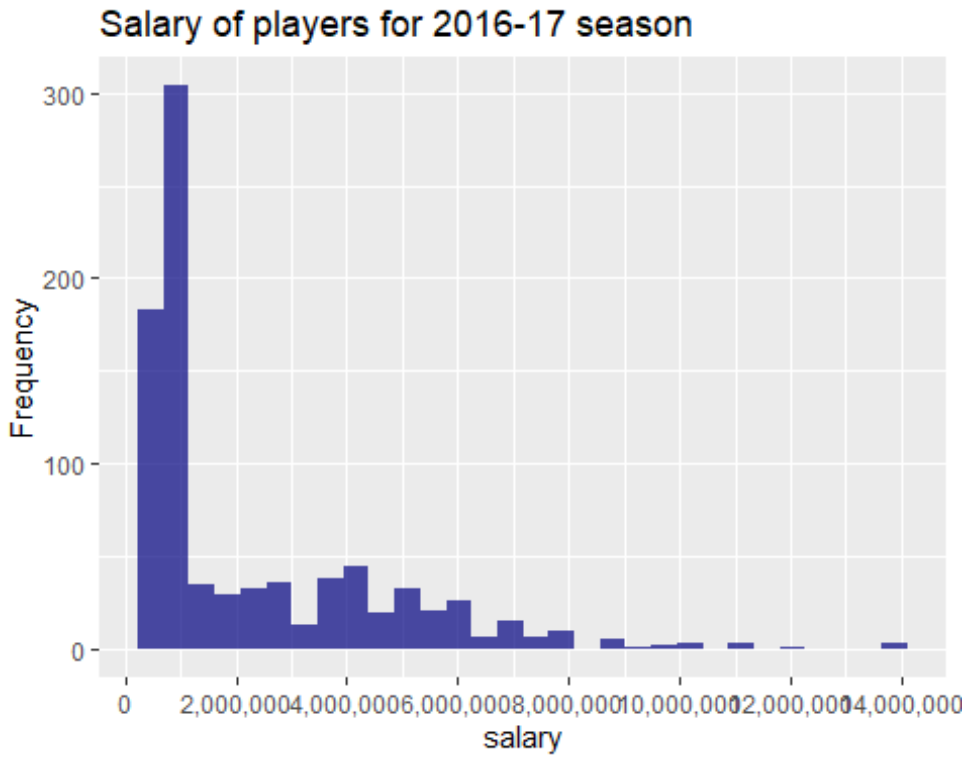
4:

```
library(dplyr)
library(ggplot2)
#load data and convert 'Born' to Age
nhl.salaries <- read.csv("./Raw_Data/salaries.csv")
varaible.set <- sub("^", "19", nhl.salaries$Born)
varaible.set <- as.Date(varaible.set, format = '%Y-%m-%d')
varaible.set <- data.frame(Born = varaible.set, salary= nhl.salaries$Salary)

date.for.age <- as.Date("2017-5-1")
varaible.set <- varaible.set %>%
  mutate(Age = as.numeric(floor((date.for.age - Born) / 365.25)))

#Age histogram
ggplot(varaible.set, aes(x= Age)) +
  geom_histogram(fill = "navy", alpha = 0.7) +
  labs(title = "Age of players at end of 2016-17 season", x = "Age", y =
"Frequency")
```



5:

```
#salary histogram
ggplot(varaible.set, aes(x= salary)) +
  geom_histogram(fill = "navy", alpha = 0.7) +
  labs(title = "Salary of players for 2016-17 season", x = "salary", y =
"Frequency") +
  scale_x_continuous(labels = scales::comma,
                     breaks= pretty(varaible.set$salary, n = 5))
```



