Neural Network Approach to NFL Position Classification

Sithija Manage[*]
Project advisor: Sai-Mang Pun[†]

June 1, 2022

Abstract

With an ever-increasing captivation of the United States sports-viewing audience, the National Football League continues to produce some of the world's most capable, physical athletes. In this work, athletes' positions C, OG, OT, DE, and DT were categorized as *on the line*, while the remaining positions were categorized as *not on the line*. In this work, a predictive neural network is applied to classify 2,022 National Football League players into the two classifications using scouting combine data of height, weight, and 40-Yard dash time, outperforming the current standard logistic regression. The two measures utilized to compare the strength of the methods were total accuracy and area under ROC curve, with the neural network yielding a slightly higher average in both. In terms of total accuracy, the neural network had an accuracy of 0.9134 to the logistic model's 0.9065, and in terms of area under ROC curve, the neural network had an area of 0.9578 compared to the logistic model's 0.9567. As a head-to-head iteration-wise comparison, the neural network had a winning Win-Loss-Tie ratio of 7-2-1 and 5-5-0 in the two measures respectively.

Some keywords: Neural Network, Logistic Regression, Classification, American Football, National Football League

1 Introduction

One of the earliest pastimes of humanity is physical sport; from the very first physical human sport of wrestling to the highly intricate game of modern American football, we have found ourselves at stadiums and arenas across the globe, captivated by the action. With the rapid development of technology, these sports have become rich in data, allowing for sports analytics to emerge and new technologies, such as machine learning, to have applications in this area. Many of the questions posed in sports analytics have binary outcomes, such

[*] Department of Mathematics, Texas A&M University, College Station, TX 77843, United States (email: sithijamanage@tamu.edu)

[†] Department of Mathematics, Texas A&M University, College Station, TX 77843, United States (email:smpun@math.tamu.edu)

as “Will this team win the championship?” or “Is a given player injury-prone” One of the most simple statistical methods to tackle such classification problems, detailed in Section 3.3, is logistic regression, which is used to explain the relationship between one dependent binary variable and one or more independent variables. Furthermore, throughout the past two decades, machine learning techniques, such as neural networks, detailed in Section 3.4, have been increasingly utilized to predict outcomes in sports as well (Bunker & Susnjak 2019). However, one key drawback of most machine learning techniques is that, unlike logistic regression models, they are very difficult to interpret. This lack of interpretability is often accepted as a tradeoff for higher accuracy. For problems that require deep interpretations, many researchers utilize methods such as logistic regression. For instance, a 1990 study by David Nieman and Jerry Lee investigated the relationship between “infectious episodes” and running training intensity in Los Angeles Marathon runners with logistic regression models (Nieman & Lee 1990). Likewise, a 2001 study by Di Zhong and Shih-Fu Chang analyzed sports video using supervised machine learning techniques, an area that has found increasing relevance in training facilities today (Zhong & Chang 2001).

This research work aims to merge machine learning and sports in an intuitive way, through a modern classification problem. Specifically, this work involves data from the National Football League (NFL). When surveying an American football field during a game, one may marvel at the range of physical builds of the players. Some players are fast and agile, whereas others may be slow and powerful. The optimal physical builds of football players seems Darwinian, with the positions that require craftiness and agility filled by seemingly small, quick players and the positions that require blocking other players or physically outmaneuvering the opponent filled by large, strong men. More specifically, the positions around the line of scrimmage (offensive line and defensive line) yield the biggest, strongest players - justified by their task of physically overtaking their opponents.

2 Background

Offensive linemen have two major roles in the sport of football - guarding the quarterback (passer) from opposing players and controlling the line of scrimmage for the surrounding teammates such as the running back. Directly opposed to these players are the defensive equivalent: defensive linemen. These linemen have two major roles as well - getting past the offensive linemen to tackle the quarterback and more generally preventing the offense from scoring (Alvarez 2011). The remaining positions in football have a diversity of roles, and one aim of this research is to observe whether or not this intuitive difference in categorizing positions as linemen or not-linemen can be modeled in machine learning and logistic regression by just using height, weight, and speed data.

Machine learning refers to the process by which a computer system utilizes data to train itself to make better decisions, and along with data science, it is transforming decision-making in sports (Pickering n.d.). The machine learning method used in this research is further discussed in Section 3.4. Current applications of machine learning in sports include predicting recovery times, injury probabilities, and play calls (NFL 2022, R 2018). Furthermore, a more standard approach to binary classification tasks, logistic regression is also employed to compare the results of the machine learning algorithm's performance, which is discussed in more detail in Section 3.3.

