Proof of Concept for a Receiver Separation Model using Long Short-Term Memory Recurrent Neural Networks on Next Gen Player Tracking Data

Jason Rock Phelps
Senior Fraud Waste and Abuse Data Scientist, Anthem, Inc.
Lead Data Scientist, Sports Injury Predictor
jason.rock.phelps@gmail.com
Cell: 614-260-7663

Introduction

The pass play is widely accepted as the most successful offensive play in football. On any pass play, a complexity of moving bodies ensues. Having predefined routes to create space between themselves and their would-be defenders in order to receive a pass from the quarterback, five eligible receivers are trying to make that design a reality and give the quarterback a viable target. All the while, the quarterback is processing the movement of the twenty-one other bodies on the field as some defenders gravitate towards his receivers and others collapse and break through the wall meant to give him time to find his targets. This time is fleeting and his opportunity to find a target before he is flat on his back is gone in a matter of seconds. A quarterback must make the most of the time he has to process the throng and find an open receiver. The receivers need to be given choreographed direction to help each other generate open targets with consistency and confidence to make the pass play successful.

A single receiver can only do so much to help himself create separation using his talent, technique and athletic ability. Working as a unit allows them to create additional space for each other by distracting and obstructing the defense. Coordinators can put their players in a situation to succeed if they can call route combinations that will generate space for their receivers. The chance of success will be elevated if the quarterback can call for the snap already knowing which of his receivers are most likely to be open and when.

Good quarterbacks are renowned for their ability to process a hoard of information quickly. What coverage personnel are on the field and what coverage scheme are they positioning themselves in? What leverage are cornerbacks playing with and with how much cushion? Where is the blitz coming from? But with all that is happening on the field, refining this thought process would be beneficial.

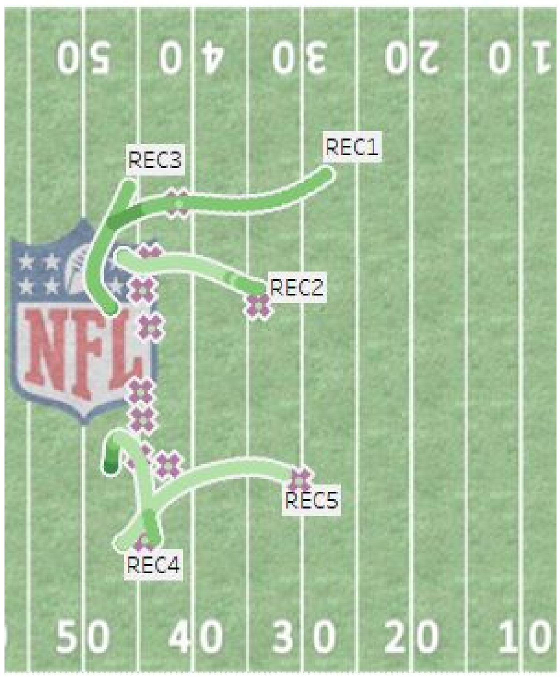
Optimizing the pass route combinations to generate the most separation from the defense is essential, as is the quarterback's ability to identify when the receiver is open. With five eligible receivers often spread across a field and depth of vision too large to process in full in the time allotted, it is imperative that they know where to look and when. Knowing this can optimize their progression reads through the route combination to increase the likelihood they find an open target on each passing play. The Next Gen player tracking data has the potential to deliver incredible insights on these fronts.

How would a separation prediction tool be utilized?

Let's look at how this works in practice for the quarterback. Say he calls a play with 3 receivers split out to the wide side of the field and 2 more to the short side. Receiver 1 is running a 15-yard corner route, receiver 2 a 7-yard slant, receiver 3 a flat route to the wide side, receiver 4 a chip block before releasing on a flat route to the short side, and receiver 5 running a seam route.