6:

```
#create clean data set, selected varaibles and fixed Posistion variable to be
2 factor
colnames(nhl.salaries)[colnames(nhl.salaries) == "X..."] <- "plus.minus"

nhl.salaries <- nhl.salaries %>%
  mutate(Born = as.Date(Born, format = "%y-%m-%d"),
         Age = as.numeric(difftime('2016-8-1', Born, units = "days")) %/%
365.25,
         Position = ifelse(grepl("D", Position), "Defense", "Forward"),
         Position= factor(Position))

nhl.salaries.clean <- nhl.salaries  %>%
  select("Salary",
         "Age",
         'Cntry',
         'Ht',
         'Wt',
         'Hand',
         'Position',
         'GP',
         'G',
         'A',
         'plus.minus',
         'Pass',
```

```
        'PIM',) %>%
  filter(Cntry %in% c('CAN','USA','SWE','RUS','SVK','FIN'))
```

7:

```
library(glmnet)
library(randomForest)
library(tree)
library(caret)

# seed for reproducability
set.seed(478)

#Train and validation set
nhl.salaries.clean <- na.omit(nhl.salaries.clean)
sal.shuffle <- nhl.salaries.clean[sample(nrow(nhl.salaries.clean)),]
indices <- sample(1:nrow(sal.shuffle), size = 0.8 * nrow(sal.shuffle))
train <- sal.shuffle[indices,]
test <- sal.shuffle[-indices,]
```

8:

```
##################Linear model with all predictors from cleaned data set
all.fit <- lm(Salary~. , data= train)
#summary(all.fit)
#Predict values for RMSE
pred <- predict(all.fit, test)
rmse <- sqrt(mean((test$Salary - pred)^2))
rmse
```

```
## [1] 1543401
```

9:

```
####################Variable selection lm model
both <- step(all.fit, direction= "both", trace= 0)
#Predict values for RMSE
stepwise.pred <- predict(both, test)
rmse <- sqrt(mean((test$Salary - stepwise.pred)^2))
rmse
```

```
## [1] 1538034
```

```
# both : Salary ~ Age + Wt + Position + GP + G + A + plus.minus
# AIC 18137.16
```

10:

```r
####################random forest
rf.model <- randomForest(Salary~., data= train)
#Predict values for RMSE
rf.pred <- predict(rf.model, test)
rmse <- sqrt(mean((test$Salary - rf.pred)^2))
rmse
```

```
## [1] 1500074
```

11:

```r
##################knn
train_numeric <- train[, sapply(train, is.numeric)]
train.scaled <- as.data.frame(scale(train_numeric))
train.scaled$Salary <- train$Salary
knn <- knnreg(Salary~., data= train.scaled, knn= 4)
#Predict values for RMSE
knn.pred <- predict(knn, test )
rmse <- sqrt(mean((test$Salary - knn.pred)^2))
rmse
```

```
## [1] 3404763
```

12:

```r
##################Elastic net
x <- model.matrix(Salary~., data= train)[,-1]
y <- train$Salary

enet <- cv.glmnet(x, y, alpha= 0.5)
best.lambda <- enet$lamdba.min
X <- model.matrix(Salary~., data= train)[,-1]
#Predict values for RMSE
enet.pred <- predict(enet, newx= X, s= "lambda.min")
rmse <- sqrt(mean((test$Salary - enet.pred)^2))
rmse
```

```
## [1] 2848689
```

13:

```r
#Variable selection lm model cross validated
all.fit <- lm(Salary~. , data= train)
# list for each fold RMSE value
rmse.list.step <- c()
for(fold in 1:10)
{
  # fir model
  both <- step(all.fit, direction= "both", trace= 0)
```

```
  #predict values then find RMSE
  stepwise.pred <- predict(both, test)
  rmse <- sqrt(mean((test$Salary - stepwise.pred)^2))
  # save value
  rmse.list.step <- append(rmse.list.step, rmse)
}
#cross validated rmse
step.rmse<- mean(rmse.list.step)
step.rmse

## [1] 1538034
```

14:

```
#random forest
set.seed(478) # for reproduction

# using train() for CV
train_control <- trainControl(method = "cv", number = 10)

tuned.rf <- train(Salary ~ .,
                  data = train,
                  method = "rf",
                  trControl = train_control,
                  tuneGrid = expand.grid(mtry = c(3, 4, 5, 6, 7, 8, 9, 10)))

best.mtry <- tuned.rf$bestTune$mtry


# Prediction and RMSE
rf.pred <- predict(tuned.rf, newdata = test)
rf.rmse <- sqrt(mean((test$Salary - rf.pred)^2))
```

All Variable Definitions

Acronym - Meaning

%FOT - Percentage of all on-ice faceoffs taken by this player.

