Expected punt outcomes in American Football

Digant Jani, Ryan Keeney, Salman Saeed, Shainu Prakash, Thanigaivel Shanmugam, Thang Trieu

Introduction and Motivation

In football, special teams are composed of members of the roster and they are placed in the game during kicking plays. Members of this group include kicker, punter, returner, and any members who are involved in the kick and return coverage [1][2]. Our objective is to utilize time-continuous and incorporate spatial information - a first for special teams players to develop a novel approach that is suited for this play type and use it to evaluate the players and strategies that have not been studied previously at this level. We will further expand our evaluation to include players who are directly impacting the outcome of the play but may not have stats associated with their performance.

The motivation behind our objective is to bring to light the importance of special teams and by analyzing their plays, we hope to use the results to the team's advantage. For our analysis, we will focus on one type of special team plays, punts, and attempt to predict the starting field position of the receiving team's offense. We will leverage player tracking data from the NFL and player performance grading from Pro Football Focus (PFF) to develop a framework for analyzing punts [3] [4] including sub-models whose results are connected to an interactive Shiny web application running on an Amazon EC2 linux server.

Problem Definition

Historically NFL analysis, both mainstream and academic, has relied on public, play-level data to generate team and player comparisons[12]. Recently the NFL released player tracking data, analysts could use this data to predict how players may perform during a game to better use them for success. Our focus was mainly on special teams as they are one of the most overlooked aspects of football. Special teams account for 30-35% of points scored in a season. Currently, football teams and scores primarily use film reviews to evaluate players and strategize. This approach can be a little risky as there are limited datasets and the person reviewing the tape must have years of experience to effectively assign grades to players. A poor choice in ranking the punters could result in a higher-scoring chance for the opposing team.

Our mission is to build a system which assesses punting at different time periods of play and predicts the expected return distance of the team receiving the ball. This analysis can later be used to improve the punting success, provide insights into the play-by-play data of a player to maximize the effectiveness of the special team, and rank players accordingly.

Survey

The punt is the most important play in football. While looking towards the revolutionary trajectory of the research on football and new technologies have made it possible to make a significant advancement in strategizing the game of Football. There are various articles published similar to our approach confirming the possibility of this project to be very likely to be successful in this field. John R. Olson and Gary R. Hunter collaborated to write a paper on anatomic analysis of punts which in-depth shows how the punter's punt to maximum distance with their body placement and strength combined. [3] There has been a development in kinematic analysis of punts based on the level of kick efforts by students in Japan. They can collectively predict the velocity of the ball based on how the angle and body were placed and it also helped in improving the distance and velocity of the ball. [4]. The study says that punts are a highly impactful part of the game and it's been measured by 33rd Team Studies. [6] There has been research into the outcome of an event at continuous time, using player tracking data it needs a general framework for continuous-time within-play valuation. [7] Similarly there are many other projects like Expected Hypothetical Completion Probability (EHCP) which is an objective framework for evaluating plays [8] have invested a lot in the similar field. Even the NFL runs an annual bigdata bowl where they allow analysts to contribute into the game and try to come up with innovative ideas. [13] Similarly other games have contributed immensely to football like Large-Scale Analysis of Soccer Matches. [16] There is multiple research on Regression applied in the outcome of games. By University of New Hampshire's student Stephen Bouzianis. [17] and also the Game day outcome paper. [18] [19] We also need to consider risk management while considering this so there are multiple papers written on Risk management similar to one written by T. Raz and E. Michael from Tel Aviv University. [20] The model we build is based on a majority of factors to predict the outcome of Punts which will be analyzed and then performed with a few optimizations so we can get accuracy on the outcome.

Proposed Method

Our analysis will utilize continuous-time assessments and will incorporate spatial information through player tracking data - made publicly available for the first time for special team plays. The NFL's tracking system captures the position and trajectory of all special team players and the ball through RFID at a frequency of 12.5 Hz, or 12.5 times per second [9]. Our aim would be to predict the expected punt distance, return type and return distance at various points during a play. The success of the model would be in terms of predicting the expected end-of-play yard line.

To establish a baseline, a Pre-Snap Model was developed which predicts the starting field position of the receiving team before the play occurs without the use of player tracking data. For this analysis, a random decision forest model was selected due to it's low bias. A random tree model provides a robust solution for evaluating high-dimensionality, non-linearity, and interacting data, and is simple to implement.

For punting, we would need to consider a classification model [8][9][10], and estimate the probability of a punt being returned, downed, fair-caught, kicked out of bounds or muffed, i.e. the punt receiver successfully catches the ball. The probability of each outcome could be estimated using a boosted random forest non-parametric model [8] which will take inputs such as the punt receiver's speed and distances to the ball and the nearest defender.