Predicted Separation (In Yards)

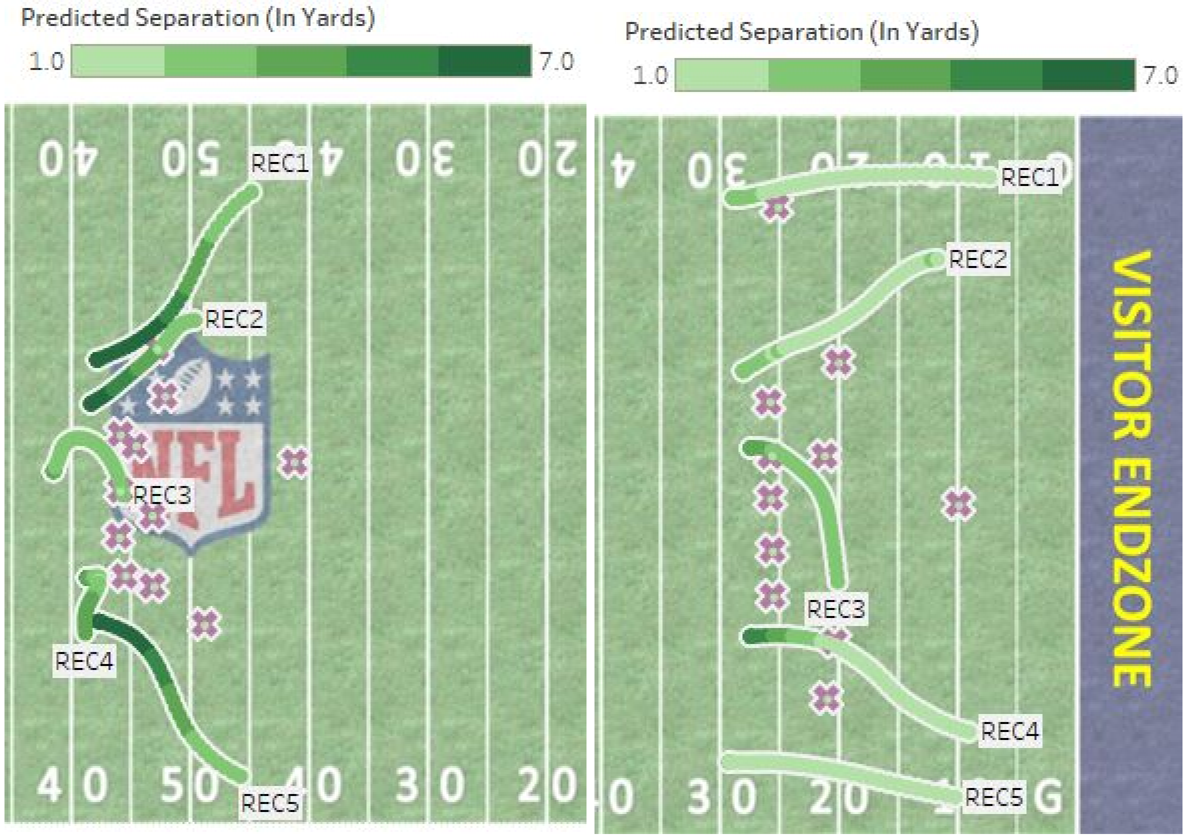
1.0 – 7.0



Seconds From Snap	Receiver On Play: REC1	REC2	REC3	REC4	REC5
0	4.2	2.8	4.3	5.7	2.5
0.2	4.2	2.7	4.2	5.7	2.5
0.4	4.1	2.6	3.9	5.4	2.5
0.6	3.9	2.3	3.6	4.9	2.4
0.8	3.6	2.0	3.2	4.2	1.9
1	3.5	1.8	3.0	3.5	1.8
1.2	3.2	1.9	2.8	2.7	1.7
1.4	3.2	1.7	2.6	2.2	1.7
1.6	3.0	1.7	2.4	1.7	1.5
1.8	2.8	1.7	2.3	1.5	1.4
2	2.9	1.8	2.3	1.4	1.3
2.2	2.8	1.9	2.4	1.4	1.2
2.4	2.6	1.8	2.4	1.6	1.3
2.6	2.7	2.1	2.6	1.7	1.3
2.8	2.5	2.2	2.7	1.9	1.2
3	2.6	2.4	2.7	2.1	1.3
3.2	2.7	2.3	2.9	2.3	1.3
3.4	2.7	2.3	3.0	2.5	1.3