+/- - Plus/minus

1G - First goals of a game

A/60 - Events Against per 60 minutes, defaults to Corsi, but can be set to another stat

A1 - First assists, primary assists

A2 - Second assists, secondary assists

BLK% - Percentage of all opposing shot attempts blocked by this player

Born - Birth date

C.Close - A player shot attempt (Corsi) differential when the game was close

C.Down - A player shot attempt (Corsi) differential when the team was trailing

C.Tied - A player shot attempt (Corsi) differential when the team was tied

C.Up - A player shot attempt (Corsi) differential when the team was in the lead

CA - Shot attempts allowed (Corsi, SAT) while this player was on the ice

Cap Hit - The player's cap hit

CBar - Crossbars hit

CF - The team's shot attempts (Corsi, SAT) while this player was on the ice

CF.QoC - A weighted average of the Corsi percentage of a player's opponents

CF.QoT - A weighted average of the Corsi percentage of a player's linemates

CHIP - Cap Hit of Injured Player is games lost to injury multiplied by cap hit per game

City - City of birth

Cntry - Country of birth

DAP - Disciplined aggression proxy, which is hits and takeaways divided by minor penalties

DFA - Dangerous Fenwick against, which is on-ice unblocked shot attempts weighted by shot quality

DFF - Dangerous Fenwick for, which is on-ice unblocked shot attempts weighted by shot quality

DFF.QoC - Quality of Competition metric based on Dangerous Fenwick, which is unblocked shot attempts weighted for shot quality

DftRd - Round in which the player was drafted

DftYr - Year drafted

Diff - Events for minus event against, defaults to Corsi, but can be set to another stat

Diff/60 - Events for minus event against, per 60 minutes, defaults to Corsi, but can be set to another stat

DPS - Defensive point shares, a catch-all stats that measures a player's defensive contributions in points in the standings

DSA - Dangerous shots allowed while this player was on the ice, which is rebounds plus rush shots

DSF - The team's dangerous shots while this player was on the ice, which is rebounds plus rush shots

DZF - Shifts this player has ended with an defensive zone faceoff

dzFOL - Faceoffs lost in the defensive zone

dzFOW - Faceoffs win in the defensive zone

dzGAPF - Team goals allowed after faceoffs taken in the defensive zone

dzGFPF - Team goals scored after faceoffs taken in the defensive zone

DZS - Shifts this player has started with an defensive zone faceoff

dzSAPF - Team shot attempts allowed after faceoffs taken in the defensive zone

dzSFPF - Team shot attempts taken after faceoffs taken in the defensive zone

E+/- - A player's expected +/-, based on his team and minutes played

ENG - Empty-net goals

Exp dzNGPF - Expected goal differential after faceoffs taken in the defensive zone, based on the number of them

Exp dzNSPF - Expected shot differential after faceoffs taken in the defensive zone, based on the number of them

Exp ozNGPF - Expected goal differential after faceoffs taken in the offensive zone, based on the number of them

Exp ozNSPF - Expected shot differential after faceoffs taken in the offensive zone, based on the number of them

F.Close - A player unblocked shot attempt (Fenwick) differential when the game was close

F.Down - A player unblocked shot attempt (Fenwick) differential when the team was trailing

F.Tied - A player unblocked shot attempt (Fenwick) differential when the team was tied

F.Up - A player unblocked shot attempt (Fenwick) differential when the team was in the lead. Not the best acronym.