2.1 Contribution and Existing Work

The aim of this study was to observe whether a predictive neural network could be applied to correctly classify NFL players into their respective *on the line* and *not on the line* categories using just height, weight, and speed data, with the benchmark goal of outperforming the potentially more standard approach, logistic regression. The details of the classification task are given in Section 2.1.

There have been several research works that employ machine learning to predict player *success* and team success in American Football such as Namhoon Lee and Kris Kitani's *Predicting wide receiver trajectories in American football* (Kitani 2016) and Yu-Chia Hsu's *Using Machine Learning and Candlestick Patterns to Predict the Outcomes of American Football Games* (Hsu 2020) . A similar study was conducted by Arpitha et.al recently in which the researchers analyzed anthropometric, physiological, and motor fitness parameters with the aid of machine learning models to predict the best suitable playing positions for basketball players (Arpitha et al. 2021). Furthermore, this approach of applying machine learning tools to predict player *position* classifications has been utilized in other sports such as soccer and international basketball, including Garcia-Aliaga et al.'s *In-game behaviour analysis of football players using machine learning techniques based on player statistics*, published in *International Journal of Sports Science & Coaching* (Garcia-Aliaga et al. 2020) and *Position-specific performance profiles, using predictive classification models in senior basketball* by Pion et al. (Pion et al. 2018). Another work, published in *Journal of Sports Analytics* in 2020 developed an interpretable machine learning model to predict offensive plays (pass versus rush) in the NFL (Fernandes et al. 2020). Kuzmits and Adams' 2008 publication in *the Journal of Strength and Conditioning Research* also shed light on the relationship between NFL players' combine data and performance in the league (Kuzmits & Adams 2008) . However, this research is a unique study in the space of applying machine learning to predict NFL player positions.

3 Methodology

3.1 Classification Task

Throughout this work, these players with positions on the offensive line (center, offensive guard, offensive tackle) and the defensive line (defensive end, defensive tackle) are categorized as "on the line," and all other players are categorized as "not on the line." In this study, using the height, weight, and forty-yard dash time of players, the aim was to apply a predictive neural network to correctly classify American football players into their respective "on the line" and "not on the line" categories, with the goal of outperforming the existing standard approach, logistic regression.

3.2 Data

Data was retrieved from `https://stathead.com/`. This database includes a searchable index of NFL Draft Combine results since 2000. The Draft Combine, or National Scouting Combine, is a four-day event in which that year's top draft-eligible players are evaluated on a variety of medical, mental and physical criteria. It is from this draft that the heights, weights, and speeds of players are retrieved for this analysis. As Norton and Olds' work on the morphological changes in modern athletes suggests, the physical attributes of athletes are increasingly converging to the optimal qualities (Norton & Olds 2001). This indicates that in general, the modern athlete is more physically suited for the sport he or she plays than a comparable athlete in a previous era. Thus, with such significant potential for difference (across decades) in mind, in this study, data from the years 2011 - 2021 was used (a span of 11 years). As stated previously, variables used for this analysis included player height (in inches), player weight (in pounds), and player 40-Yard Dash time (in seconds). These variables were selected as although other variables are available in the StatHead data set, such as vertical jump, bench press, broad jump, 3 cone drill time, and shuttle time, these contained many missing entries. On the other hand, height, weight, and 40-Yard dash time are key metrics that are recorded for nearly every player. There were a total of 2,022 players analyzed in this study, and 10-fold cross-validation was utilized to train and test the 10 neural networks and 10 logistic regression models. Thus, the training and testing sets are the same for both algorithms.

3.2.1 Statistical Summary of Data

The following tables depict a simple statistical summary of the variables and their pairwise correlations, respectively.

Variable	Mean	StdDev
Height	73.8062	2.7050
Weight	242.7956	45.1072
40Yard	4.7623	0.2973

Table 1: Statistical Summary of Predictors

	Height	Weight	40Yard
Height	1	0.6943	0.5906
Weight	0.6943	1	0.8592
40Yard	0.5906	0.8592	1

Table 2: Pairwise Correlations of Predictors

3.3 Logistic Regression

One popular method of tackling such a problem of predicting a binary response from multiple input variables is logistic regression. Logistic regression is a transformation of a linear regression using the sigmoid function, and the response variable is given as a probability of classification into one of the binary categories. The logistic function is given by the following equation:

$$\theta(x) = \frac{exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k)}{1 + exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k)} \tag{1}$$

where $\theta(x)$ is the probability that the response variable is 1, and $x_1, \cdots, x_k$ are the input variables with coefficients $\beta_1, \cdots, \beta_k$ (Sheather 2009). Note that $\theta(x)$ is equivalent to $P(Y = 1)$, with Y being the response variable.