The model predicts that receiver 4 will have the most separation immediately off of the snap, but we know he is preparing to engage in a block. The recommended route progression read optimized for separation of each read would be to receiver 1, then to receiver 3, then back to receiver 1 or hold on receiver 3. If the quarterback is flushed quickly to the short side then he should look to receiver 4 before receiver 2. No model is going to provide guarantees, but this does suggest where a quarterback should be looking for his most open target.

Now, prior to running this play, the offensive coordinator was in his office sketching out route combinations to add to the playbook. This model would provide feedback on whether the route combinations were likely to create space for his wide receivers. Let's say he had sketched up two route combinations with likely defensive pre-snap alignments and wanted to add them to the game plan for that week's game. And let's say he only had enough practice time during the week to introduce one of them. Which should he choose. Let's consider the two plays below.



For receiving options 1 second or more after the snap, play 1 on the left produces an average predicted receiver separation of 1.9 yards and a max predicted separation of 8.2 yards. Compare that to play 2 on the right where the average predicted receiver separation is 0.9 yards with a max predicted separation of 3.0 yards. Both measures on play 2 are less than half that of play 1. It should probably be a higher priority to install play 1 in the playbook.

Not only would the model allow for route combination comparisons, but also comparisons between the same route combination with different field positions or expected defensive configurations.

How can we get from the Next Gen tracking data to the insights above?

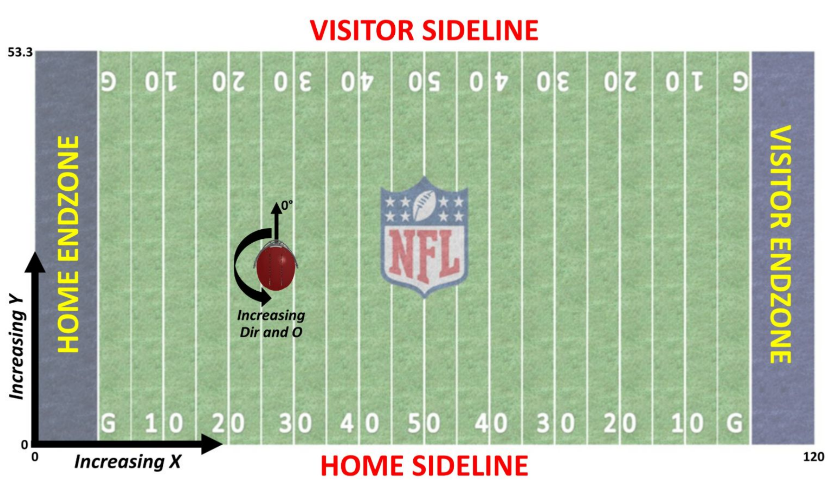
What is known by the quarterback before the ball is snapped? First, the quarterback knows the planned pass routes of each of his eligible receivers. So, he knows roughly where they will be on the field at any given moment of the designed play. He also knows the starting location of each defender. What is trickier to know is the defensive coverages that will be deployed during the play. He also has limited information as to how his receivers might deviate from their routes when the quarterback holds the ball longer than designed. Based on this information, we have a framework for what we can plan for and predict within reason. We can build a model to predict the separation of a receiver at any given moment in the designed play based on the factors a quarterback can assess pre-snap to inform our gameplan.

Preparing the tracking data

The NFL's Next Gen tracking data includes player location on the field at 10 frames per second for all 22 players on any given play, as well as for the football itself. Significant frames are tagged for major events within the play, such as the moment the ball was snapped or when a pass forward was attempted. As part of the NFL's Big Data Bowl, additional descriptive player and play data was included. The tracking data is very raw and requires significant preprocessing before tackling this problem. Some normalization is required to mitigate the relatively small sample provided and make more reliable comparisons between plays. Certain predictive and predicted fields need to be calculated. Finally, the data must be restructured into a format that will work in our model.