F/60 - Events For per 60 minutes, defaults to Corsi, but can be set to another stat

FA - Unblocked shot attempts allowed (Fenwick, USAT) while this player was on the ice

FF - The team's unblocked shot attempts (Fenwick, USAT) while this player was on the ice

First Name -

FO% - Faceoff winning percentage

FO%vsL - Faceoff winning percentage against lefthanded opponents

FO%vsR - Faceoff winning percentage against righthanded opponents

FOL - The team's faceoff losses while this player was on the ice

FOL.Close - Faceoffs lost when the score was close

FOL.Down - Faceoffs lost when the team was trailing

FOL.Up - Faceoffs lost when the team was in the lead

FovsL - Faceoffs taken against lefthanded opponents

FovsR - Faceoffs taken against righthanded opponents

FOW - The team's faceoff wins while this player was on the ice

FOW.Close - Faceoffs won when the score was close

FOW.Down - Faceoffs won when the team was trailing

FOW.Up - Faceoffs won when the team was in the lead

G - Goals

G.Bkhd - Goals scored on the backhand

G.Dflct - Goals scored with deflections

G.Slap - Goals scored with slap shots

G.Snap - Goals scored with snap shots

G.Tip - Goals scored with tip shots

G.Wrap - Goals scored with a wraparound

G.Wrst - Goals scored with a wrist shot

GA - Goals allowed while this player was on the ice

Game - Game Misconduct penalties

GF - The team's goals while this player was on the ice

GP - Games Played

Grit - Defined as hits, blocked shots, penalty minutes, and majors

GS - The player's combined game score

GS/G - The player's average game score

GVA - The team's giveaways while this player was on the ice

GWG - Game-winning goals

GWG - Game-winning goals

HA - The team's hits taken while this player was on the ice

Hand - Handedness

HF - The team's hits thrown while this player was on the ice

HopFO - Opening faceoffs taken at home

HopFOW - Opening faceoffs won at home

Ht - Height

iBLK - Shots blocked by this individual

iCF - Shot attempts (Corsi, SAT) taken by this individual

iDS - Dangerous shots taken by this player, the sum of rebounds and shots off the rush

iFF - Unblocked shot attempts (Fenwick, USAT) taken by this individual

iFOL - Faceoff losses by this individual

iFOW - Faceoff wins by this individual

iGVA - Giveaways by this individual

iHA - Hits taken by this individual

iHDf - The difference in hits thrown by this individual minus those taken

iHF - Hits thrown by this individual

iMiss - Individual shots taken that missed the net.

Injuries - List of types of injuries incurred, if any

iPEND - Penalties drawn by this individual

iPenDf - The difference in penalties drawn minus those taken

iPENT - Penalties taken by this individual

IPP% - Individual points percentage, which is on-ice goals for which this player had the goal or an assist

iRB - Rebound shots taken by this individual

iRS - Shots off the rush taken by this individual

iSCF - All scoring chances taken by this individual

iSF - Shots on goal taken by this individual

iTKA - Takeaways by this individual

ixG - Expected goals (weighted shots) for this individual, which is shot attempts weighted by shot location

Last Name -

Maj - Major penalties taken

Match - Match penalties

MGL - Games lost due to injury

Min - Minor penalties taken

Misc - Misconduct penalties

Nat - Nationality

NGPF - Net Goals Post Faceoff. A differential of all goals within 10 seconds of a faceoff, relative to expectations set by the zone in which they took place

NHLid - NHL player id useful when looking at the raw data in game files

NMC - What kind of no-movement clause this player's contract has, if any

NPD - Net Penalty Differential is the player's penalty differential relative to a player of the same position with the same ice time per manpower situation

NSPF - Net Shots Post Faceoff. A differential of all shot attempts within 10 seconds of a faceoff, relative to expectations set by the zone in which they took place

NZF - Shifts this player has ended with a neutral zone faceoff

nzFOL - Faceoffs lost in the neutral zone

nzFOW - Faceoffs won in the neutral zone

nzGAPF - Team goals allowed after faceoffs taken in the neutral zone

nzGFPF - Team goals scored after faceoffs taken in the neutral zone

NZS - Shifts this player has started with a neutral zone faceoff

nzSAPF - Team shot attempts allowed after faceoffs taken in the neutral zone

nzSFPF - Team shot attempts taken after faceoffs taken in the neutral zone

OCA - Shot attempts allowed (Corsi, SAT) while this player was not on the ice

OCF - The team's shot attempts (Corsi, SAT) while this player was not on the ice

ODZS - Defensive zone faceoffs that occurred without this player on the ice

OFA - Unblocked shot attempts allowed (Fenwick, USAT) while this player was not on the ice