In logistic regression, a linear relationship between the predictor variables and the log-odds (or logit) of the event that $Y = 1$ is assumed. Mathematically, this assumption is

$$\ell = \log\left(\frac{\theta(x)}{1 - \theta(x)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k \tag{2}$$

Logistic Regression allows some interpretability of the produced model in two manners. Firstly, the β coefficients can be interpreted as a multiplier when they are used as exponents of Euler's number, e. For example, a β coefficient of 1 indicates that for every increase of 1 unit in the corresponding predictor variable, the odds of the response variable being 1 increases by e. Secondly, the significance of the variables as inputs can be determined. Thus, the variables that contribute highly to the model can be identified (as done in Section 4.3).

The cost function minimized in logistic regression is as follows:

$$Cost(\theta(x), y) = \begin{cases} -\log(\theta(x)) & \text{if } y = 1, \\ -\log(1 - \theta(x)) & \text{if } y = 0. \end{cases} \tag{3}$$

which can be expressed as

$$Cost(\theta(x), y) = -ylog(\theta(x)) - (1 - y)log(1 - \theta(x))$$

where y is the true categorization of a given player, x (1 if *on the line*, 0 otherwise) (Sheather 2009).

3.4 Neural Network

Neural Networks are computer systems modeled on the human brain and nervous system (*Neural Network* n.d.). These neural networks are structures that contain nodes and connections that pass information from input to an output. Each of these nodes takes a value from zero to one that indicates the strength of the information that it passes. These nodes are arranged in layers, yielding a general mathematical form of

$$F(x) = \sigma(W^{[m]}\sigma(W^{[m-1]}...\sigma(W^{[2]}x + b^{[2]})... + b^{[m-1]}) + b^{[m]}) \in \mathbb{R}^2, \tag{4}$$

where m is the the number of layers, W and b are the weights and biases respectively (Higham & Higham 2019). These parameters are the *learnable parameters* of the model, meaning that they are adjusted during the training process to minimize some cost function. Biases transform the input to a node by adding or subtracting from Wx, the input transformed by the weight. The weight, W, influences the steepness of the sigmoid function mentioned below, determining the amount of influence an input has on the output. σ is the activation function. These activation functions are mathematical functions that help the algorithm learn complex patterns in data, and there are a variety of options for such functions. In this study, the *sigmoid* function is used:

$$\sigma(x) = \frac{1}{1+e^{-x}}.$$

The sigmoid function has the property that

$$\sigma'(x) = \sigma(x)(1-\sigma(x)),$$

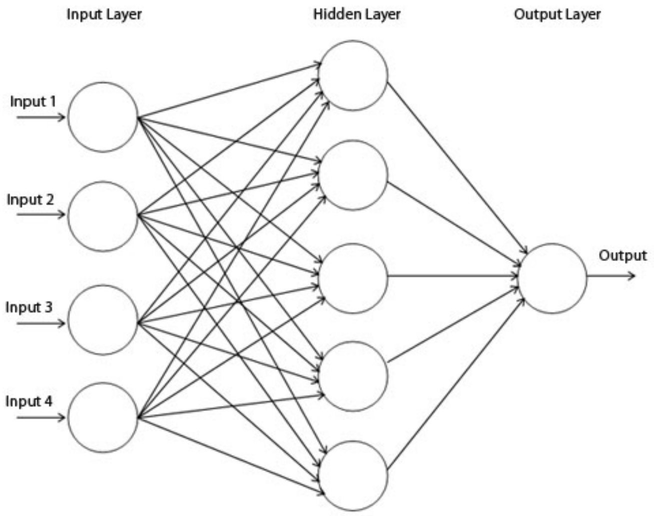
a very simple form for its derivative. Using standard transformations, the steepness and location of the function can be varied as mentioned above (with W and b) (Higham & Higham 2019). Each “node” or “neuron” outputs a single real number (from the sigmoid function), which is then inputted into the next neuron. Sets of these neurons compose the aforementioned layers. To represent the weight that is used to modify the function and the bias, W and b are used respectively:

$$\sigma(Wa+b)$$

Because each sigmoid function takes input that has passed through all of the previous layers, the function becomes nested, eventually passing through the n^{th} function. In a neural network with four layers, it reads

$$F(x) = \sigma(W^{[4]}\sigma(W^{[3]}\sigma(W^{[2]}x + b^{[2]}) + b^{[3]}) + b^{[4]}) \in \mathbb{R}^2.$$

A visual depiction of single-layer neural network with 5 hidden nodes is shown below (Manzini n.d.):



For a neural network with four layers, all parameters that define $W^{[2]}, W^{[3]}, W^{[4]}, b^{[2]}, b^{[3]}$, and $b^{[4]}$ must be optimized over.