The basic steps required are as follows:

- Combine player tracking data from all available games; in this case 91 games from the 2017 season
- Determine which plays are pass plays
    - A pass forward, quarterback sack or quarterback strip sack event happened on the play
    - Exclude any Special Teams plays
- Normalize the plays so all offensive plays are moving in the same direction (home endzone to visitor endzone)
    - Determine the direction of the offense by comparing the average distance from the back of the home endzone value of the offensive and defensive players at the time the ball is snapped.
    - For any plays where average offensive distance from the back of the home endzone is greater than average for the defensive reflect over the 50 yard line



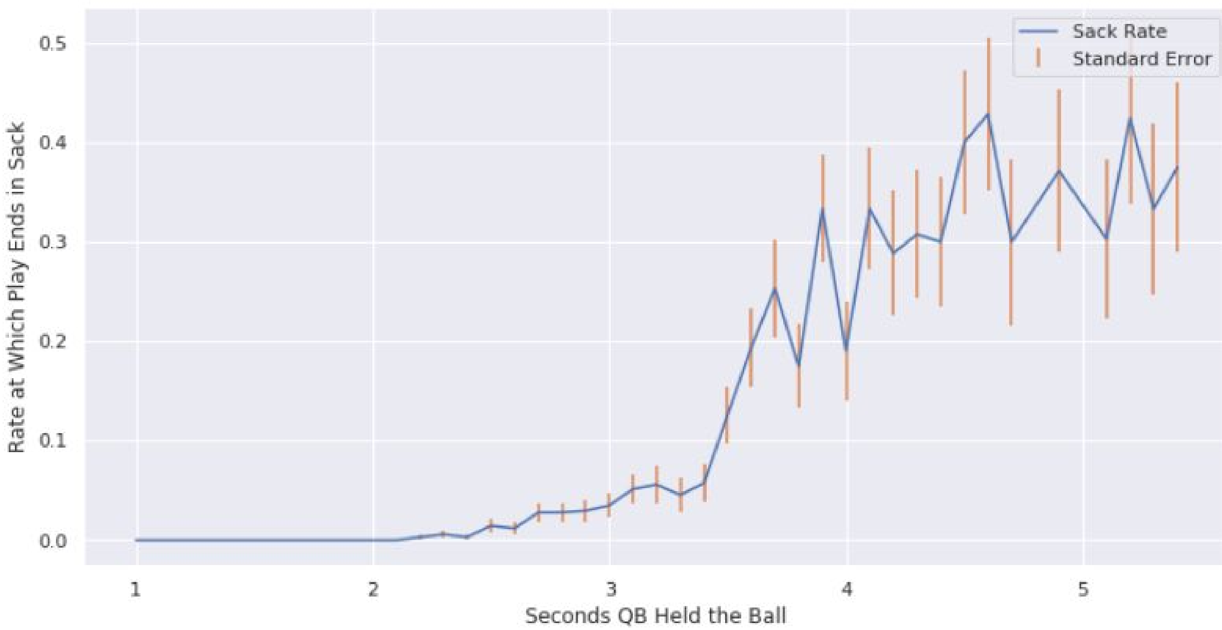
- Normalize each play so the short side of the field is always to the home sideline
    - For any plays where the football at the snap is above the midfield line, Y=53.3/2, reflect y values over this line
- Determine the seconds from snap for any given frame
    - The frame at which the ball is snapped is set as 0, with each subsequent frame being an additional tenth of a second post snap.
- Calculate the separation in yards between each eligible receiver and the nearest defender.
    - Use Euclidean distance between the receiver and each defender on each frame of a play.
    - Determine the distance of closest defender for each receiver and each frame and assign this as the separation.
- Give distinct receiver id (REC1,REC2,REC3,REC4,REC5) to each receiver for each play
    - Ranked by y value so REC1 is on the wide side of the field and REC5 is on the short side of the field when ball is snapped
    - This is used to pivot the data from records for each player-play-frame to records for each play frame to be used as input into the model.
    - This is used in lieu of numbering the wide receivers, running backs, and tight ends separately since these numbers could vary from pass play to pass play and would introduce null values into similarly pivoted data.
- Give distinct defender id (DEF1,...,DEF11) to each defender for each play
    - Ranked by y value so DEF1 is on the wide side of the field and DEF11 is on the short side of the field when ball is snapped
    - This is used to pivot the data from records for each player-play-frame to records for each play frame to be used as input into the model.
    - This is used in lieu of numbering the defensive backs, defensive lineman, and linebackers separately since these numbers could vary from pass play to pass play and would introduce null values into similarly pivoted data.
- Restructure player tracking data so each play frame has one record capturing location data for each receiver and each defender and the separation of each receiver from the closest defender on a given play frame.

