Issues in Information Systems
Volume 23, Issue 4, pp. 279-292, 2022

DOI: https://doi.org/10.48009/4_iis_2022_124

Identifying factors that lead to injury in the NFL

Matthew Toner, *Bryant University, mtoner1@bryant.edu*
Suhong Li, *Bryant University, sli@bryant.edu*

Abstract

This study hypothesizes that injury-causing factors can be identified through training machine learning models with NFL injury data. The machine learning process entailed web scraping, pre-processing, cleaning, modeling, and analyzing NFL injury data to identify these factors. The features used to model injuries included the following: games played, games started, weight, height, age, year, years of experience, starting position, and team. The four models used to model NFL injuries were Logistic Regression, Decision Trees, Random Forests, and Gradient Boosted Trees. The model with the best performance was the Gradient Boosted Trees model, with an F1 score of 0.508. In addition, the Gradient Boosted Trees model selected teams, games played, and starting position to be the most influential factors when determining the probability of a player getting injured during the season. The most notable takeaway from this study is that the left guard position is the most injury-prone, followed by defensive secondary player positions such as safety and cornerback.

Keywords: athlete injury prevention, machine learning, undergraduate students

Introduction

Since the 1920s, the National Football League has moved away from being a brute, leather-helmeted game and transitioned into the sport we love today. Nevertheless, this inherently violent sport still poses serious health risks to players. Recent reports of severe brain injuries affecting the long-term mental health of players have placed the NFL under increased scrutiny from fans and players. For the NFL to survive as a sport, the league needs to find ways to minimize the risk of injury without sacrificing the physical contact relished by fans. The scope of this study seeks to address the incidence of injuries in the NFL to determine the factors causing injuries. This study illustrates that accurate data modeling of NFL injuries will prove insightful in identifying factors that cause injuries in the NFL.

The NFL has gained a poor reputation when it comes to protecting its players. In recent years, the league has attempted to regain the trust of its players by implementing rule changes such as increasing the number of yards a team receives for committing a touchback on a kickoff to discourage kick returns. This change was made because the kickoff is the most dangerous play in football. This literature review seeks to uncover other aspects of football that are more dangerous than others to predict player injuries. This review explores injuries to the head and body and identifies other factors such as race that contribute to the likelihood of a player being injured. This study hypothesizes that by incorporating these factors into a dataset and applying machine learning algorithms, factors that contribute to injuries in the NFL can be identified.

Literature Review

This section will discuss prominent injuries in the NFL, concussions, and machine learning methodologies used to forecast/prevent injuries.

Prominent Injuries in the NFL

Current literature suggests that specific injuries in the NFL have higher incidence rates than others. For example, knee injuries account for 29.3% of lower-body injuries, followed by injuries to the ankle (22.4%), thigh (17.2%), and foot (9.1%) (Mack et al., 2020). Lower body injuries are common in the NFL as the legs are typically the target point when a defensive player is tackling an offensive player. A similar study assesses the incidence of spine and axial skeleton injuries, a much more severe injury in the NFL. The study collected data from the NFL's injury surveillance database from 2000 to 2010 and learned that these injuries represented 7% of total injuries. This percentage equates to 2,208 injuries, where 987 occurred in the cervical spine. Additionally, time missed from play was most significant for thoracic disc herniations, where the average recovery period was 189 days (Mall et al., 2012). Furthermore, offense linemen followed by defensive backs, defensive linemen, and linebackers were most likely to suffer this injury. Compared to other positions, these positions prone to spinal injury engage in the most tackling and blocking on the field.

Additional injuries that NFL.com reports include concussions, ACL tears, and MCL tears from 2015 to 2021. For example, from 2015 to 2021, the occurrence of ACL tears increased from 30 to 49. And the occurrence of MCL tears has decreased from 110 to 98. On the other hand, the NFL's data claims that concussions have dropped from 192 in 2015 to 135 in 2021. In addition, the NFL argues that its rule changes have reduced concussions in the sport. Yet the NFL continues to face lawsuits over these injuries.

Concussions in the NFL

Concussions consistently taint the reputation of the NFL. An estimated 3.8 million concussions occur each year, with football causing more concussions than any other professional sport (Mostafavi et al., 2021). Although player equipment and league rules have evolved to increase player safety, concussions jeopardize the game's longevity. For example, during the 2010-2011 NFL season's opening day, Eagle's quarterback and middle liner backer Kevin Kolb and Stewart Bradley sustained concussions during the second quarter of the game (Lazarus, 2011). These injuries were caused by Kolb being tackled by an opposing player headfirst into the turf, whereas Bradley was accidentally hit in the head by a teammate's thigh. Kolb returned to play after twenty minutes and Bradley after four plays. Medical experts criticized the medical staff. However, Andy Reid, former coach of the Packers and current coach of the Kansas City Chiefs, explained, "The football players were fine […] All the mental status questions they answered on the sideline and the things they did with the team docs registered well. However, as the game went on, they weren't feeling well, so we took them out" (Lazarus, 2011). Two days after Kolb's concussion, he wrote on his Twitter, "Does anyone have good remedies for a killer headache" (Lazarus, 2011). This situation illustrates the negative impact of concussions on the game of football. By addressing the concussion issue within the NFL, players can sustain more NFL seasons and generate more wealth throughout their careers. Football is a lifelong investment for many of these athletes, and by irresponsibly managing their injuries, they are destroying a player's potential return on this investment. By making the game of football safer, the NFL will also benefit because a large pool of players may choose to pursue football professionally over other sports, which in turn increases the amount of skill within the league. This increase in agility will bring the game to new heights and retain football's image as a spectator's sport.