The cost function utilized in this project is the default function in the *neuralnet* package, sum of squared errors:

$$Cost = \frac{1}{2}\sum_{i=1}^{n}(y(x_i) - \phi(x_i))^2$$

where x_i is a given input vector of height, weight, and speed, $\phi(x_i) \in \{0,1\}$ is the category output from the algorithm, $y(x_i)$ is the true categorization of the $i^t h$ player (1 if *on the line*, 0 otherwise), and n is the number of players in the data set.

3.5 Application of Logistic Regression

The logistic regression models were programmed using R programming language. After dividing the data into the appropriate training and testing blocks, a new column of data, "onTheLine" was created to denote whether the player's position was "on the line" or not. In this study, a value of 1 represents on the line, and 0 represents otherwise. Using the *glm* function in the *stats* package in R, the logistic regression models were fit with weight, height, and forty-yard time as predictor variables and onTheLine as the response. The probabilities predicted by the model were then applied to the testing set, rounding at 0.5. Thus, if the predicted probability from the model was > 0.5, onTheLine $= 1$. Otherwise, onTheLine $= 0$. Subsequently, the true onTheLine values for the testing set were compared to the predicted values to retrieve the total accuracy. In the final computation, the *auc* function of the *pROC* package in R was used to calculate the area under the ROC curve.

In this case, $x_1 =$ Player Weight (in pounds), $x_2 =$ Player Height (in inches), and $x_3 =$ Player Forty-Yard Dash Time (in seconds):

$$\theta(x) = \frac{exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)}{1 + exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)} \tag{5}$$

3.6 Application of Neural Network

The neural networks were programmed using Python. After dividing the data into the appropriate training and testing blocks, the "onTheLine," variable was created. Using the $tensorflow$ package, the neural networks were fit utilizing weight, height, and 40-yard time as predictors and onTheLine as response. The prediction step was conducted using the $compile$ and fit functions, and in this method as well, probabilities were rounded at 0.5. In the same fashion as for logistic regression, the total accuracy and area under ROC curve were computed for each iteration.

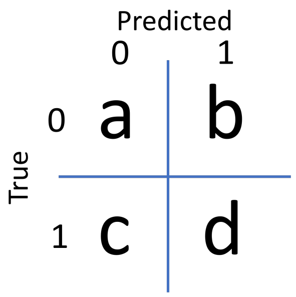
3.7 Model Validation

3.7.1 *K*-Fold Cross-Validation

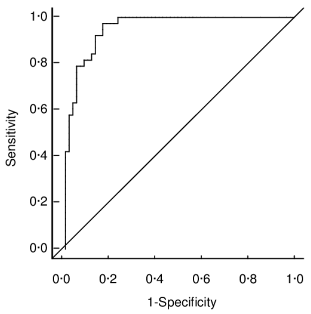
In a typical prediction model scenario, one may split a data set into a training set and a testing set, using the training set to train the model and the testing set to evaluate it. However, this method may not be reliable for evaluating the performance of a model as the results may vary significantly based on the specific selection of training v.s. testing data. Thus, in order to evaluate the model's performance in a more reliable and rigorous fashion, K-Fold Cross-Validation is used. In K-Fold Cross-Validation, a given data set is split into K groups, and through K iterations, each individual group is evaluated as the testing set while using the others to train the model. In this study, $K = 10$, indicating that in the first iteration, the first tenth of the data (by row) is set aside as the testing set. Then, the remaining nine tenths are used to train the model. The trained model is finally evaluated using the testing set. This entire process is done once for each model (neural network and logistic regression) and has 10 iterations of cross-validation within each. Comparing the model produced by the neural network and the model produced by logistic regression at each iteration allows for a valid and clear conclusion of which model is stronger.