Scope and considerations

This model is designed to predict separation of receivers on traditional passing plays. So, only plays where a total of five wide receivers, tight ends, running backs and fullbacks are kept. Although this criterion does not capture all possible passing play personnel scenarios defined by the NFL rulebook, it does capture 97.7% of the pass plays identified. These should represent traditional passing play offensive personnel and provide an adequate sample for building a model.

Further, not every frame of a play is relevant to our analysis. The tracking data often begins a second or more before the play and continues after the quarterback has passed the ball. For this model we only want to focus on the time frame from the snap and while the separation is unaffected by the release of the ball from the quarterback. We also only want to capture the designed patterns and not ad-libbed movement of the receivers when the quarterback can't find a target.

How much time does a quarterback have to release a pass before his chance of getting hit is unreasonable?



The quarterbacks are well aware of the limited time they have in the pocket. To reliably avoid a sack, they need to get rid of the ball within 3.4 seconds of the snap. After that point, the likelihood of a sack rises steeply from 10% to roughly 30% within half a second and to roughly 40% in another half second. At 3.4 seconds, the likelihood of the quarterback suffering a sack greatly increases, making it unwise to design routes and dropbacks that require more time than that to develop.

In defining our training data, only the path of the receiver prior to the quarterback releasing the pass will be considered. Also, only the receiver path prior to 3.4 seconds after the snap will be considered if the quarterback has held the ball beyond that time.

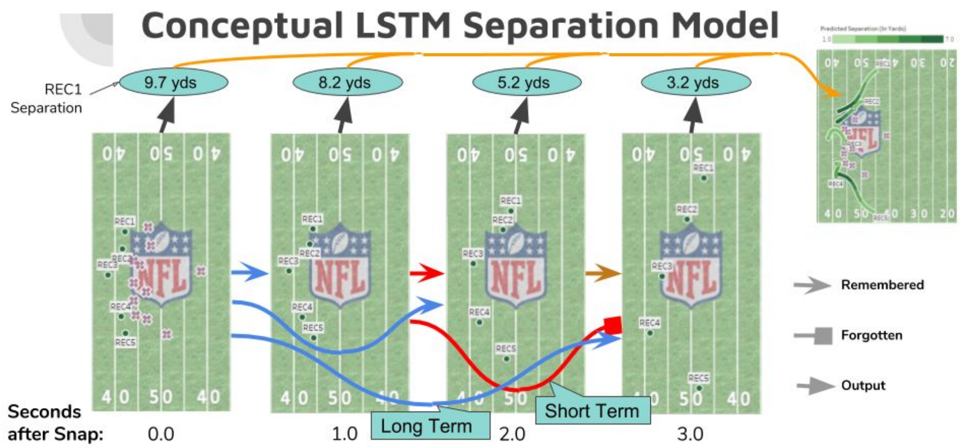
Modeling Approach

Traditional time series models like Auto Regressors (ARs), Moving Average (MAs) models or Autoregressive Integrated Moving Average Models (ARIMAs) are typically used in univariate time series modeling. Modeling receiver separation at a given point in a play based on the separation of the receiver during the play up to that point may be predictive, but the separation of the receiver throughout the play is not something that is readily known prior to the play and thus the insights from such a model would be inapplicable in a live setting.