Current Methods for Injury Forecasting

Each year at the NFL, combine teams scout players based on their physical capabilities. Studies have recently been conducted to assess these players' bones and muscles' structural integrity at the combine to analyze the likelihood they will sustain specific injuries during their careers. Comparing the hamstring to quadriceps ratio of combine players resulted in an accurate prediction of injury later in their career 51.3% of the time. This low accuracy score was not improved upon when researchers measured side-to-side peak torque differences of legs (Zvijac & Toriscelli, 2013). However, a separate study by Wang et al. (2018) found that when analyzing a player's medical history, athletes with a history of cervical or lumbar spine injury, rotator cuff repair, superior labrum anterior-posterior repair, anterior cruciate ligament reconstruction, and other specific played in significantly fewer games in their NFL career (Wang, et al., 2018). The distinction between these two studies is that the first assessed muscle strength while the second study assessed prior injuries. This research suggests that previous injuries correlate to a greater chance of damage in the future. In contrast, the ratio of muscles in the body does not perform as a strong indicator of injury.

Machine Learning Approach

Another method for predicting injuries involves using machine learning algorithms. A critical component of any model is the accuracy of data collected. The majority of current models collect data from league-wide electronic health records. These records are used to record injuries that come to the attention of the team's athletic trainers and physicians, NFL medical spotters, or unaffiliated neurotrauma consultants (Dreyer et al., 2019). The primary concern with NFL medical data is the possibility of underreporting. This concern is supported because after switching from targeted reporting to an EHR, 20% more injuries were recorded for the same "reportable" injuries (Dreyer et al., 2019). However, underreporting is not the only hurdle in analyzing NFL data. Often the data is unbalanced and enormous. Data scientists have used feather, a fast disk format capable of reading and writing data frames at defined chunks, with each chunk containing around 5000 records. (Arulanantham et al., 2019). This format made reading data into the models easier for the data scientists and improved modeling accuracy.

Various machine learning models have been experimented with in modeling NFL injuries with varying results. For example, the SVM model, which stands for support vector machines, required extensive run times to classify injuries. A possible improvement for this study by Arulanantham et al. (2019) would be experimenting with kernelized SVMs. This algorithm is advantageous to traditional SVMs because they enable classification in a higher-dimensional space without calculating the new, possibly extensive representation. On the other hand, Random forests can be used as a potential model. An essential difference between SVMs and Random forests is that this Decision Tree-based classifier does not require meticulous pre-processing of the data. However, the data scientists in this study found that the model developed overfitting complications with the training data leading to poorer scores on the test data (Arulanantham et al., 2019). According to the study by Arulanantham et al. (2019), the best performing model was XGBoost, which displayed adequately higher scores without overfitting concerns. This model uses fit gradient boosting, which minimizes the max depth of decision trees. Instead of having the complex decision trees used in a Random Forest classifier, the XGBoost classifier uses a series of smaller trees to create complex models while reducing overfitting concerns. Given this study's success with gradient boosting, it is anticipated that gradient boosted trees will have the most substantial results when modeling NFL injury data in this study.

Researched Variables for Modeling Injuries

All machine learning models require an output variable (did an injury occur?) with various input variables. The best models incorporate all variables that have a statistically significant impact on predicting the output variable. Potential variables the current literature addresses and considers when forecasting injuries include: race, day of the week the game is played on, position, time in the season, playing surface, and shoe surface. Specific racial groups dominate certain positions in the NFL. Specifically, center, quarterback, and punter are 78%, 78%, 97% white, respectively (Krill et al., 2017). In contrast, there are four critical positions dominated by black players. They are cornerback (100% filled by black players), running back (90%), wide receiver (88%), defensive tackle (82%) (Woods et al., 2018). For black-dominated positions (defensive end, linebacker, running back, wide receiver and tight end), 605 black players were injured compared to146 white players during the 2013-2014 season. For white positions (quarterback, left and right tackle, and center), the rate of injury for black players was 29.4% compared to 22.1%. Since black players play primarily defensive positions, they should have a higher rate of injury since another study found defensive players suffered the majority of achilles tendon ruptures (73.7%) (Krill et al., 2017). Since black players are injured at higher rates and play more dangerous positions, a player's race is a potentially statistically significant variable in a predicting model.

As previously mentioned, the time of season plays a factor in determining the likelihood of an injury. Most injuries occur in games, with a higher rate of injury in the preseason than the regular season (11.5 vs. 9.4) (Mack et al., 2020). A separate study focusing on achilles tendon ruptures determined that 72.7% of AT ruptures were sustained in the first eight games of the regular season. A possible theory to explain a drop off in injury incidences later in the season may be underreporting. Player's are motivated to play in the playoffs for a chance to earn financial incentives included in their contract and are more likely to attempt to hide or play through an undisclosed injury.

Playing through minor injuries occurs throughout the season, suggesting more extended recovery periods between games may lower injury. Yet, studies on the impact of Thursday Night Football suggest otherwise. During Sunday and Monday Night Football games, the injury rate was 7,598 per 1000 athletic exposures compared to 6,072 per 1000 athletic exposures during Thursday night games (Baker et al., 2019). Although Thursday night games provide players with fewer days of rest, the data indicates this does not impact the incidence of injury. This example demonstrated less injury incidence than a regularly scheduled game.

Another variable in predicting injury is the playing surface and shoe design. Currently, 14 of the 32 stadiums in the NFL have grass fields rather than synthetic turf. Turf fields have the potential to reduce injury because "they yield a significantly higher peak torque and rotational stiffness than the natural grass surfaces" (Villwock et al., 2009). However, game data doesn't support this evidence. During 8 NFL seasons, there was an injury incidence rate of 1 per 100 total games on grass fields compared to 1.08 per 100 total games on artificial turf (Krill et al., 2017). Since the playing surface is not a statistically significant factor in whether a player will be injured, this variable will be excluded from the model unless a more compelling study refutes these statistics is discovered in the future.

Overall, race, player position, and time of season appear to be significant based on the research conducted. In contrast, field type, shoe type, and days of rest seem to have minimal impact on the likelihood of a player being injured. As more variables are considered, the complexity and accuracy of machine learning models will be improved.