3.7.2 Total Accuracy & ROC Curves

Both the neural network's and logistic regression model's efficacy was tested using Total Accuracy of prediction and area under ROC curve. Total Accuracy is the number of correct classifications (into onTheLine and not-onTheLine) out of the total number of classifications (equivalently, total number of players). This measurement gives an estimation for the overall strength of prediction by the model. Likewise, ROC curves utilize two alternate measures for prediction efficacy - sensitivity and specificity. The following diagrams help illustrate these measures. The first is the "confusion matrix" used to generate the ROC curves in this research, and the second is an example of an ROC curve, showing the relation between sensitivity and 1-specificity (Ormerod & Williams 2002):



		Predicted	
		0	1
True	0	a	b
	1	c	d



Sensitivity denotes the percentage of truly onTheLine players being selected $(d)/(c+d)$, and specificity denotes the percentage of truly not-onTheLine players being identified as not-onTheLine $(a)/(a+b)$. ROC curves utilize different cut-off values for the probability of a player being predicted as onTheLine to plot a graph that displays Sensitivity v.s. 1-Specificity. A stronger "arch" of the curvature on an ROC curve indicates more accurate prediction. Furthermore, an area under the curve of 1 represents a perfect test while an area of 0.5 represents a valueless test (*Area Under an ROC Curve* n.d.). Hosmer et al. provide the following rule of thumb used in this study in *Applied Logistic Regression* (Hosmer et al. 2013):

$$
\begin{gathered}
.90\text{-}1 \rightarrow \text{outstanding discrimination}\\
.80\text{-}.90 \rightarrow \text{excellent discrimination}\\
.70\text{-}.80 \rightarrow \text{acceptable discrimination}\\
.60\text{-}.70 \rightarrow \text{poor discrimination}\\
.50\text{-}.60 \rightarrow \text{no discrimination}
\end{gathered}
$$

4 Experimental Results

Results are presented iteration-wise for clear comparison between this study's neural network and the standard logistic regression model. An overall comparison summarized in tables is presented at the end of this section. The following "iterations" are each of the ten blocks from 10-fold cross-validation introduced in Section 3.5.1. For each iteration, total accuracy and area under the ROC curve are compared, and both methods are visually represented. The ROC curves are plotted for each method as well, and the R coefficient output for each iteration of the logistic model is given in Appendix A

4.1 Summary of Results

The overall Total Accuracy and Area Under ROC Curve are summarized in the following tables:

Total Accuracy		
Iteration	Neural Net Acc.	Logistic Reg. Acc.
1	0.9113	0.8911
2	0.9310	0.9059
3	0.9109	0.9059
4	0.9257	0.9257
5	0.8960	0.8911
6	0.9257	0.9158
7	0.9109	0.9059
8	0.9208	0.9257
9	0.8812	0.8867
10	0.9208	0.9113
Average	**0.9134**	**0.9065**
Std.Dev.	0.0144	0.0138

Table 3: Total Accuracy

Area Under ROC Curve		
Iteration	Neural Net AUC	Logistic Reg. AUC
1	0.9598	0.9519
2	0.9653	0.9671
3	0.9571	0.9579
4	0.9784	0.9745
5	0.9436	0.9425
6	0.9476	0.9487
7	0.9691	0.97
8	0.9663	0.9597
9	0.9286	0.9329
10	0.9618	0.9616
Average	**0.9578**	**0.9567**
Std.Dev.	0.0136	0.0129

Table 4: Area Under ROC

On average, the neural network has a higher total accuracy across all $k-$fold cross-validation iterations, and the neural network has a higher area under ROC curve across all $k-$fold cross-validation iterations. Consequently, on average, the neural network outperforms the existing logistic regression method.

The following table displays the best method with regards to Total Accuracy and Area Under ROC Curve for each iteration. For instance, because the total accuracy of the neural net in the first iteration is ≈ 0.91, and the total accuracy of the logistic model is ≈ 0.89, "NeuralNet" is displayed under Total Accuracy for iteration 1 ($0.91 > 0.89$):

Head-to-Head Comparison of Methods		
Iteration	Total Accuracy	Area Under ROC Curve
1	NeuralNet	NeuralNet
2	NeuralNet	LogReg
3	NeuralNet	LogReg
4	Equal	NeuralNet
5	NeuralNet	NeuralNet
6	NeuralNet	LogReg
7	NeuralNet	LogReg
8	LogReg	NeuralNet
9	LogReg	LogReg
10	NeuralNet	NeuralNet
Totals	**NeuralNet: 7**	**NeuralNet: 5**
	LogReg: 2	LogReg: 5
	Equal: 1	Equal: 0