Traditional machine learning techniques are powerful when attacking multivariate problems of independent instances. So, they may be applied to this problem to predict the separation of a receiver based on the player's location on the field at a given moment and the starting position of the defender (amongst other variables). But the separation of a receiver is theoretically largely based on how they arrived at that position. What makes a receiver effective at getting open is their ability to "set up" a defender with unexpected directional changes or acceleration. It would be possible to introduce additional variables to a regression model to capture the time taken to arrive at the current location or the type of pattern the receiver is executing. However, this does not capture the incremental accelerations during the route and predictions of separation mid-route may suffer from noise introduced from the part of the route not yet performed.

Deep learning and neural networks are the cutting edge of machine learning techniques. Deep neural networks have shown to be superior to traditional machine learning techniques when analyzing multidimensional data like images or video. Theoretically, they can learn to make standard calculations like distance between two players, perceived line of scrimmage or whether the receiver has qualities of a running back or tight end based on their location in the formation without being explicitly instructed to do so. This ability allows them to also learn latent or conceptually difficult patterns without direction. A particular variation suites this problem well: Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNN). LSTM RNNs have the ability retain information from previous related training samples to predict the current state of a system. In our case, the separation at a given frame can be predicted based on not only the current location of the receiver, but also the route he took to arrive there AND the routes fellow receivers have run on the play. And unlike time series models, these LSTM RNNs can process any additional predictive variables with similarity to traditional machine learning models. Another significant capability of LSTM RNNs is that they can accept time series of different lengths, which is the case with play tracking data since not every play has the same length.

For this concept development, a limited number of LSTM RNN designs were trained and validated using Python and keras with a tensorflow backend. The most successful of which had two LSTM hidden layers with 100 nodes each. These multilayer LSTM RNNs were built using an Adam optimizer with a learning rate of 0.0001, batch sizes of 64, and early stopping after 50 epochs without improvement using 90% training and 10% validation sets measuring error using root mean squared error as the loss function. Separate models were trained to predict the separation for each of the receivers 1 through 5. As a proof of concept, only X and Y location values for each of REC1-5 frame and the DEF1-11 locations at the snap of the ball were used as inputs and only the 1349 plays where the QB held the ball at least 3.4 seconds for each model.



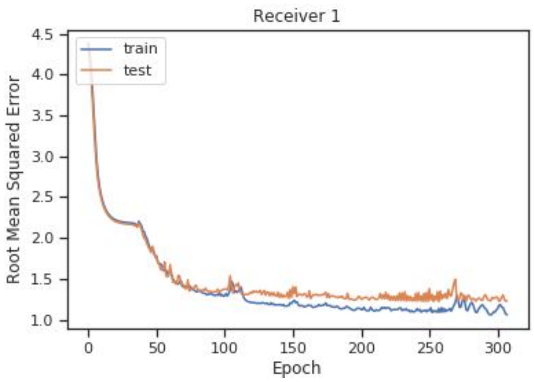
In this conceptual map of the LSTM separation model, we can see how location information from previous time points in the play can inform the predictions at future time points. We can also see how the model can learn to remember or forget a “memory” of a previous time point depending on whether it contributes to a better prediction.

