Predicting NFL Play Calling

Using Historic Play-By-Play Data

Hans Missenheim

Executive Summary

- Top NFL teams look to improve strategic decision making
- Detailed NFL Play-by-Play data is available
- Machine learning models trained on past season
- A neural network model predicted ***72.2%*** of the 2022 season's plays

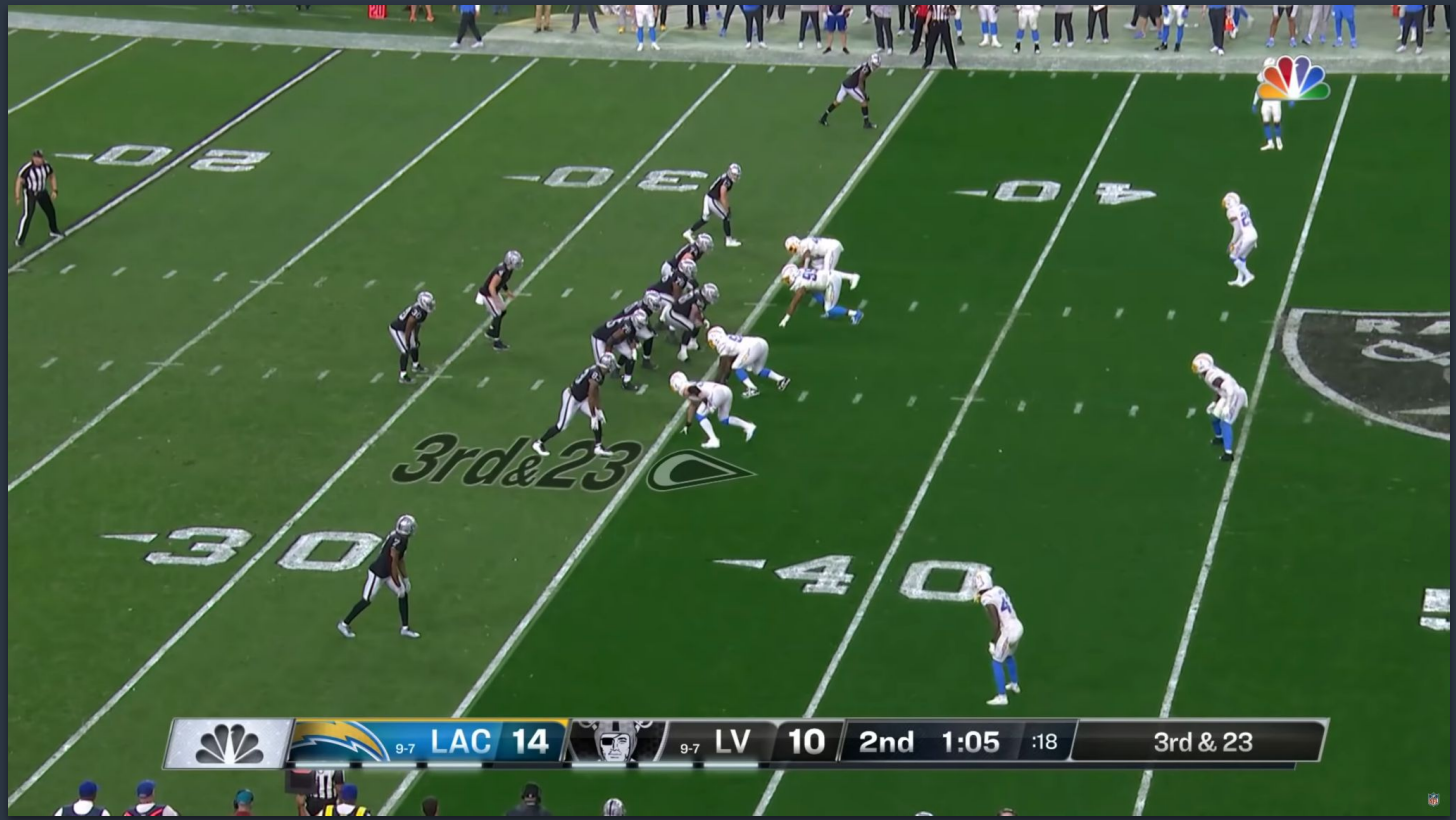
Research Question:

What is the next type of play (run or pass) given the current game state and previous plays?