Table 5: Head-to-Head Comparison of Methods

As depicted in the table, through most iterations, the total accuracy of the neural network is greater than or equal to the accuracy of the logistic regression model. Furthermore, out of the 10 tested iterations, the neural network outperforms the logistic model with respect to area under ROC curve in 5 iterations, indicating that the neural network only marginally outperforms the logistic model (when looking at total average AUC). Fitting the theme of the study, the comparison can be summarized as a head-to-head "Win-Loss-Tie" Record:

Total Accuracy	**Area Under ROC Curve**
7-2-1	5-5-0

Table 6: Win-Loss-Tie Record

4.2 Example Iteration

For example, in the first iteration, the neural network had a total accuracy of 91.13% while the logistic regression model had a total accuracy of 89.11%. The area under the ROC curve for the neural network was 95.98%, while the area under the ROC curve for the logistic model was 95.19%. The following figure depicts the neural network for iteration 1:

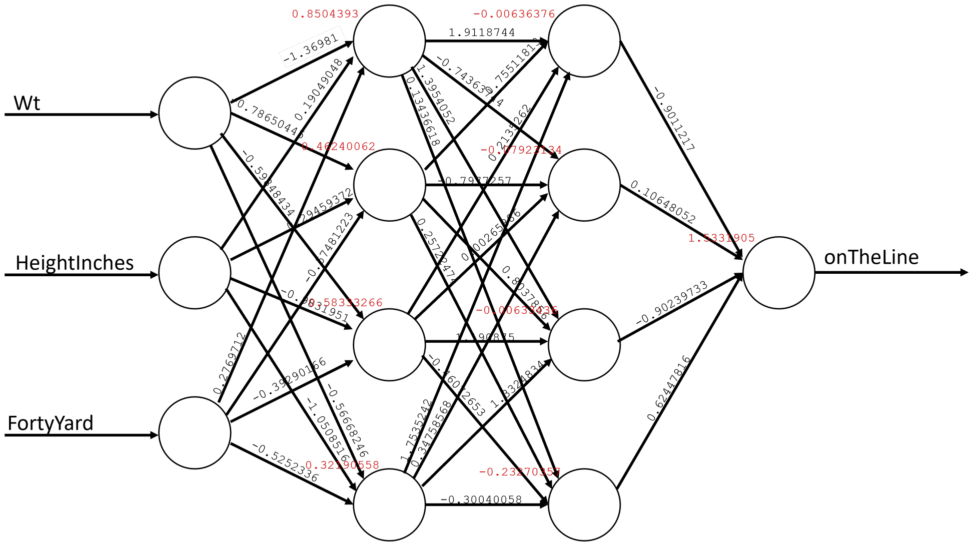


Figure 1: Neural Network Architecture - Iteration 1

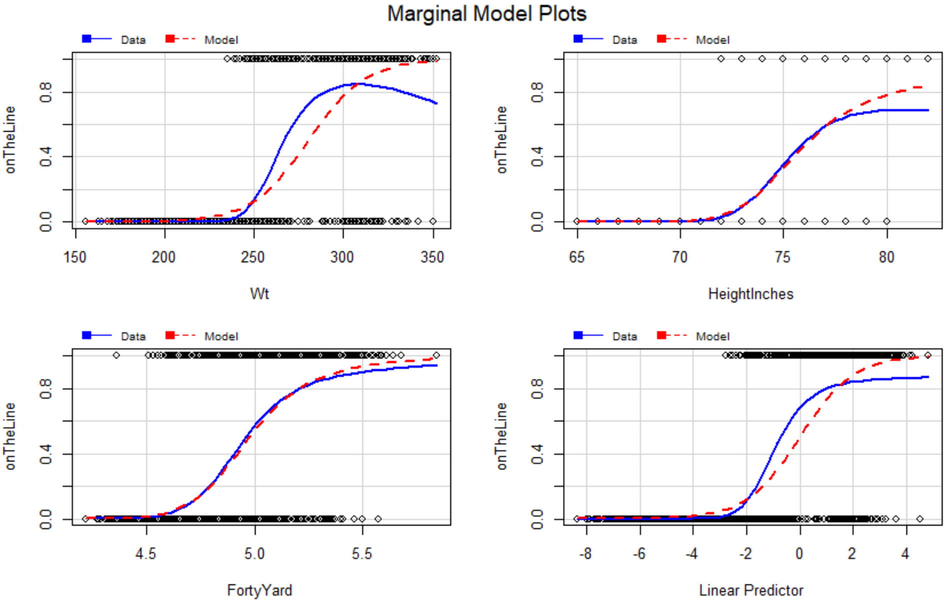


Figure 2: Marginal Model Plots - Iteration 1

The marginal model plots for the logistic model and ROC curves for the neural network and logistic model are depicted on the following page. With regard to *both* total accuracy and area under ROC curve, the neural network outperforms the logistic model.