OFF - The team's unblocked shot attempts (Fenwick, USAT) while this player was not on the ice

OGA - Goals allowed while this player was not on the ice

OGF - The team's goals while this player was not on the ice

ONZS - Neutral zone faceoffs that occurred without this player on the ice

OOZS - Offensive zone faceoffs that occurred without this player on the ice

OpFO - Opening faceoffs taken

OpFOW - Opening faceoffs won

OppCA60 - A weighted average of the shot attempts (Corsi, SAT) the team allowed per 60 minutes of a player's opponents

OppCF60 - A weighted average of the shot attempts (Corsi, SAT) the team generated per 60 minutes of a player's opponents

OppFA60 - A weighted average of the unblocked shot attempts (Fenwick, USAT) the team allowed per 60 minutes of a player's opponents

OppFF60 - A weighted average of the unblocked shot attempts (Fenwick, USAT) the team generated per 60 minutes of a player's opponents

OppGA60 - A weighted average of the goals the team allowed per 60 minutes of a player's opponents

OppGF60 - A weighted average of the goals the team scored per 60 minutes of a player's opponents

OppSA60 - A weighted average of the shots on goal the team allowed per 60 minutes of a player's opponents

OppSF60 - A weighted average of the shots on goal the team generated per 60 minutes of a player's opponents

OPS - Offensive point shares, a catch-all stats that measures a player's offensive contributions in points in the standings

OSA - Shots on goal allowed while this player was not on the ice

OSCA - Scoring chances allowed while this player was not on the ice

OSCF - The team's scoring chances while this player was not on the ice

OSF - The team's shots on goal while this player was not on the ice

OTF - Shifts this player started with an on-the-fly change

OTG - Overtime goals

OTOI - The amount of time this player was not on the ice.

Over - Shots that went over the net

Ovrl - Where the player was drafted overall

OxGA - Expected goals allowed (weighted shots) while this player was not on the ice, which is shot attempts weighted by location

OxGF - The team's expected goals (weighted shots) while this player was not on the ice, which is shot attempts weighted by location

OZF - Shifts this player has ended with an offensive zone faceoff

ozFO - Faceoffs taken in the offensive zone

ozFOL - Faceoffs lost in the offensive zone

ozFOW - Faceoffs won in the offensive zone

ozGAPF - Team goals allowed after faceoffs taken in the offensive zone

ozGFPF - Team goals scored after faceoffs taken in the offensive zone

OZS - Shifts this player has started with an offensive zone faceoff

ozSAPF - Team shot attempts allowed after faceoffs taken in the offensive zone

ozSFPF - Team shot attempts taken after faceoffs taken in the offensive zone

Pace - The average game pace, as estimated by all shot attempts per 60 minutes

Pass - An estimate of the player's setup passes (passes that result in a shot attempt)

Pct% - Percentage of all events produced by this team, defaults to Corsi, but can be set to another stat

PDO - The team's shooting and save percentages added together, times a thousand

PEND - The team's penalties drawn while this player was on the ice

PENT - The team's penalties taken while this player was on the ice

PIM - Penalties in minutes

Position - Positions played. NHL source listed first, followed by those listed by any other source.

Post - Times hit the post

Pr/St - Province or state of birth

PS - Point shares, a catch-all stats that measures a player's contributions in points in the standings

PSA - Penalty shot attempts

PSG - Penalty shot goals

PTS - Points. Goals plus all assists

PTS/60 - Points per 60 minutes

QRelCA60 - Shot attempts allowed per 60 minutes relative to how others did against the same competition

QRelCF60 - Shot attempts per 60 minutes relative to how others did against the same competition

QRelDFA60 - Weighted unblocked shot attempts (Dangeorus Fenwick) allowed per 60 minutes relative to how others did against the same competition

QRelDFF60 - Weighted unblocked shot attempts (Dangeorus Fenwick) per 60 minutes relative to how others did against the same competition

RBA - Rebounds allowed while this player was on the ice. Two very different sources.