Methodology

Data Collection/Pro-processing

The first step of the analysis to obtain data about the players and injuries. The player's data was retrieved from the website www.jt-sw.com, which held a collection of historic roster data. The following player attributes were collected for analysis: Team, Position, Jersey number, Name, GP – Games Played, GS – Games Started, Starting position, Years of experience, Date of birth , Height , Weight and College attended. In addition, the injury data was scraped from prosportstransactions.com. This website contained every injury incident in the NFL from 2016 to 2022 with the following columns: Date of Injury, Team, Name of the player acquired to the injury reserve, Name of the player relinquished from the injury reserve and Notes containing a description of the injury.

The player's data was then joined with the injury data. To ensure COVID-19 did not influence results, all COVID-19 related injury data was removed from the injury dataset. The players' data were filtered to include the years 2016-2019 to match the timeframe of the injury dataset. In addition, if a player was injured in a particular year, there would have no performance data for that player in that year. In this case, the player's performance data from the previous year will be used. A tremendous time was spent to pre-process and clean the data. For example, both the name of each player and team they are in was stored differently in two datasets and had to be cleaned/matched.

Additional features were created for further analysis. Each player's age was calculated using the player's date of birth. An injury column was created to represent whether a player was injured in a specific year. This column serves as the prediction target in this analysis. Various tools (python, pyspark and SQL) were used to engineer the data and run machine learning models.

The final features included in the DataFrame used for modeling include the following:

- Label (1 = injured, 0 = not injured)
- Player
- GP (Games Played)
- GS (Games Started)
- Wt (Weight)
- Year
- Age
- Height
- Experience (Seasons played in the NFL)
- StartPos1 (player's primary starting position)
- Pos1 (player's primary position)
- Team Name

Running Machine Learning Models

The data was split into a testing and training set, with 80% of the data being used to train the model and 20% reserved for testing the model's performance. To improve model performance, oversampling was applied to the training set. This mechanism involved randomly selecting examples from the minority class "not injured" with replacement and adding them to the training dataset.

Four different types of machine learning algorithms were used in this study, including logistic regression, decision trees, random forests, and gradient boosted trees. A parameter grid was created for each model to test various combinations of parameters specific to each. A 3-fold cross validator assessed each combination of parameters in the parameter grid and selected the best parameters for each model.

Discussion and Results

This section will first discuss whether injuries differ by type, weekday and time of season followed by discussing the results of machine learning models and identifying the most significant factors impacting injuries.

Injuries by Type

Figure 1a shows the dominant injury is knee injury and account for 52% of the total injuries, followed by 25% foot injury and 22% ankle injuries. It can be seen that most injuries are related to lower-body. The result is different from Mack et al. (2020) as this study contained significantly more knee and foot injuries. The difference may come from the different time periods in each study and how data was recorded in each database. In addition, there are 105 spinal injuries overall total 4,598 injuries, representing 2% of the spinal injuries.

Figure 1b shows the ACL tears, MCL tears and concussion between 2016 and 2021. ACL tears have the highest number of injuries compared to the other two injuries and has increased in 2021. MCL tears have a steady number over the years. Concussion has the lowest number in 2019. Some number in this study is inconsistent with the data reported on NFL.com. This may be due to the lack of access to the electronic health record systems of the NFL in this study.

Injuries by Weekday

Previous research shows that injury is associated with the day of the week a game was played. It is expected that most injuries would be recorded on the days NFL football is played: Monday, Thursday, and Sunday. Figure 1c shows that many injuries were recorded on days games were not held in this study. Sunday has the greatest total of injuries recorded (782), but this should be expected considering that most NFL games are held on Sunday. These findings align with Baker et al. (2019), who found that the injury rate was less during Thursday night games than during Sunday and Monday games. However, since the distribution of injuries in the dataset of this study occurs across weekdays, it does not seem logical to include weekdays as a feature as injuries are recorded days after they occur. Therefore, day of the week will not be used as a feature when building models.

Injuries by the time of season

Figure 1d shows that most injuries occurred during August (preseason), with 962 injuries, and September (first four weeks of the regular season) totaling 1123 injuries. This result is consistent with previous studies. Mack et al. (2020) discovered a higher rate of injury in the preseason than in the regular season (11.5 vs. 9.4). In addition, a separate study focusing on achilles tendon ruptures determined that 72.7% of AT ruptures were sustained in the first eight games of the regular season.

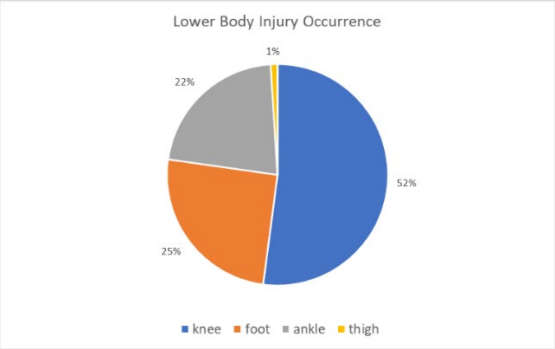

Figure 1a Lower Body Injury Occurrence from Web-Scapred Data Source

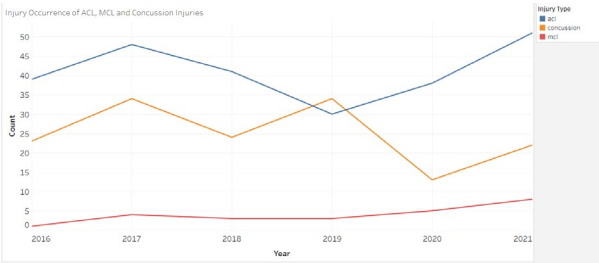

Figure 1b ACL, MCL and concussion between 2016 and 2021

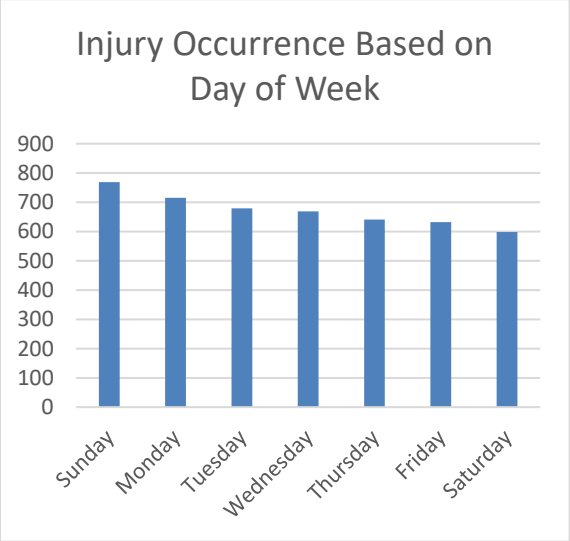

Figure 1c – Sum of injury occurrence based on day of the week the injury occurred