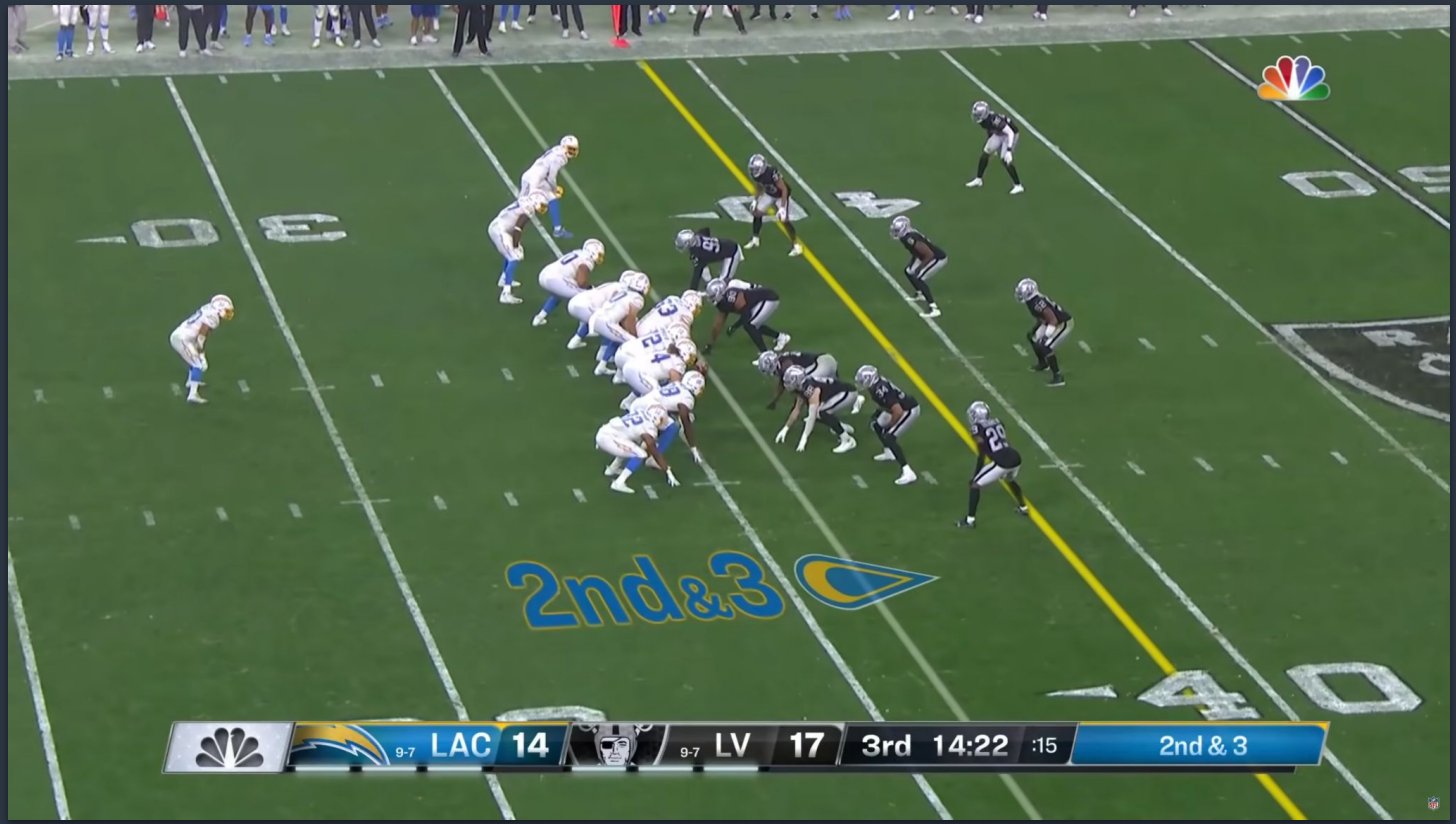
Dataset

- Made available through the ***nflfastR*** package
- Contains:
  - **24** seasons
  - **6,418** games
  - **1,148,717** plays
- Each play has **300+** variables