Additionally, a second "post-snap" model will assume the punt returner has fielded the ball and predict the expected end-of-play yard line. One way to do this is to use a long short-term memory (LSTM) recurrent neural network to construct a model which estimates the number of yards the punt receiver would be expected to gain from their current position [7]. This would be subject to the locations and trajectories of the punt receiver, teammates and opposing team members. A LSTM model provides a robust solution for evaluating high-dimensionality, non-linearity, interactions, and temporal structure of the data. The team explored a LSTM model but ultimately chose to use a decision-tree-based ensemble algorithm that had a lower complexity to implement and provided a robust prediction.

A decision-tree-based ensemble Machine Learning algorithm that uses a gradient boosting framework will also be evaluated via the XGBoost package in R. A boosted tree model provides a robust solution for evaluating high-dimensionality, non-linearity, and interacting data, and is easier to implement. However, it does not address the temporal structure of the data; but, it is simpler to implement and has shown to be successful in previous applications [7].

The sequence would begin at the point where the punt receiver receives the ball after the punt. The end of the sequence would be when the punt receiver is either tackled, runs out of bounds or fumbles. All players would be split into either defense or offense. The NFL has provided tracking data for all players on the field but rather than using raw tracking coordinates for the players, the accurate way to estimate the expected end-of-play yard line would be to calculate coordinates and direction of each player on the field with respect to the punt receiver's position and target endzone.

Previous analysis [7] has shown that the most important features on the accuracy of the results are the punt receiver's speed and the distance that the punt receiver is to the nearest defender. This is consistent with intuition that the faster the punt receiver moves and the farther he is from defenders, the further the expected end-of-play yard line.

The results will then be collected and displayed in a Shiny R application running on a Linux Ubuntu EC2 server from AWS. The interface will allow users to dynamically explore the change in return probability and expected return distance dynamically throughout a play. Additionally, the results collected from the players performance-vs-expectation will provide a framework to evaluate and compare players.

Key Innovations

Our approach significantly expands upon the current approaches and tools available publicly:

1. Continuous predictions: Previous evaluations techniques rely on play description and result statistics. Our approach provides improved accuracy and predictions that react to what is happening through the play.
2. Non-primary player evaluation: Previous evaluation techniques are focused on the primary players who impact a play, the punter, returner, and tackler. Our approach broadens the evaluation to include secondary plates who may not appear a recorded stat but did affect the plays outcome.
3. Probability modeling and evaluation of hypotheticals: By using a pre-snap, post-snap-classification, and post-snap-tracking model, our approach provides more detailed information and predictions about what is occurring during a play, and also allows for evaluation of hypothetical scenarios (e.g. Should the punt returner have fielded the ball in this situation?).
4. User interface: Our user interface would be the first publicly available tool to display the continuous prediction of punt outcomes as well as providing a framework for evaluating gunners and jammers.

If successful, this framework could be applied to other special teams plays such as kick-off returns. Additionally, our approach could be adopted for use in player evaluation, contract negotiation, and media.

Experiments / Evaluation

Collecting Data

Datasets were collected via downloads provided by the NFL, or were available publicly:

1. Player Tracking Data: Ball and player location and trajectories collected via RFID at a frequency of 12.5 Hz during each play.
2. NFL Play-By-Play Data: Play descriptions provided by the NFL, including the starting field position, result, and players involved in the play.
3. NFL game Data: NFL game information, including team, location, and game outcome.
4. Lee Sharpe game Data: Detailed NFL game information, including team, location, weather, game outcome, and betting market information.
5. PFF Play-By-Play Data: Detailed play descriptions provided by PFF, including descriptions of personnel, scheme, and results for each play.

After collecting the data, we merged, filtered, and cleaned the datasets, resulting in a play-by-play dataset used for the pre-snap model and player tracking data used for the post-snap models. Additionally, reference tables were created to link player and team tags across datasets. There are five-thousand valid plays in our dataset, resulting in over 13 million data points in the tracking data.