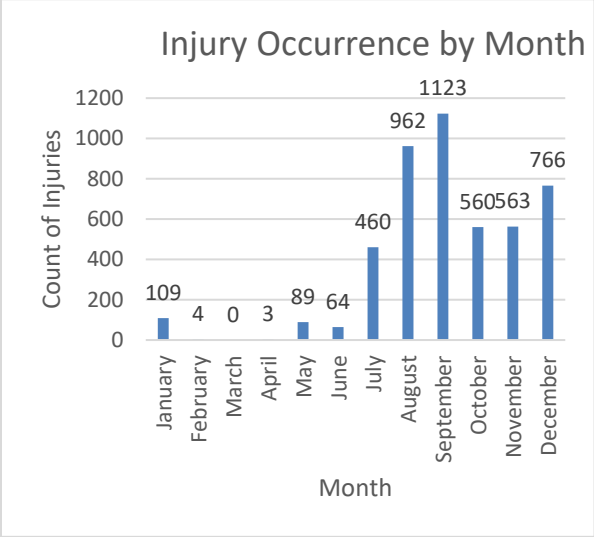

Figure 1d – Time of Season's effect on injury occurrence

Results of Machine Learning Models

Accuracy Scores

Four machine learning (logistic regression, decision trees, random forests, and gradient boosted trees) were used to find the best model and identify factors that are important in predicting injury. The results are shown in Figure 2 and evaluated based on the following metrics: accuracy, precision, recall and F1 score.

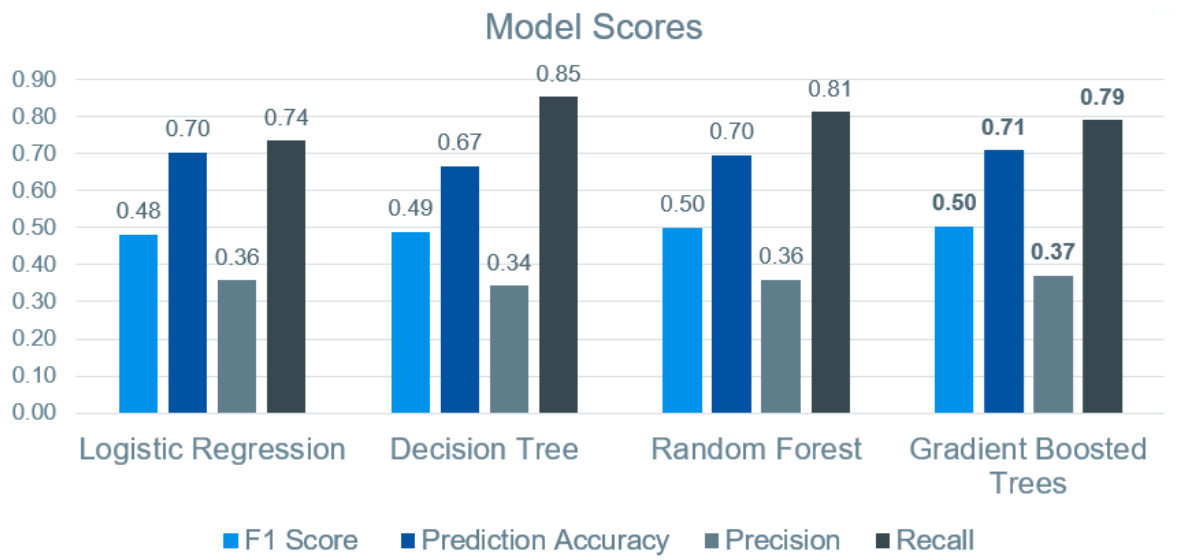

Figure 2 – Accuracy scores for each model used to predict injury label

The Gradient Boosted Tree model had the most remarkable accuracy score of 0.71. This model correctly predicted the target variable of whether or not a player was injured during a specific year 71% of the time. On the other hand, the Decision Tree was the model with the worst accuracy score at 0.67. However, a more accurate way to assess a model's performance is by using the F1 score. This is because the F1 score considers a model's precision and accuracy. Below is the formula to calculate the F1 score.

F1 = 2 x ((precision x recall) / (precision + recall))

It is essential to consider a model's precision because precision measures the quality of a positive prediction made by the model. Precision refers to the number of true positives divided by the total number of positive predictions. Gradient Boosted Trees had the best precision out of the four models tested. However, the Random Forest model had a better recall score than the Gradient Boosted Trees model. The recall score measures how well the model correctly identifies true positives. For example, the Random Forest recall score of 0.81 implies that the model correctly identified 81% of the injured players between 2016 and 2019. Overall, the best model was the Gradient Boosted Trees because it had a slightly higher F1 score, accuracy and precision score than Random Forest model.

To further assess the performance of each model, the lift charts for each model were compared side by side, as seen in Figure 3. The maximum lift point, the lift's intersection, and the sample's percentage at 0 were relatively similar for all models except Logistic Regression. The Logistic Regression model had a maximum lift of 3.5, whereas the other three had a maximum lift of around 5. The maximum lift implies that the logistic regression model is the worst when predicting injury with the players with the highest probability of being injured. On the other hand, when analyzing 20% of the sample, the Gradient Boosted Trees model has the best results. Based on the lift curve, the group of 20% of the population with the highest probability predicted by the algorithm would have around a 30% (3 times the 20% mean) proportion of injured players.