NFL
WEEK 18
GAME HIGHLIGHTS


3rd&23
9-7 LAC 14
9-7 LV 10
2nd 1:05
:18
3rd & 23


NFL
WEEK 18
GAME HIGHLIGHTS


2nd&3
9-7 LAC 14
9-7 LV 17
3rd 14:22 :15
2nd & 3

Utility

- **Defense:** Predict offensive play to limit success
- **Offense:** Understand current play call predictability and evaluate
- **NFL:** Increased strategic decision making
  = Increased entertainment value
  = Increased viewership

Methodology

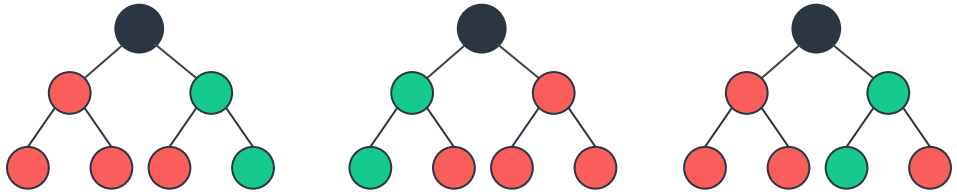
Naive Bayes

- Classification algorithm based on *Bayes' Theorem*
- Assumes independent variables
- Calculates probability per variable

$$P(A|B) = \frac{P(A|B)\,P(A)}{P(B)}$$

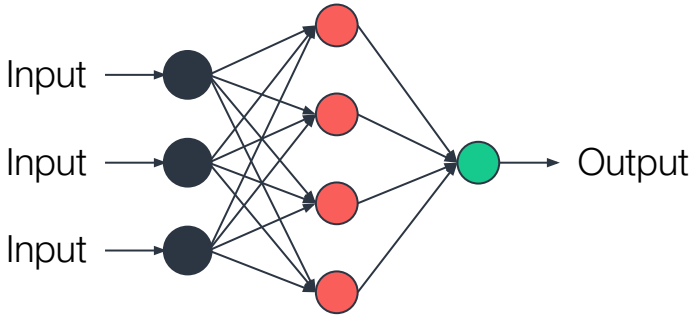
Random Forest

- Ensemble of decision trees
- Subset of samples and features per tree
- Aggregates predictions of every tree