Completed Evaluation and Experimentation

To establish a baseline, a Pre-Snap Model was developed which predicts the starting field position of the receiving team before the play occurs without the use of player tracking data. A simple random tree model using the factor set: distance to the opponent's end zone, number of gunners, number of rushers, and number of jammer resulted in a RSQ 0.56, RSME of 9.97, and MAE of 6.88 on a 5-fold cross validation split 80/20 (training/testing). Our team will utilize this prediction to evaluate the added benefit of our post-snap model.We performed another approach called the Post-Snap model where we used player location, velocity and continuously predicted the starting field position of the receiving team during play. Within the post-snap model there are two primary sub models. First, a classification model which predicts the probability of each type of return occurring: Return, touchback, fair-catch, downed, muffed or out of bounds. Second post-snap model predicts the final field position of the receiving team assuming the punt is returned. The classification model gives the probability of play given location, trajectory of the ball, return and closest defender. Return model predicts using the ball carriers, and other players, location and trajectory. This will evaluate multiple independent and dependent model aspects of the model and predict the final output of punts using that output. Based on Analysis we performed a few experiments and evaluated models based on it to find out what is needed for optimization and tuning the approach. Finally we created a website where you can see play-by-play evaluation and return demos.

<u>Experiment Questions and Answers</u>

1. How accurately can we predict punt outcome locations without player tracking data?
    a. What are the most important factors for these predictions?
2. How accurately can we predict the punt outcome classification with tracking data?
    a. Can we visualize it over the duration of a play?
    b. What factors are important to this prediction?
3. How accurately can we predict the punt outcome locations using player tracking data?
    a. How does it compare with the baseline when the ball is fielded?
4. How can we compare and rank players with these models?

Using the both models, we can predict the punt outcome with a 91.1% accuracy at the time of the punt for plays from the test data. A visualization of the post-snap regression is shown in Figure 1 and the probability of different play outcomes is shown in figure 2. Predicting continuous outcomes provides deeper insight into player performance and in-play dynamics.

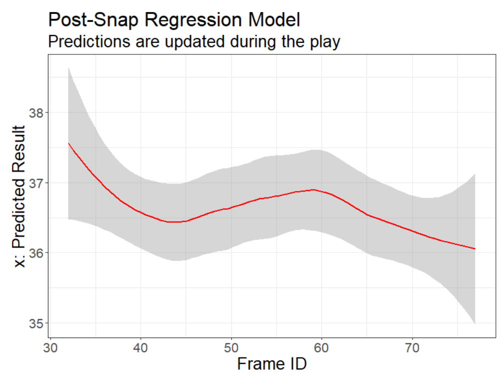

Figure 1

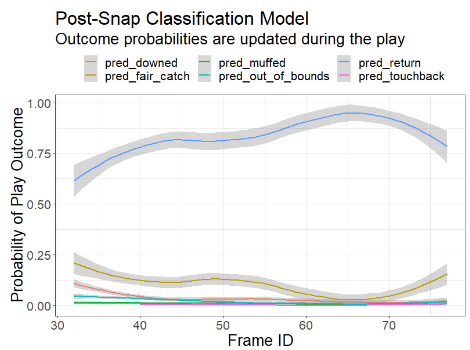

Figure 2

Our algorithms utilize player tracking data, including: starting location, ball location, relative location of the closest defender and returner and their orientation, velocity, and heading.

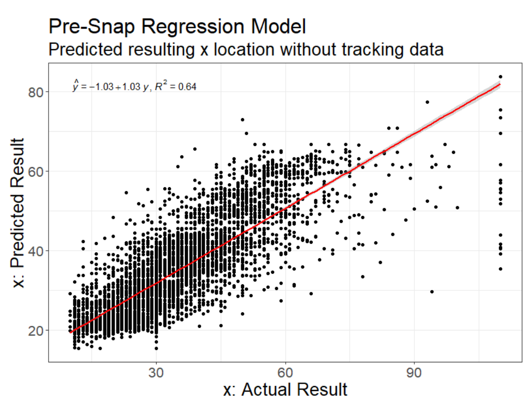

Figure 3

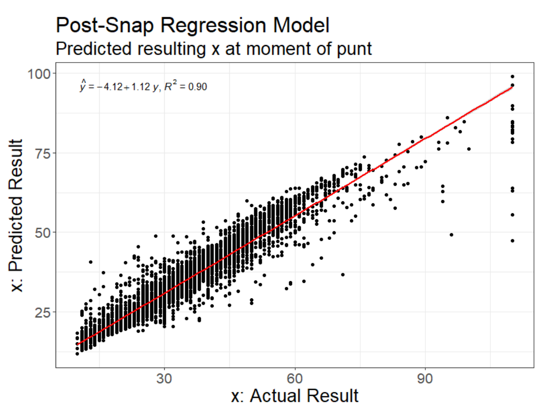

Figure 4

By comparing our approach with the baseline model, the RSQ increases from 0.64 to 0.90. This is also by assuming we only measure at the start of the beginning of the punt, when predictions are updated throughout the play.