Another curve used to assess the performance of each model was the cumulative gains curve as shown in Figure 4. Similar to the lift curve, both are constructed by ordering data points from most likely to be injured on the left to least likely on the right. The cumulative gains curves for each model are pictured above in

Figure 4. Visually, the Gradient Boosted Tee model has the greatest amount of area between the class 1 line and the baseline. However, when analyzing the cumulative gains further, the Gradient Boosted Trees model captured 90% of injuries with 40% of the sample population. In contrast, the other three models captured 80% of injuries with the same portion of the sample population. This indicates that the Gradient Boosted Tree model is the best model when assessed using the cumulative gains curve.

Overall, the Gradient Boosted Tree model was the best model for predicting injuries in the NFL. This is in line with existing literature, as the study by Arulanantham et al. (2019) also found that gradient boosted classification algorithms were best for modeling injuries in the NFL. Therefore, considering this model was the strongest, important features were assessed using the Gradient Boosted Trees algorithm.

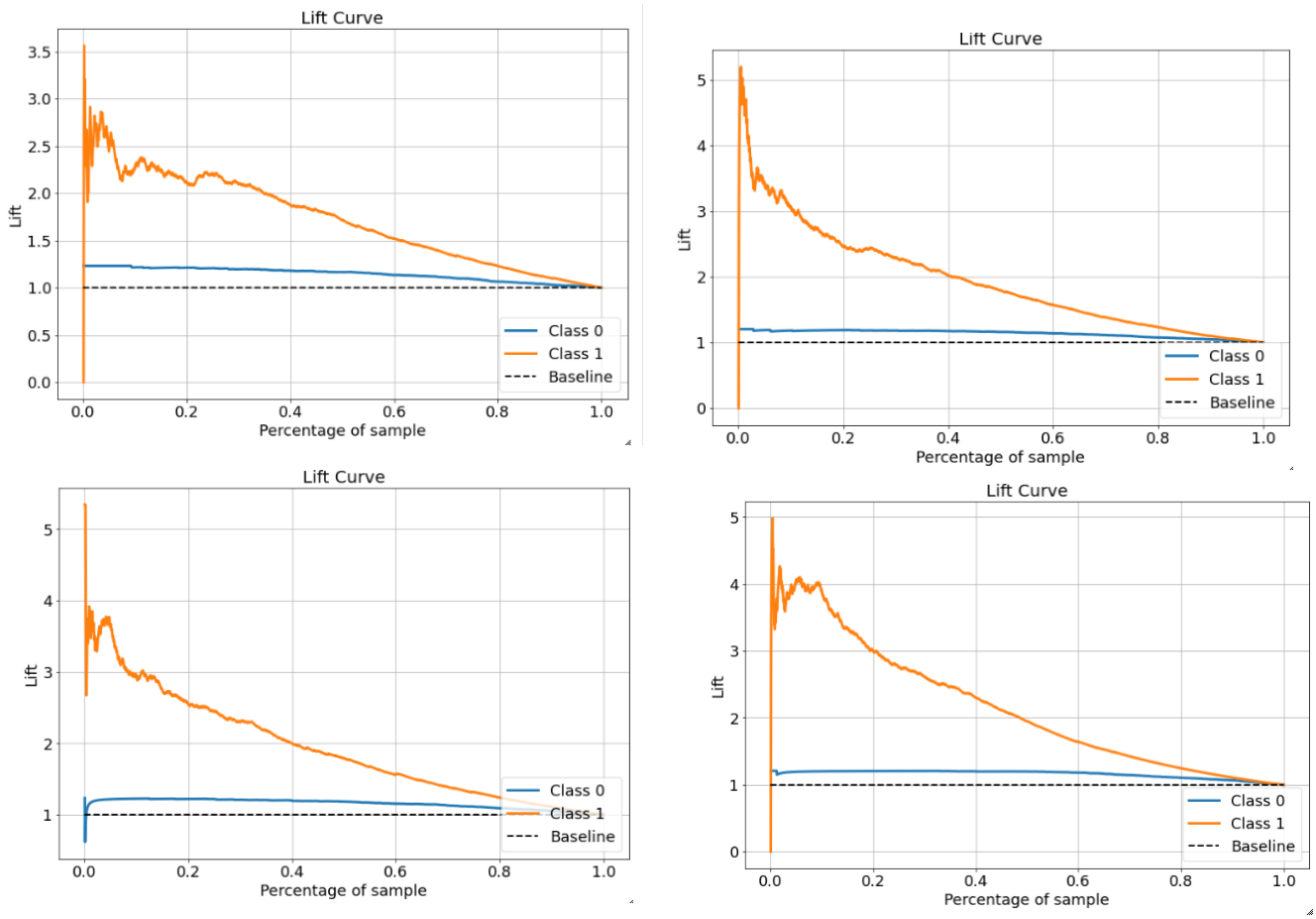

Figure 3 - Lift curves for each model. Logistic Regression (Top Left), Decision Trees (Top Right), Random Forest (Bottom Left), Gradient Boosted Trees (Bottom Right)

Feature Importance

The critical features identified by the Gradient Boosted Trees algorithm are displayed in Figure 5. The model identified team, games played, and starting position as the three most important features when determining whether an NFL player will be injured.

Team: Team has been identified as an important factor in predicting injury. A further analysis shows that there was a wide range of injury occurrences depending on the team. For example, the top five teams with

the highest injuries are Jets (72 injuries), Lions (66 injuries), Dolphins (65 injuries), Cardinals (63 injuries) and Texans (62 injuries). While the bottom five teams with the lowest injuries are Titians (30 injuries), Vikings (30 injuries), Falcons (29 injuries), Steelers (24 injuries) and Chargers (15 injuries). This range of injuries enabled the model to decrease the probability a player was injured if they played on the Chargers team.

