

Caden Weigel  
CS 472  
5/8/25

## Project Proposal

For my project, I am working with the dataset from the Kaggle NFL Big Data Bowl 2025. The aim of this competition is to predict post-snap events based on pre-snap information. In other words, the goal is to predict what will happen during the play based on what happens before the play (but after the previous play). I want to specifically focus on offensive strategy based on defensive position/movement (best option for a play (and adjustments before the play) based on what the defense is doing, and how that option changes if the defensive formation changes).

<https://www.kaggle.com/competitions/nfl-big-data-bowl-2025/overview>

Goal: Determine optimal formations/motions for offense (specifically wide receivers and running backs) based on defensive formations/motions.

### Dataset:

The dataset consists of game data, play data, player data, player play data (what players do on given plays), and tracking data.

Player play data and tracking data provides the information needed for analysis, whereas everything else is more supplemental.

The tracking data is provided for 9 weeks of the season.

### Models:

- Clustering similar formations
  - Since there's a lot of similar offensive/defensive formations, clustering similar ones might make predictions easier
- Effectiveness of different formations
- Optimal motion predictor
- Counter formation predictor