RBF - The team's rebounds while this player was on the ice. Two very different sources.

RelA/60 - The player's A/60 relative to the team when he's not on the ice

RelC/60 - Corsi differential per 60 minutes relative to his team

RelC% - Corsi percentage relative to his team

RelDf/60 - The player's Diff/60 relative to the team when he's not on the ice

RelF/60 - The player's F/60 relative to the team when he's not on the ice

RelF/60 - Fenwick differential per 60 minutes relative to his team

RelF% - Fenwick percentage relative to his team

RelPct% - The players Pct% relative to the team when he's not on the ice

RelZS% - The player's zone start percentage when he's on the ice relative to when he's not.

RopFO - Opening faceoffs taken at home

RopFOW - Opening faceoffs won at home

RSA - Shots off the rush allowed while this player was on the ice

RSF - The team's shots off the rush while this player was on the ice

S.Bkhd - Backhand shots

S.Dflct - Deflections

S.Slap - Slap shots

S.Snap - Snap shots

S.Tip - Tipped shots

S.Wrap - Wraparound shots

S.Wrst - Wrist shots

SA - Shots on goal allowed while this player was on the ice

Salary - The player's salary

SCA - Scoring chances allowed while this player was on the ice

SCF - The team's scoring chances while this player was on the ice

sDist - The average shot distance of shots taken by this player

SF - The team's shots on goal while this player was on the ice

SH% - The team's (not individual's) shooting percentage when the player was on the ice

SOG - Shootout Goals

SOGDG - Game-deciding shootout goals

SOS - Shootout Shots

Status - This player's free agency status

SV% - The team's save percentage when the player was on the ice

Team -

TKA - The team's takeaways while this player was on the ice

TMCA60 - A weighted average of the shot attempts (Corsi, SAT) the team allowed per 60 minutes of a player's linemates

TMCF60 - A weighted average of the shot attempts (Corsi, SAT) the team generated per 60 minutes of a player's linemates

TMFA60 - A weighted average of the unblocked shot attempts (Fenwick, USAT) the team allowed per 60 minutes of a player's linemates

TMFF60 - A weighted average of the unblocked shot attempts (Fenwick, USAT) the team generated per 60 minutes of a player's linemates

TMGA60 - A weighted average of the goals the team allowed per 60 minutes of a player's linemates

TMGF60 - A weighted average of the goals the team scored per 60 minutes of a player's linemates

TMSA60 - A weighted average of the shots on goal the team allowed per 60 minutes of a player's linemates

TMSF60 - A weighted average of the shots on goal the team generated per 60 minutes of a player's linemates

TmxGF - A weighted average of a player's linemates of the expected goals the team scored

TmxGA - A weighted average of a player's linemates of the expected goals the team allowed

TMGA - A weighted average of a player's linemates of the goals the team scored

TMGF - A weighted average of a player's linemates of the goals the team allowed

TOI - Time on ice, in minutes, or in seconds (NHL)

TOI.QoC - A weighted average of the TOI% of a player's opponents.

TOI.QoT - A weighted average of the TOI% of a player's linemates.

TOI/GP - Time on ice divided by games played

TOI% - Percentage of all available ice time assigned to this player.

Wide - Shots that went wide of the net

Wt - Weight

xGA - Expected goals allowed (weighted shots) while this player was on the ice, which is shot attempts weighted by location

xGF - The team's expected goals (weighted shots) while this player was on the ice, which is shot attempts weighted by location

xGF.QoC - A weighted average of the expected goal percentage of a player's opponents

xGF.QoT - A weighted average of the expected goal percentage of a player's linemates

ZS% - Zone start percentage, the percentage of shifts started in the offensive zone, not counting neutral zone or on-the-fly changes