Quantifying NFL Special Teams Play

Team Camel

Introduction

Special teams play can significantly impact the outcome of a game in the National Football League (NFL). The rising use of advanced metrics and data analytics in American football can help NFL analysts and coaches better understand what features influence special teams play, which has been relatively limited to date. The 2022 NFL Big Data Bowl Challenge seeks to address this. Our team worked with a research scientist from MinedXAI, an Ohio based explainable artificial intelligence company, to quantify special teams play using topological approaches for this competition.

Approach

We focused on analyzing two out of four types of special teams plays: extra points and field goals. Prior to modeling, we engineered features with potential predictive power for successful kick attempts. A metric called 'core distance' allowed us to quantify defensive pressure by considering the distances between the kicker and the kth nearest defensive player. For kicker accuracy, we considered the curvature of a kick, measured by its deviation from a straight-line trajectory calculated from the moments immediately following the kick. This also showed how far wide the ball swings left or right.

Next, the topologically derived UMAP algorithm created a low-dimensional embedding of the data (see Fig. 1). This cleanly separated regions of local connectedness in the data while capturing as much global structure as possible. Finally, we used HDBSCAN, a clustering algorithm with foundations in computational topology.

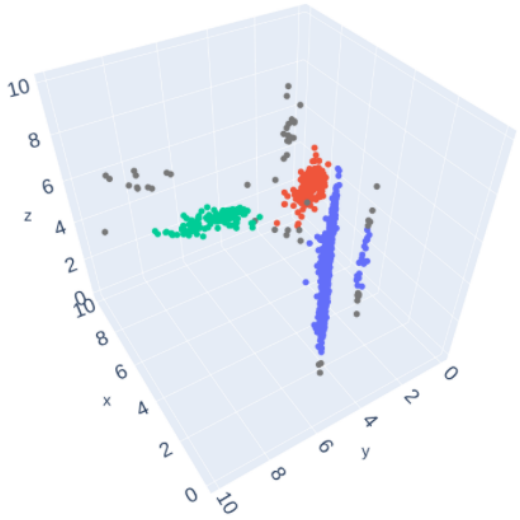

Fig. 1: Field Goal UMAP Embedding

Results

HDBSCAN separated our data into three main clusters, with a low number of outliers in the case of field goals. The ball's measured deviation from a straight-line trajectory appears to distribute equally between successful and unsuccessful kicks. From this we conclude failure to account for potential curvature in the ball's trajectory does not account for any systematic failure in kicks. Though the axes of the UMAP embedding need not maintain any meaningful relationship to the original features, in the case of field goals our axes retained correlation with two features: the location of the ball as it passes through the endzone and the result of the special teams play (e.g. "good" or "no good").

Using linear regression we show the y-axis of the UMAP embedding of field goals data and endzone location significantly correlate (R-squared = .5551). Logistic regression showed a very strong correlation between the UMAP x-axis and the play result (mean accuracy = .9987). This indicates these two features were most impactful in explaining the high-dimensional structure of the data.

Recommendations and Future Work

Our results indicate that the feature with the greatest impact on successful extra point and field goal attempts is kicker accuracy. Thus, to improve special teams play we recommend offensive coaches begin with a focus on kicker accuracy rather than strategic responses to defensive formations. This recommendation follows our results indicating that defensive pressure has a trivial effect on the success of extra point and field goal attempts. We speculate that the primary role of the defensive team is to force the kick and prevent trick plays.

Our next step is to focus on punt returns and kickoffs, which are significantly more complex. We will consider the motion of the receiving team as they attempt to progress across the field and avoid tackles. Additionally, there is more data on play type/kick type and strategy on which we may perform our analysis. We expect more differentiation between clusters along these features.