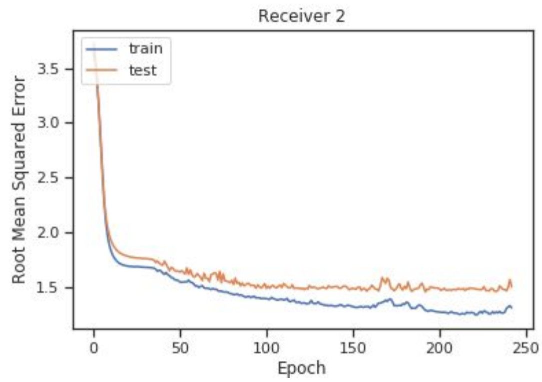
Model Results

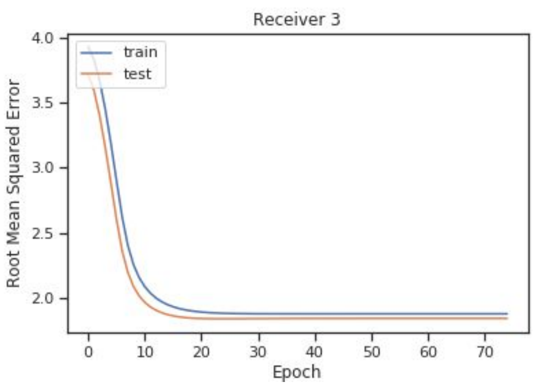
Receivers 1 and 5 being the outside most receivers on a play were the easiest to model having the lowest validation root mean squared error (RMSE). Receiver 3 has the most error. This can be explained by the fact receiver 3 captures the middle most reciever on the field, which might be a running back in the backfield, a slot receiver or tight end. In looking at the training vs test RMSE learning history plot for receiver 3 below, we can see that the learning (indicated by a drop in RMSE) plateaued similarly to that of the other receiver models for the other receivers. However, unlike for receivers 1, 2, 4 and 5, it never broke through the plateau. Because of the complexity of the 3rd designated receiver, it will likely take more time for the neural network to learn the underlying patterns explaining separation and would benefit from a more lenient early stopping threshold.

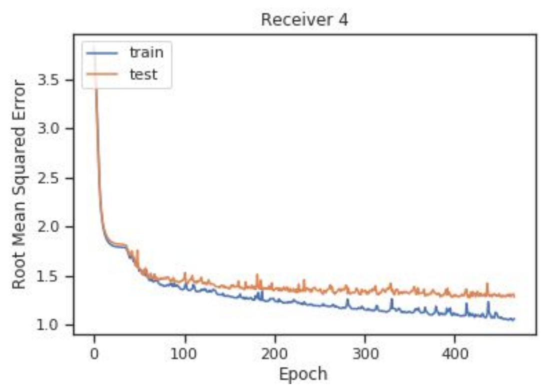
Receiver	Validation RMSE	Validation MAE
Receiver 1	1.2 yards	1.1 yards
Receiver 2	1.5 yards	1.3 yards
Receiver 3	1.8 yards*	1.6 yards*
Receiver 4	1.3 yards	1.0 yards
Receiver 5	1.1 yards	0.9 yards

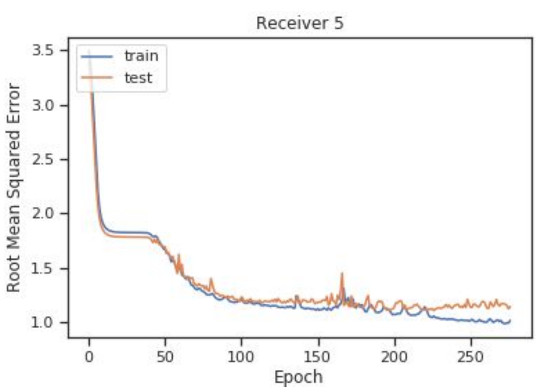
The errors from these LSTM RNN models are incredibly encouraging, especially considering the limited network designs tested. RMSEs and mean absolute errors (MAE) of ~1 to 1.5 are usable uncertainties. Let’s say we classify an open receiver as having at least 2 yards of separation from his defender. Then with these errors we can be pretty confident that when the model predicts a receiver will have 3.5 or more yards of separation he will be open in a game situation.











Overfitting can be an issue in building neural networks because of their power to learn nearly any pattern, even those that may seem to exist in the training data but don’t exist in out of sample cases. Further analysis of the training vs test RMSE learning history plots shows that overfitting is not a big issue in this application. As the training RMSE decreases so does the validation RMSE with limited separation between the two.