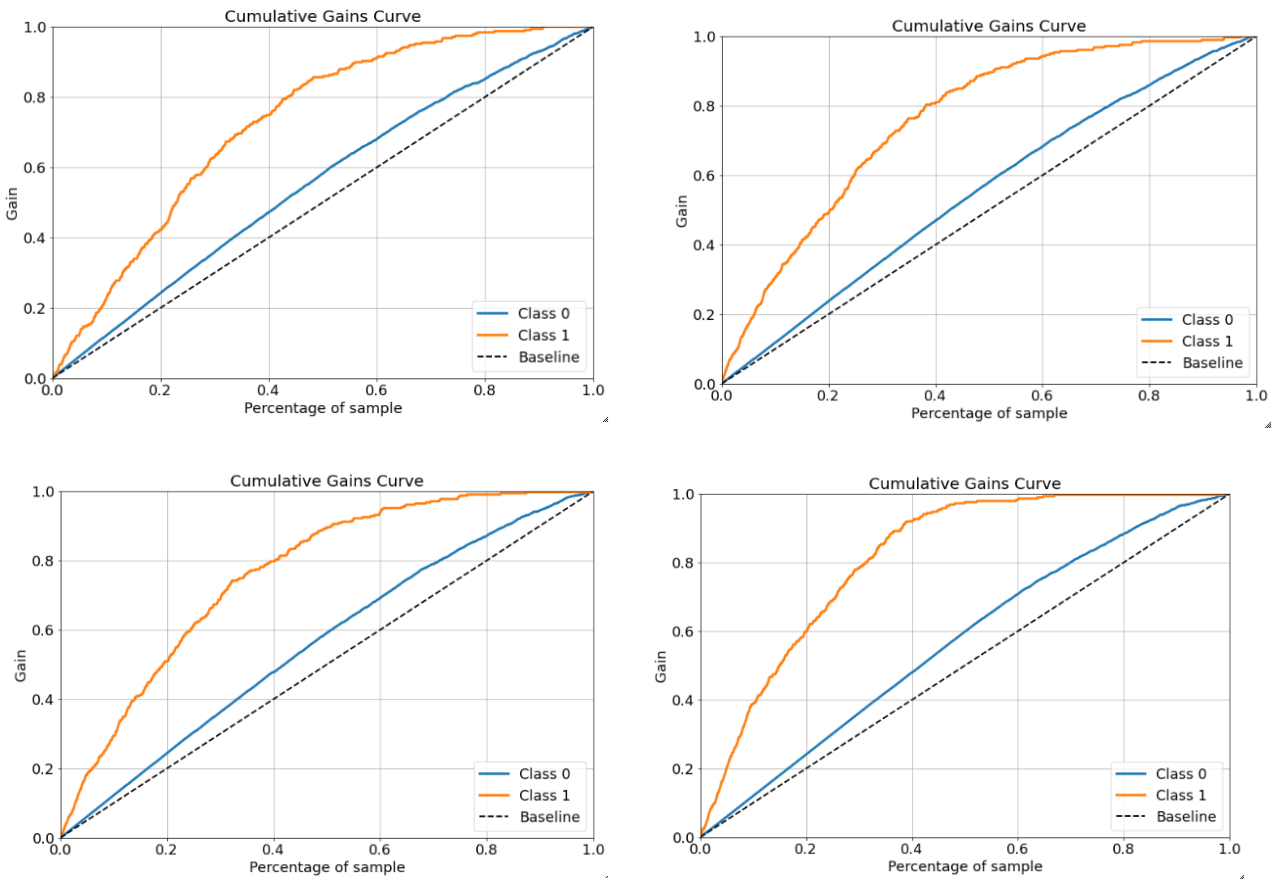

Figure 4 – Cumulative gains curves for each model. Logistic Regression (Top Left), Decision Trees (Top Right), Random Forest (Bottom Left), Gradient Boosted Trees (Bottom Right)

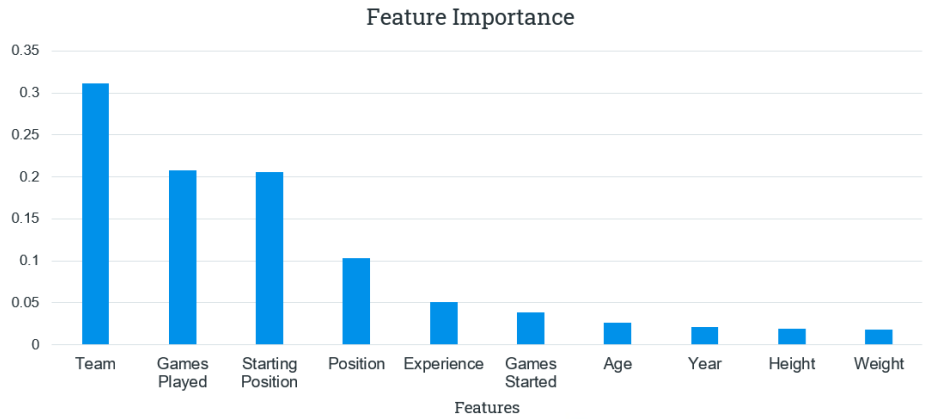

Figure 5 – Feature Importance from the Gradient Boosted Trees model

Games Played: Figure 6 shows the average number of injuries by the number of games played. Each number of games was started as its own column to analyze how the number of games started influenced injury occurrence. The y-axis in Figure 6 is the average of the injury label. The average over the sum was chosen to account for variation in the number of entries for each game played. Injury occurrence decreases as the number of games played increases because the probability of them being injured in that particular season diminishes as fewer games are left. This trend is most apparent after playing in 8 games, the halfway point in the season; this feature appears to follow a normal distribution. Therefore, it makes logical sense that the model could predict that a player was not injured if they played more games in the season than other players.