Plan of Activities

Data identification, exploration, profiling, enhancing data quality, feature engineering
Identified and prepared: NFL Player tracking data, NFL play by play data, NFL game data, Lee sharpe game data, PFF play-by-play data. In progress by Salman and Shainu (100% completed on 11/08).

Model identification, exploration
Explored Spatial analysis techniques, Passing model [8][9][10], Bayesian non-parametric model [8], next-gen-scarPy (K-means++ clustering approach),Classification model [8][9][10], LSTM recurrent neural network, Tensorflow and keras packages, Decision-tree based model, that uses a gradient boosting framework. In progress by Thanigaivel and Thang (100 % completed on 11/10).

Modelling
Using Classification model [8][9][10], LSTM recurrent neural network, Tensorflow and keras packages and Decision-tree based model that uses a gradient boosting framework. In progress by Thanigaivel and Thang (100% completed on 11/23).

Model performance evaluation and representation
Evaluate the performance of the pre and post-snap models and related submodels. In progress by Salman and Digant (100% completed on 11/26).

UI design and development
Develop the Shiny R application. In progress by Shainu and Ryan (100% completed on 11/28).

Outcome Analysis and Visualization creation
Create the supporting visualizations for the UI, demonstrations, and model evaluation. In progress Thanigaivel,Thang, Digant, Ryan (100% completed on 11/30).

All team members contributed a similar amount of effort [Appendix: Table 2].

Conclusions and Discussion

By the end of the project we were able to predict the starting field position of the receiving team before the play occurs, estimate the probability of a punt being returned, downed, fair-caught, kicked out of bounds or muffed and last the expected end-of-play yard line using the spatial information available through player tracking data for special team plays. Our system is able to assess punting at different time periods of play and predicts the expected return distance of the team receiving the ball. This approach allows for additional insights into player performance and strategy. For example, the average yards over expectation can be calculated based on the predicted starting field position at the time of the punt. This analysis can later be used to improve the punting success, provide insights into the play-by-play data of a player to maximize the effectiveness of the special team, and rank players accordingly. This model is the first of its kind using spatial information and specifically targeting the special teams.

References

[1] ByCoach Martin. (2020, November 16). *What are special teams in football? (explained)*. Football Advantage. Retrieved October 10, 2021, from https://footballadvantage.com/special-teams-football/

[2] Kristan, Greg (2021, August 3rd). *What are Special Teams in Football?.* The Stadium Reviews. Retrieved October 11, 2021, from https://thestadiumreviews.com/blogs/info/what-are-special-teams-in-football/

[3] Olson, John R., Hunter, Gary R, Dunnam, Laura A. (April 1986) An anatomic analysis of football punting with implications for conditioning *National Strength and Conditioning Association Journal April 1986 - Volume 8 - Issue 2 - p 5-13* https://journals.lww.com/nsca-scj/Citation/1986/04000/An_anatomic_analysis_of_football_punting_with.1.aspx

[4] Tago T., Nishino K., Kaneko K., Wada T., Nishiyama T. & Funato K. (2014). Kinematic Analysis of Punt Kick in Football Goalkeepers Based on the Level of Kick Effort *NSSU Journal of Sport Sciences Vol.3 36-43* https://www.nittai.ac.jp/souken/katsudou/pdf/vol3_36-43.pdf

[5] Woroszylo Jr., T.R., *Rating System for Punting in the National Football League.* Ball State University M.S. Thesis. Retrieved October 11, 2021, from https://cardinalscholar.bsu.edu/bitstream/handle/123456789/198503/WoroszyloT_2014-3_BODY.pdf?sequence=1&isAllowed=y

[6] Simpson, R., & Engler, P. (2021, June 10). *Study: Impact of special teams on NFL Success*. The 33rd Team. Retrieved October 10, 2021, from https://www.the33rdteam.com/case-study-impact-of-special-teams-on-nfl-success/

[7] Yurko, R., Matano, F., Richardson, L., Granered, N., Pospisil, T., Pelechrinis, K. & Ventura, S. (2020). Going deep: models for continuous-time within-play valuation of game outcomes in American football with tracking data. Journal of Quantitative Analysis in Sports, 16(2), 163-182. https://doi.org/10.1515/jqas-2019-0056

[8] Deshpande, S. & Evans, K. (2020). Expected hypothetical completion probability. Journal of Quantitative Analysis in Sports, 16(2), 85-94. https://doi.org/10.1515/jqas-2019-0050

[9] Mallepalle, S., Yurko, R., Pelechrinis, K. & Ventura, S. (2020). Extracting NFL tracking data from images to evaluate quarterbacks and pass defenses. Journal of Quantitative Analysis in Sports, 16(2), 95-120. https://doi.org/10.1515/jqas-2019-0052

