The Efficacy of NFL Defenses Against Passing Plays

Mohamad Quteifan, Data Scientist, M.S. in Data Science

Kayla Thompson, Data Scientist, M.S. in Data Science

Gloria Moore, Data Scientist, M.S. in Data Science

Bellevue University

Author

Mohamad Quteifan, Department of Science and Technology, Bellevue University

Kayla Thompson, Department of Science and Technology, Bellevue University

Gloria Moore, Department of Science and Technology, Bellevue University

Brett Werner, Department of Science and Technology, Bellevue University

Project Overview

In this project we plan to create an algorithm/model which will predict the efficacy of a defensive scheme against a passing play in the NFL. We were inspired by the NFL Big Data Bowl 2021 on Kaggle. We will measure the efficacy of the defensive scheme by looking at both the play result (complete pass, incomplete pass, quarterback sack or intercepted pass) as well as the net number of yards gained by the offense. The goal of our model will be to determine which defensive schemes work best for creating situations of zero or negative net yards.

We plan to perform Exploratory Data Analysis that will lead to model selection and we plan to evaluate multiple models. We are considering using Logistic Regression, Random Forest and K Nearest Neighbor to start, but this will evolve as our work begins. In order to create a reliable model we will also need to ensure our data is balanced and if it is not balanced we will use SMOTE to upsample the undersampled data. Our model will be successful if it effectively determines the outcomes of a play based on several specifics regarding the defensive scheme.

We plan to communicate using Slack and Zoom to work through this project.

References:

https://www.kaggle.com/c/nfl-big-data-bowl-2021/rules