Starting Position: Figure 7 shows that the position with the highest ratio of injury occurrence was the left guard position. This offensive linemen position is responsible for blocking the running back on rushing players and protecting the quarterback on passing plays. Considering this position involves physical contact on each play, the selection of this feature suggests that players involved in more contact are more prone to injury. Furthermore, the positions making contact with the left guard are the defensive end (DE), defensive tackle (DT), and occasionally a linebacker (LB). All three of these defensive positions are also within the top 10 positions with the greatest injury occurrence. Figure 7 separates the offensive and defensive positions by color, with offensive positions in red and defensive positions in blue. Out of the top 10 injury-prone positions, two are offensive (LG, TE), whereas the other 8 positions are defensive. This preliminary finding suggests that defensive players making the tackles are more likely to be injured than the players being tackled on the field.

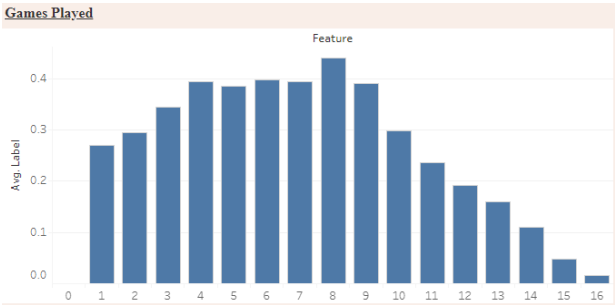

Figure 6 The average injury occurrence based on the number of games played

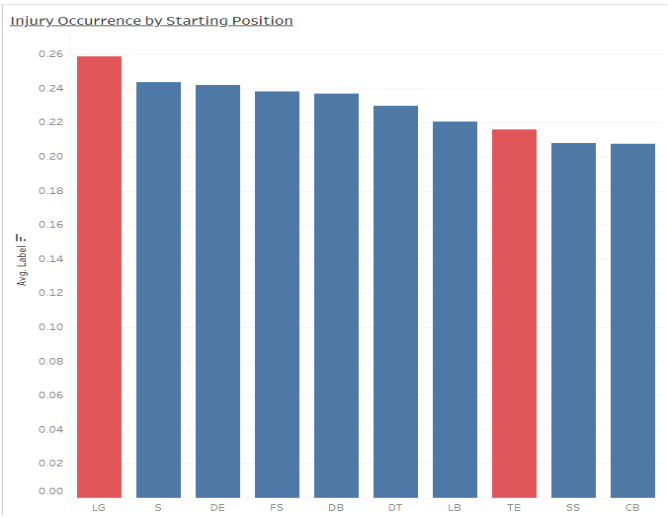

Figure 7 Injury Occurrence by Starting Position

Another important consideration is that 50% of positions with the highest injury occurrence are located in the secondary portion of the field. These positions include safety (S), free safety (FS), defensive back (DB), cornerback (CB), and strong safety (SS). These positions involve tackling players rushing up the field with speed and intercepting passes from receivers sprinting upfield. Considering these plays involve collisions and contact at high speed, these defensive positions initiating contact are more prone to injury than the offensive players being tackled. An interesting future study would be to collect more data to calculate the force of players colliding to see if the magnitude of a collision would be significant in modeling player injuries. These defensive players' injury occurrence suggests that players initiating contact are more likely to be injured than those receiving the tackle.

Conversely, when viewing players with the lowest injury occurrence, the positions include kicker (K), punter (P), and quarterback (QB). These positions were expected to have the lowest injury occurrence because they are football's most well-protected positions. These players receive minimal contact due to league rules restricting how they are tackled. Also, they have numerous players, such as the left guard, whose job assignment is to protect them from contact. Therefore, the rate of injury for well-protected positions supports the idea that more physical contact required by each position plays a role in injury occurrence.

Experience and Age: Figure 8 shows the impact of age and experience on a player's injury. The y-axis lists the years of experience, whereas the x-axis contains the age of players. Figure 7 uses the average of the injury label (1 = injured, 0 = not injured) to visualize injury occurrence depending on age and experience. The combination of age and experience with the highest probability of injury occurs at age 33 with 13 years of experience. Based on the dataset, 66.7% of players that fit this criterion were injured between 2016 and 2019. The general trend appears to be that as a player gets older and gains more experience, their probability of injury increases.

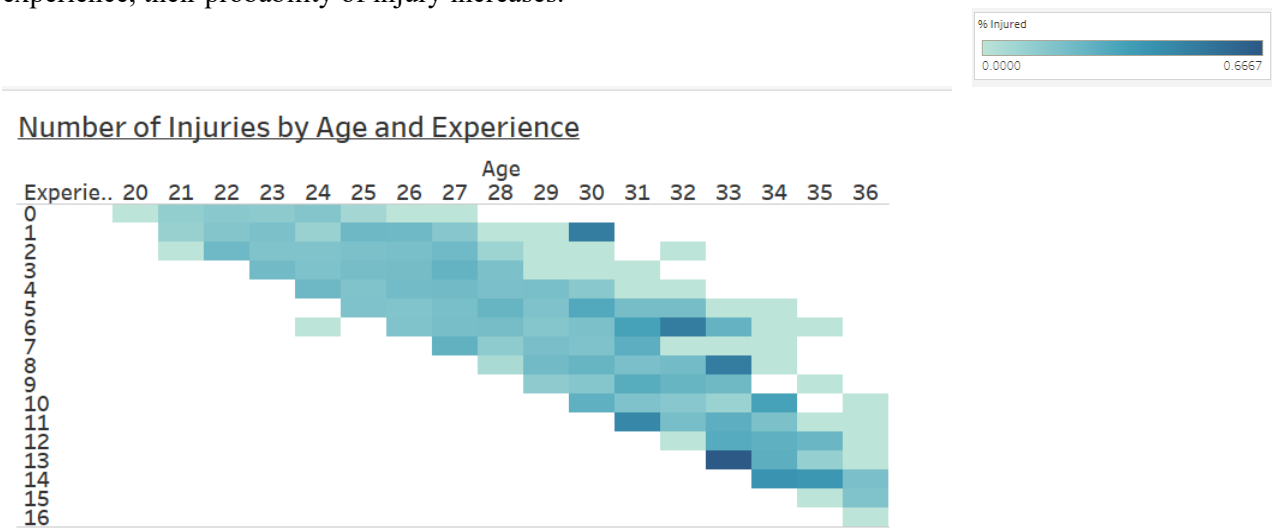

Figure 8 injury occurrence based on age and experience of players. Darker color represents a higher probability of injury