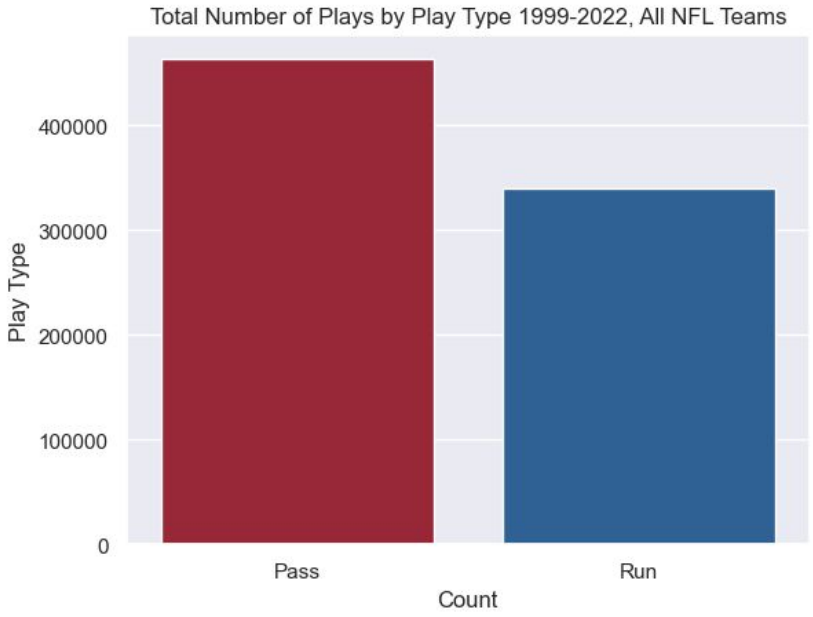


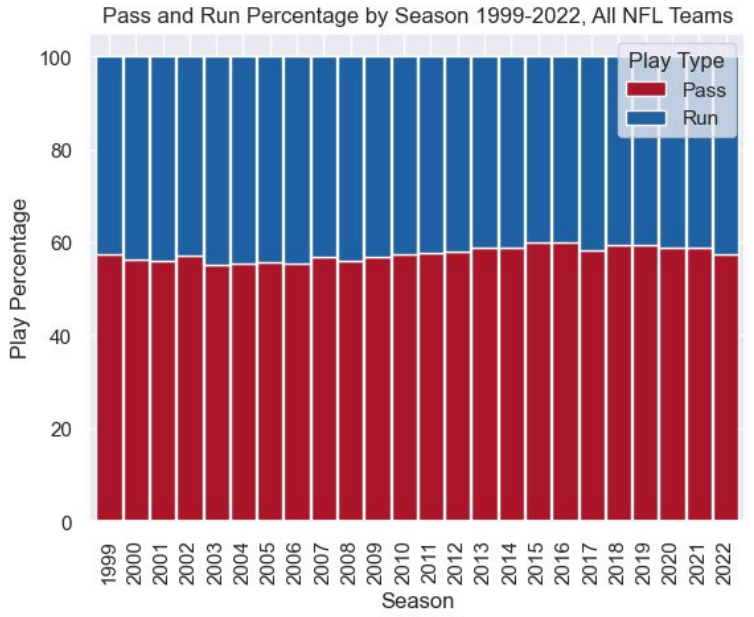
Feed Forward Neural Network

- Basic artificial neural network
- Composed of input, hidden, and output layers of neurons
- Weights given to inputs and activation function applied with neuron