The relative success of a minimally trained Long Short Term Memory Recurrent Neural Network should prove the efficacy of using Neural Networks and Deep Learning to inform optimal route combination design and quarterback progression reads.

Additional Model Applications

The defensive coaching staff could also enter expected opponent route combinations and adjust starting defensive configurations to reduce the predicted separation. This could also be used as a player evaluation model by determining which receivers create the most separation above expectation.

Improvement and Enhancement Opportunities

There are a variety of ways we might improve on and/or enhance this model.

A big benefit could be had from more training data. We can capture more pass plays by determining which quarterback runs were scrambles rather than designed runs and include them in pass plays. Modifying the receiver identification to match the NFL rulebook definitions for a receiver would introduce a few more passing scenarios and formations. Simply having tracking data from more games would obviously be helpful, too.

As mentioned earlier, deep learning models are theoretically capable of learning latent information from the input variables despite the information not being explicitly introduced to the model. But, we can ease that pain by providing more explicit information to the model.

Any receiver separation model can likely be improved by introducing quarterback location and/or direction. A quarterback shifting left or right in the pocket will lure any defenders looking into the backfield in that direction. Looking off a safety is also effective at giving his receivers space down field.

Receiver qualities can be introduced, as well. Whether the receiver is a wideout, a tight end or running back could make a difference in their ability to separate from defenders. The historical separation of that player on similar routes, or a receiver's tendency to create more separation than expected or their maximum speed and acceleration are examples of viable qualities.

Even with offensive pre-snap motion techniques designed to identify defensive coverage schemes, this typically only provides partial coverage information and only if the defense isn't actively trying to disguise their coverage pre-snap. So, the full defensive coverage scheme might not be clearly known by the quarterback. That being said, the model may be improved by introducing pre-snap behavior of the defense.

We can flip most of the receiver enhancement possibilities and apply them to the defense as well. Tagging each defender a corner, safety, etc. might make it easier for the model to learn the characteristics and capabilities of each role. Expected separation for the defender based on their coverage history or number of men in the box could add value, too. Another source of information the quarterback is gathering and could be captured by the player tracking data is movement of the defense pre-snap. Is a defender following a receiver in motion? Is a linebacker showing an urge to blitz? Or how about the movement of the safeties and/or linebackers during the quarterback's drop back when he is processing the play?

Game situation could help explain some noise, as well. Time on the clock, the line of scrimmage, down and distance, current score, etc. all affect the defensive coverage schemes and thus the expected separation of receivers. Weather and field surfaces probably have an affect on separation, too.

We could further normalize the data. For example, it may be the case that it doesn't matter whether the line of scrimmage is your own twenty-yard line or your opponent's forty-yard line. All plays in that area of the field could start at the same reference line while another reference line is used for plays nearer the end zones.

To further improve the model performance, output from other machine learning models could be used as input. Or multiple models can be built and the results ensembled.

Conclusion

This model does not need to be an end point, but can be a complement or starting point to further analysis.

Recent research shows that Convolutional Neural Networks, which are adept at capturing spatial information, have the ability to outperform LSTM RNNs models on some time series problems. And the combination of the two could be especially effective here to model the spatio-temporal tracking data.

The model could be rebuilt to predict whether a player would be open or not based on some predetermined separation cutoff, perhaps based on likelihood of completion given that separation. Going beyond separation, a model to predict the likelihood of completing a pass at any given point in the route could be built. And incorporating expected value added or win probability added of a completion at that depth of target and field position.

Regardless, the use of Long Short-Term Memory Recurrent Neural Network models to predict receiver separation is viable and could be exceptionally effective with further refinement. A tool like this can provide valuable information to coaches designing route combinations and to quarterbacks planning the progression reads. It could also help improve defensive coverage schemes or enhance player evaluation. With further resources this concept can become a go-to tool for an NFL organization.