Conclusion, Limitation and Recommendations for Future Study

This study aimed to understand the factors leading to a player's injury using various machine learning models. The model with the best performance was the Gradient Boosted Trees model, with an F1 score of 0.508. In addition, the Gradient Boosted Trees model selected teams, games played, and starting position to be the most influential factors when determining the probability of a player getting injured during the season. The most notable takeaway from this study is that the left guard position is the most injury-prone, followed by defensive secondary player positions such as safety and cornerback. Future NFL rule changes and adaptions should focus on these players in the positions highlighted in the study to improve their safety.

One limitation of this research is that the model only considers a limited number of features due to the limitation of the data and time constraint. The model can be improved by incorporating more features. It is possible that previous injuries of a player and a particular match and range may have an impact on the injury of a player. In addition, further study should assess the team's performance metrics, such as win-to-

loss ratio and salary cap. Future studies should account for the temperature and weather conditions of the stadium where each game injury occurred. In addition, individual player statistics should also be considered as possible injury-causing variables.

References

Alizadeh, H. V., Fanton, M. G., Domel, A. G., Grant, G., & Camarillo, D. B. (2019, October). Prevention of Traumatic Brain Injury with Liquid Shock Absoprtion. *Journal of Biomechanical Engineering*, 1-9. Retrieved March 7, 2021

Arthur Lazarus, M. M. (2011, Jan/Feb). NFL Concussions and Common Sense. *Physician Executive, 37*(1), 6-7. Retrieved March 7, 2021

Arulanantham, A., Anzar, A. A., Rajalingam, G., Ingran, K., & Haddela, P. S. (2019, September 25). Effective Mechanism to Mitigate Injuries During NFL Plaus. *Conference on Industrial and Information Systems*. Retrieved March 8, 2021

Baker, H. P., Lee, K., Dayton, S., Terry, M., & Tjong, V. (2019, March 8). Thursday Night Football's impact on all-cause injuries in NFL players during 2012-2017. *The Physician and Sportsmedicine*, 350-352. Retrieved March 7, 2021

Cai, Y., Wu, S., Zhao, W., Li, Z., Wu, Z., & Ji, S. (2018). Concussion classification via deep learning using whole-brain white matter fiber strains. *PLOS ONE, 13*(5). doi:10.1371/journal.pone.0197992

Cohen, G. I., Lynch, H. F., & Deubert, C. R. (2016, Nov/Dec). A Proposal to Address NFL Club Doctors' Conflicts of Interest and to Promote Player Trust. *Hastings Center Report, 46*, 2-24. Retrieved March 7, 2021

Dreyer, N., Mack, C., & Anderon, R. (2019, September 1). Lessons on Data Collection and Curation from the NFL Injury Surveillance Program. *Sports Health: A Multidisciplinary Approach, 11*(5), 440-445. Retrieved March 8, 2021

Krill, M. K., Borchers, J. R., Hoffman, J. T., & Hewett, T. E. (2017). Effect of position, time in the season, and playing surface on Achilles tendon ruptures in NFL games: a 2009-10 to 2016-2017 review. *The Physician and Sportsmedicine, 45*(3), 259-264. Retrieved March 7, 2021

Mack, C. D., Kent, R. W., & Coughlin, M. J. (2020, June 2). Incidence of Lower Extremity Injury in the National Football League: 2015 to 2018. *The American Journal of Sports Medicine*, 1. Retrieved March 7, 2020

Mall, N. A., Buchowski, J., & Zebala, L. (2012, May 30). Spine and Axial Skeleton Injuries in the National Football League. *The American Journal of Sports Medicine*, 1. Retrieved March 7, 2021

Ruestow, P., Duke, T. J., Finley, B. L., & Pierce, J. (2015, Decemeber). Effects of the NFL's Amendments to the Free Kick Rule on Injuries during the 2010 and 2011 Seasons. *Journal of Occupational and Environmental Hygiene*, 875-882. Retrieved March 8, 2021

Siegel, B. (2019). Concussions and Capital: Tom Brady, CTE, and the NFL's Crisis of Identity. *Journal of Sport and Social Issues, 43*(6), 551-574. Retrieved March 7, 2021

Villwock, M. R., Meyer, E. G., Powell, J. W., & Fouty, A. (2009, March). Football playing surface and shoe design affect rotational traction. *Am J Sports Medicine*, 518-525. Retrieved March 7, 2021

Wang, D., Weiss, L., Abrams, M., Barnes, R., Warren, R., Rodeo, S., & Taylor, S. (2018, December). Athletes With Musculoskeletal Injuries Identified at the NFL Scouting Combine and Prediction of Outcomes in the NFL: A Systematic Review. *Orhopedic Journal of Sports Medicine*. Retrieved March 7, 2021

Woods, A., Chen, S. S.-C., Chen, C.-y., & Street, E. (2018). A Case of Racial Tasking and Prejudice: Analyses of Race, Playing Position, Rate of Injuries, and Salaries in the NFL. *KAHPERD, 55*(2), 28. Retrieved March 7, 2021

Zvijac, J., & Toriscelli, T. (2013, May 28). Isokinetic Concentric Quadriceps and Hamstring Strength Variables From the NFL Scouting Combine Are Not Predictive of Hamstring Injury in First-Year Professional Football Players. *The American Journal of Sports Medicine*, 1. Retrieved March 7, 2021