
# FIELD GOAL PROBABILITY MODEL

Predicting field goal success probabilities to support game decision-making and kicker evaluation.

Lauren Manis

# MODEL OVERVIEW & **PURPOSE**

**The goal of this model is to estimate the probability of field goal success in the NFL**, leveraging key contextual and environmental factors such as distance, weather, game situation, and kicker performance. The predictions are designed to inform critical coaching and front office decisions, including whether to attempt a field goal or go for it on fourth down, evaluating kicker reliability, or making roster and contract calls.

The model is a **Logistic Regression built on football-intuitive features**—selected based on domain knowledge and practical relevance. This model was chosen for its strong balance of accuracy, interpretability, and real-world reliability, making it especially well-suited for high-leverage football decision-making.

# MODEL DETAILS

The model was trained on 12,829 field goal attempts from all NFL regular and postseason games between 2010 and 2023, and tested on all field goal attempts from the 2024 season.

- Data was sourced from the nflfastR package, which provides detailed play-by-play tracking for every NFL game.

An initial baseline model - a simple logistic regression using only kick distance - achieved an AUC of 0.774. The final model builds on this by adding a small, targeted set of football-intuitive contextual features, recognizing that distance remains the primary driver of field goal success.

- Only features that are consistently available, interpretable, and impactful were included to ensure simplicity and generalizability.



## Features

| Game Context | Conditions | Kicker |
|---|---|---|
| **Season** | **Kick Distance** (yards) | **Adjusted FG%** |
| **Last 2 Minutes** (Binary) | **Wind** (mph) | **Rookie** (Binary) |
| | **Precipitation** (Binary) | |

*Kicker FG%: cumulative within-season overall FG%, regressed to the league average.*
*Last 2 Minutes: 1 if the kick took place within the final 2 minutes of the $1^{st}$ or $2^{nd}$ half of the game.*

# ACCURACY METRICS

**AUC, Precision, and Log Loss were prioritized as the key evaluation metrics**

- **AUC**: Measures overall classification performance.
- **Precision**: Emphasized because false positives (predicting a make when the kick is missed) are riskier in coaching decisions—better to err on the side of caution.
- **Log Loss**: Penalizes overconfident incorrect predictions, helps assess calibration.

## AUC

**0.77**

Indicates strong ability to distinguish between made and missed field goals.

## ACCURACY

**82%**

82% of all field goal attempts were correctly classified by the model.

## PRECISION

**89%**

89% of predicted made field goals were actually made.

## LOG LOSS

**0.35**

Indicates well-calibrated predictions, predicted probabilities are close to actual frequencies.

Overall predicted FG probability of **86.3%**, closely aligned with the actual make rate of **85.6%** across the test set.

*Metrics shown are based on 2024 test set performance*

*Note: Accuracy is reported for completeness, though less informative in imbalanced datasets*

# MODEL SELECTION

A total of 8 models were built and evaluated, including logistic regression, lasso, random forest, and weighted variants.

Ultimately, the **Logistic Regression with Football-Intuitive Features** ("Football Model") model provided the best balance of performance, calibration, and interpretability.

- Logistic regression is well-suited for binary outcomes, produces easily interpretable coefficients, ideal for communicating insights to technical and non-technical audiences.
- This specific "Football" model matched or outperformed more complex alternatives while remaining simple, transparent, and coach-friendly.
  - While Lasso performed slightly better on some metrics, it introduced unnecessary complexity and offered no practical gain over the simpler Football Model.


| Model | AUC | Accuracy | Precision | Log Loss | Features |
|---|---|---|---|---|---|
| Lasso Regression | 0.783 | 0.862 | 0.867 | 0.339 | 17 |
| **Football Model** | **0.782** | **0.862** | **0.867** | **0.340** | **7** |
| Weighted Model | 0.782 | 0.860 | 0.870 | 0.342 | 7 |
| Significant Model | 0.781 | 0.861 | 0.866 | 0.340 | 9 |
| Interaction Model | 0.780 | 0.862 | 0.866 | 0.341 | 8 |
| Baseline Model | 0.774 | 0.862 | 0.864 | 0.344 | 1 |
| Full Model | 0.774 | 0.862 | 0.864 | 0.344 | 30 |
| Random Forest | 0.757 | 0.859 | 0.865 | 0.369 | 15 |

*Results are from the training set, used for comparing models and selecting the final one*
*Baseline = distance only. Full = all available features including*
*Significant Model: features selected based on p-values*
*Interaction Model: includes 2-way interactions between core features*

# MODEL ASSUMPTIONS

- Blocked field goals were excluded from the dataset
  - They account for <2% of observations and tend to behave like noisy misses, so were excluded for clarity.
- Snaps and holds were not accounted for, as there is no reliable data source to identify botched or mishandled snaps/holds.
- Kick direction (left/right hash) is not included due to lack of standardized data.
- False positives are more costly than false negatives, given coaching preference for conservative decisions in high-leverage spots.
- Kicker FG% inputs are cumulative within-season stats, including only attempts prior to each kick (no data leakage).
  - Adjusted FG% is regressed to the league average, especially early in the season, to stabilize estimates and account for sample size.
- Model trained on attempts from 2010–2023, excluding earlier years due to significant changes in league-wide field goal success rates.