The following are the confusion matrices for the Neural Network and Logistic model respectively:

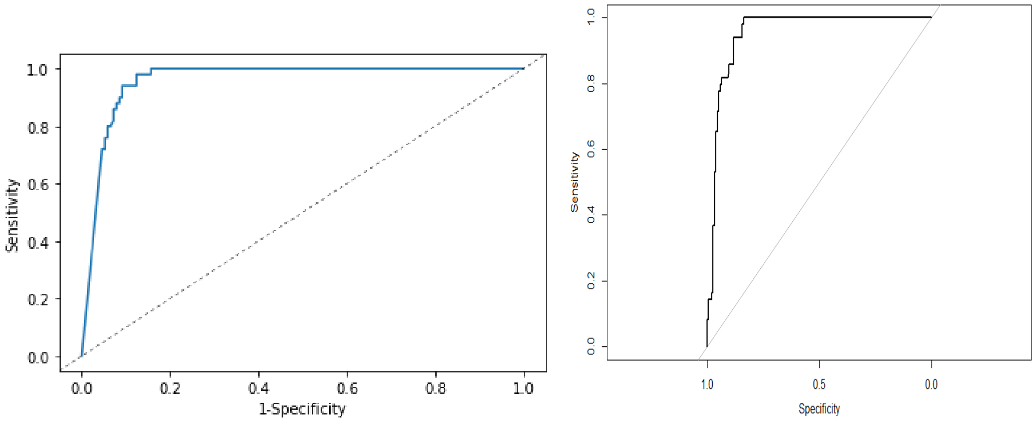


Figure 3: ROC Curves - Iteration 1

		Predicted onTheLine	
		0	1
True	0	**142**	**11**
onTheLine	1	**7**	**43**

Figure 4: Neural Network Confusion Matrix - Iteration 1

		Predicted onTheLine	
		0	1
True	0	**145**	**8**
onTheLine	1	**14**	**35**

Figure 5: Logistic Model Confusion Matrix - Iteration 1

The details of iterations 2-10 are similarly presented in Appendix B.

4.3 Interpretation of Logistic Coefficients

As evidenced by the calculated coefficients for the logistic models (given in Appendix A), player height and weight data are highly significant in determining probability of playing a position that is "on the line."

The player weight variable is deemed significant by every logistic regression model and has an average coefficient of 0.0564934. This indicates that an increase of 1 pound in weight increases the log-odds of the model predicting an "on the line" position by 0.05812. Thus, as expected, heavier players have a higher probability of playing a line position.

The player height variable is deemed significant in 7 of the 10 iterations, with an average coefficient of 0.175842. This indicates that an increase of 1 inch in height increases the log-odds of the model predicting an "on the line" position by 0.19225. Thus, as expected, taller players have a higher probability of playing a line position.

Although the coefficient for 40-Yard Dash Time is positive in each iteration, it is not deemed significant in a single one. This indicates that faster players are less likely to be predicted as "on the line," which is coherent with the hypothesis that shiftier, speedier players are more likely to play non-line positions. However, the measure does not have a statistically significant impact on the model.

5 Conclusion

In this work, a neural network approach was applied to precisely predict whether or not an NFL player is a line player given his height, weight, and 40-yard dash time. This application of machine learning allowed the researcher to more accurately (in terms of both total accuracy and area under ROC curve) predict this measure compared to logistic regression. With a head-to-head record of [7-2-1] and [5-5-0] in the categories respectively, and greater average accuracy/AUC, the neural network models, on average, outperformed the logistic models. Furthermore, the total accuracy average of 91.34% and area under ROC of 95.78% indicate that the created model is robust in such a volatile, real-data environment. Furthermore, the use of 10-fold cross validation further enforces that the method utilized in this research is rigorous and sound.

One reason the neural network was able to outperform the more traditional logistic regression model is perhaps that it is not constrained by the rigid structure of logistic regression. Although logistic regression coefficients are easily interpretable, this interpretability is due to the set structure of the model. This lack of flexibility perhaps allows a more flexible, adaptable algorithm (such as a neural network) to outperform logistic regression in certain classification tasks. Furthermore, the decision boundary in this research work may potentially have a complicated structure, favoring the neural network.