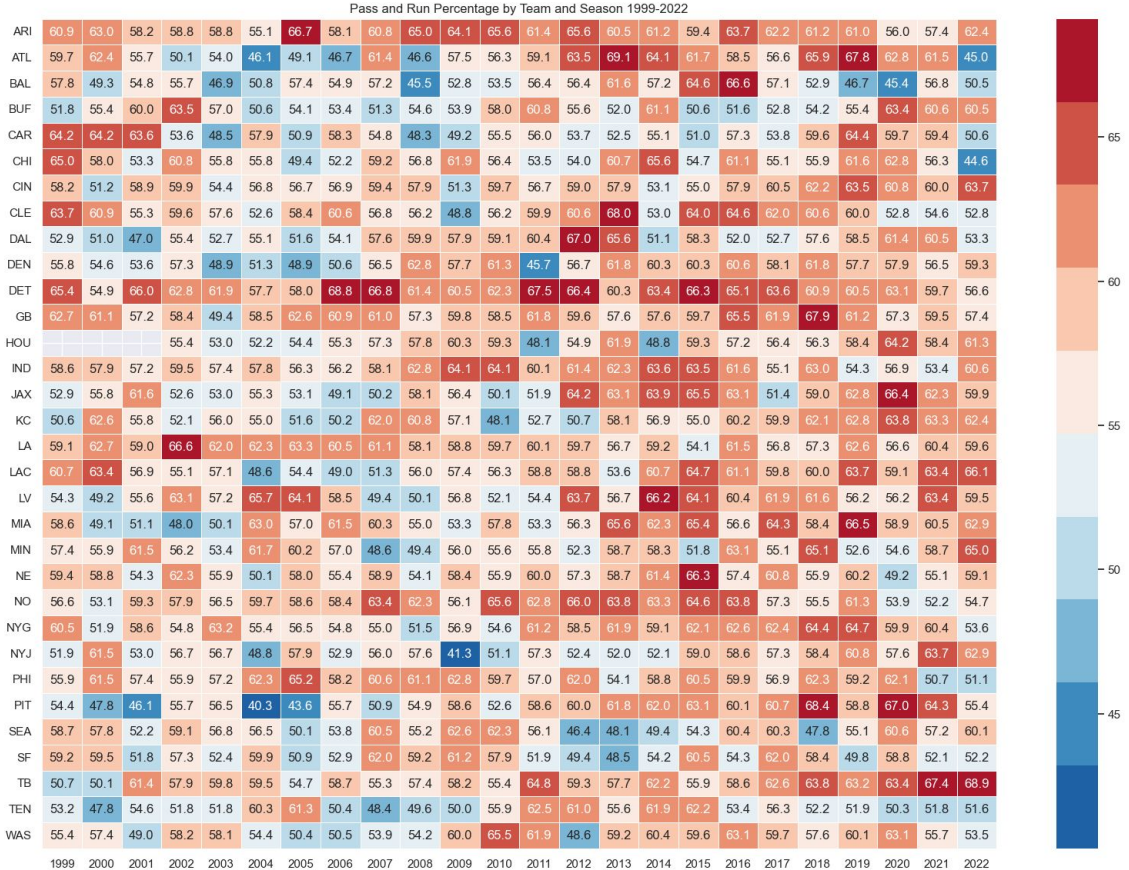


Analysis


Total Number of Plays by Play Type 1999-2022, All NFL Teams
Play Type
400000
300000
200000
100000
0
Pass
Run
Count


Pass and Run Percentage by Season 1999-2022, All NFL Teams
Play Type
Pass
Run
Play Percentage
100
80
60
40
20
0
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
Season