[10] Chu, D., Reyers, M., Thomson, J. & Wu, L. (2020). Route identification in the National Football League: An application of model-based curve clustering using the EM algorithm. Journal of Quantitative Analysis in Sports, 16(2), 121-132. https://doi.org/10.1515/jqas-2019-0047

[11] Dutta, Rishav, Yurko, Ronald and Ventura, Samuel L.. "Unsupervised methods for identifying pass coverage among defensive backs with NFL player tracking data" Journal of Quantitative Analysis in Sports, vol. 16, no. 2, 2020, pp. 143-161. https://doi.org/10.1515/jqas-2020-0017

[12] Lopez, Michael J.. "Bigger data, better questions, and a return to fourth down behavior: an introduction to a special issue on tracking data in the National football League" Journal of Quantitative Analysis in Sports, vol. 16, no. 2, 2020, pp. 73-79. https://doi.org/10.1515/jqas-2020-0057

[13] Howard , A. (n.d.). NFL announces fourth annual Big Data Bowl powered by AWS   // . NFL Communications. Retrieved October 10, 2021, from https://nflcommunications.com/Pages/NFL-Announces-Fourth-Annual-Big-Data-Bowl-Powered-by-AWS.aspx

[14] Ver Hoef, J.M., Cressie, N. Multivariable spatial prediction. *Math Geol 25, 219–240 (1993)*. https://doi.org/10.1007/BF00893273

[15] Yurko, R., Matano, F., Richardson, L., Granered, N., Pospisil, T., Pelechrinis, K. & Ventura, S. (2020). Going deep: models for continuous-time within-play valuation of game outcomes in American football with tracking data. *Journal of Quantitative Analysis in Sports*, *16*(2), 163-182. https://doi.org/10.1515/jqas-2019-0056

[16] Alina B., Patrick L., Peter C., Yisong Y., Sridha S., &Iain M.. Large-Scale Analysis of Soccer Matches using Spatiotemporal Tracking Data. (2014) *IEEE International Conference on Data Mining, ICDM '14, pages 725–730, Shenzen, Dec 2014*. https://ieeexplore.ieee.org/document/7023391

[17] Stephen Bouzianis (Spring 2019) :Predicting the Outcome of NFL Games Using Logistic Regression https://scholars.unh.edu/cgi/viewcontent.cgi?article=1472&context=honors

[18] Klein, Josh; Frowein, Anna; and Irwin, Chris (2018) "Predicting Game Day Outcomes in National Football League Games," *SMU Data Science Review: Vol. 1 : No. 2 , Article 6*. Available at: https://scholar.smu.edu/datasciencereview/vol1/iss2/6

[19] Constantinou, A.C. Dolores: a model that predicts football match outcomes from all over the world. *Mach Learn 108, 49–75 (2019).* https://doi.org/10.1007/s10994-018-5703-7

[20] *Defense Technical Information Center*. (n.d.). Retrieved October 10, 2021, from https://apps.dtic.mil/sti/pdfs/ADA319533.pdf

[21] Raz, T., & Michael, E. (2000, October 27). *Use and benefits of tools for Project Risk Management*. International Journal of Project Management. Retrieved October 10, 2021, from https://www.sciencedirect.com/science/article/pii/S0263786399000368?casa_token=r3DgD93rfuMAAAAA%3AITpmHkMcN7w-giOn7hRxBcAzIv5Q4-FDbMZHTzdNL4nG7Dh9PDs4UmlId5OCcN3Kwm8hnzpJDg

[22] Journal of Computer Information Systems. (n.d.). Retrieved October 10, 2021, from https://www.tandfonline.com/doi/pdf/10.1080/08874417.2007.11645981?casa_token=gpPokY5P-LwAAAAA%3AZVA4ot-_b1ZLRS7J2AkyYb2wNrRGk-LLwMp8KKGJw3iqu8XhG7HZqFYyRSXQhuogLKocox-3Qy2U

Appendix

All team members have contributed a similar amount of effort.

Task	Salman	Shainu	Thanigaivel	Thang	Digant	Ryan
Proposal	✓	✓	✓	✓	✓	✓
Progress Report	✓	✓	✓	✓	✓	✓
Final Report	✓	✓	✓	✓	✓	✓
Data Collection	✓	✓				
Data Cleaning	✓	✓				
Modelling			✓	✓		
UI		✓				✓
Model Evaluation	✓				✓	
Visualization			✓	✓	✓	✓
Finalization	✓	✓	✓	✓	✓	✓

Table 2 : Task Assignment & Progress (✓ Completed, ○ InProgress)