This research work was conducted to apply machine learning to an intuitive question involving a sports classification problem. The results of this work may potentially be used to better predict the optimal position for players to play as they transition from college to the NFL. However, the main outcome is that machine learning models can be applied to test intuitions of classification and can be applied to other sports in which physical metrics are strong indicators of player position.

Possible improvement and direction for future work includes using data measured on player's wingspan, hand size, jumping distance, and bench press to gain further accuracy. However, it is important to note that the data in such categories is rather sparse and adding complexity to the model may yield diminishing returns with higher computational costs.

References

Alvarez, R. (2011), 'MS Windows NT american football positions'.
URL: *https://americanfootballdatabase.fandom.com/wiki/Offensive$_{L}$inemanDe f ense*

Area Under an ROC Curve (n.d.), Available at `http://gim.unmc.edu/dxtests/roc3.htm` (2021/08/8).

Arpitha, T., Sanjay, H., Kumar, H. K., Aradhya, R. & Prithvi, B. (2021), 'Machine learning based prediction of the best suitable playing positions of the players in the game of basketball', *IEEE Mysore Sub Section International Conference (MysuruCon)* pp. 232–237.

Bunker, R. & Susnjak, T. (2019), 'The application of machine learning techniques for predicting results in team sport: A review'.

Fernandes, C. J., Yakuubov, R., Li, Y., Prasad, A. & Chan, T. (2020), 'Predicting plays in the national football league', *Journal of Sports Analytics* **6**(1), 35–43.

Garcia-Aliaga, A., Marquina, M., Coteron, J., Rodriguez-Gonzalez, A. & Luengo-Sanchez, S. (2020), 'In-game behaviour analysis of football players using machine learning techniques based on player statistics', *International Journal of Sports Science Coaching* **16**(1), 148–157.

Higham, C. & Higham, D. (2019), 'Deep learning: An introduction for applied mathematicians', *SIAM REVIEW* **61**(4), 860–891.

Hosmer, D., Lemeshow, S. & Sturdivant, R. (2013), *Applied Logistic Regression*, Wiley.

Hsu, Y.-C. (2020), 'Using machine learning and candlestick patterns to predict the outcomes of american football games', *Applied Sciences* **10**(13), 4484.

Kitani, N. L. . K. (2016), 'Predicting wide receiver trajectories in american football', *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)* pp. 1–9.

Kuzmits, F. & Adams, A. (2008), 'The nfl combine: Does it predict performance in the national football league?', *the Journal of Strength and Conditioning Research* **22**(6), 1721–1727.

Manzini, N. (n.d.), 'Single hidden layer neural network', Available at `http://www.nicolamanzini.com` (2021/08/8).

Neural Network (n.d.), Available at `https://www.google.com/search?q=definition+of+neural+network&rlz=1C1ONGR_enUS939US939&oq=definition+of+neural+network&aqs=chrome..69i57j0i512l5j0i457i512j0i512l3.3624j1j7&sourceid=chrome&ie=UTF-8` (2005/06/12).

NFL (2022), 'The digital athlete and how it's revolutionizing player health safety'.
URL: *https://www.nfl.com/playerhealthandsafety/equipment-and-innovation/aws-partnership/digital-athlete-spot*

Nieman, D. & Lee, J. (1990), 'Infectious episodes in runners before and after the los angeles marathon', *The Journal of Sports Medicine and Physical Fitness* pp. 315–328.

Norton, K. & Olds, T. (2001), 'Morphological evolution of athletes over the 20th century', *Sports Medicine* **31**, 763–783.

Ormerod, S. & Williams, H. (2002), 'Evaluating presence-absence models in ecology: The need to account for prevalence', *Journal of Applied Ecology* **38**(5), 921–931.

Pickering, C. (n.d.), 'Simplifaster how the rise of machine learning is impacting sport'.
URL: *https://simplifaster.com/articles/machine-learning-sports/*

Pion, J., Segers, V., Stautemas, J., Boone, J., Lenoir, M. & Bourgois, J. (2018), 'Position-specific performance profiles, using predictive classification models in senior basketball', *International Journal of Sports Science Coaching* **13**(6), 1072–1080.

R, D. (2018), 'Football and chess: How machine learning can improve playcalling in the nfl'.
URL: *https://digital.hbs.edu/platform-rctom/submission/football-and-chess-how-machine-learning-can-improve-playcalling-in-the-nfl/*

Sheather, S. (2009), *A Modern Approach to Regression with R*, Springer, College Station.

Zhong, D. & Chang, S.-F. (2001), 'Structure analysis of sports video using domain models', *IEEE International Conference on Multimedia and Expo, 2001. ICME 2001* p. 182.