Analysis


Pass and Run Percentage by Team and Season 1999-2022
ARI 60.9 63.0 58.2 58.8 58.8 55.1 66.7 58.1 60.8 65.0 64.1 65.6 61.4 65.6 60.5 61.2 59.4 63.7 62.2 61.2 61.0 56.0 57.4 62.4
ATL 59.7 62.4 55.7 50.1 54.0 46.1 49.1 46.7 61.4 46.6 57.5 56.3 59.1 63.5 69.1 64.1 61.7 58.5 56.6 65.9 67.8 62.8 61.5 45.0
BAL 57.8 49.3 54.8 55.7 46.9 50.8 57.4 54.9 57.2 45.5 52.8 53.5 56.4 56.4 61.6 57.2 64.6 66.6 57.1 52.9 46.7 45.4 56.8 50.5
BUF 51.8 55.4 60.0 63.5 57.0 50.6 54.1 53.4 51.3 54.6 53.9 58.0 60.8 55.6 52.0 61.1 50.6 51.6 52.8 54.2 55.4 63.4 60.6 60.5
CAR 64.2 64.2 63.6 53.6 48.5 57.9 50.9 58.3 54.8 48.3 49.2 55.5 56.0 53.7 52.5 55.1 51.0 57.3 53.8 59.6 64.4 59.7 59.4 50.6
CHI 65.0 58.0 53.3 60.8 55.8 55.8 49.4 52.2 59.2 56.8 61.9 56.4 53.5 54.0 60.7 65.6 54.7 61.1 55.1 55.9 61.6 62.8 56.3 44.6
CIN 58.2 51.2 58.9 59.9 54.4 56.8 56.7 56.9 59.4 57.9 51.3 59.7 56.7 59.0 57.9 53.1 55.0 57.9 60.5 62.2 63.5 60.8 60.0 63.7
CLE 63.7 60.9 55.3 59.6 57.6 52.6 58.4 60.6 56.8 56.2 48.8 56.2 59.9 60.6 68.0 53.0 64.0 64.6 62.0 60.6 60.0 52.8 54.6 52.8
DAL 52.9 51.0 47.0 55.4 52.7 55.1 51.6 54.1 57.6 59.9 57.9 59.1 60.4 67.0 65.6 51.1 58.3 52.0 52.7 57.6 58.5 61.4 60.5 53.3
DEN 55.8 54.6 53.6 57.3 48.9 51.3 48.9 50.6 56.5 62.8 57.7 61.3 45.7 56.7 61.8 60.3 60.3 60.6 58.1 61.8 57.7 57.9 56.5 59.3
DET 65.4 54.9 66.0 62.8 61.9 57.7 58.0 68.8 66.8 61.4 60.5 62.3 67.5 66.4 60.3 63.4 66.3 65.1 63.6 60.9 60.5 63.1 59.7 56.6
GB 62.7 61.1 57.2 58.4 49.4 58.5 62.6 60.9 61.0 57.3 59.8 58.5 61.8 59.6 57.6 57.6 59.7 65.5 61.9 67.9 61.2 57.3 59.5 57.4
HOU 55.4 53.0 52.2 54.4 55.3 57.3 57.8 60.3 59.3 48.1 54.9 61.9 48.8 59.3 57.2 56.4 56.3 58.4 64.2 58.4 61.3
IND 58.6 57.9 57.2 59.5 57.4 57.8 56.3 56.2 58.1 62.8 64.1 64.1 60.1 61.4 62.3 63.6 63.5 61.6 55.1 63.0 54.3 56.9 53.4 60.6
JAX 52.9 55.8 61.6 52.6 53.0 55.3 53.1 49.1 50.2 58.1 56.4 50.1 51.9 64.2 63.1 63.9 65.5 63.1 51.4 59.0 62.8 66.4 62.3 59.9
KC 50.6 62.6 55.8 52.1 56.0 55.0 51.6 50.2 62.0 60.8 57.1 48.1 52.7 50.7 58.1 56.9 55.0 60.2 59.9 62.1 62.8 63.8 63.3 62.4
LA 59.1 62.7 59.0 66.6 62.0 62.3 63.3 60.5 61.1 58.1 58.8 59.7 60.1 59.7 56.7 59.2 54.1 61.5 56.8 57.3 62.6 56.6 60.4 59.6
LAC 60.7 63.4 56.9 55.1 57.1 48.6 54.4 49.0 51.3 56.0 57.4 56.3 58.8 58.8 53.6 60.7 64.7 61.1 59.8 60.0 63.7 59.1 63.4 66.1
LV 54.3 49.2 55.6 63.1 57.2 65.7 64.1 58.5 49.4 50.1 56.8 52.1 54.4 63.7 56.7 66.2 64.1 60.4 61.9 61.6 56.2 56.2 63.4 59.5
MIA 58.6 49.1 51.1 48.0 50.1 63.0 57.0 61.5 60.3 55.0 53.3 57.8 53.3 56.3 65.6 62.3 65.4 56.6 64.3 58.4 66.5 58.9 60.5 62.9
MIN 57.4 55.9 61.5 56.2 53.4 61.7 60.2 57.0 48.6 49.4 56.0 55.6 55.8 52.3 58.7 58.3 51.8 63.1 55.1 65.1 52.6 54.6 58.7 65.0
NE 59.4 58.8 54.3 62.3 55.9 50.1 58.0 55.4 58.9 54.1 58.4 55.9 60.0 57.3 58.7 61.4 66.3 57.4 60.8 55.9 60.2 49.2 55.1 59.1
NO 56.6 53.1 59.3 57.9 56.5 59.7 58.6 58.4 63.4 62.3 56.1 65.6 62.8 66.0 63.8 63.3 64.6 63.8 57.3 55.5 61.3 53.9 52.2 54.7
NYG 60.5 51.9 58.6 54.8 63.2 55.4 56.5 54.8 55.0 51.5 56.9 54.6 61.2 58.5 61.9 59.1 62.1 62.6 62.4 64.4 64.7 59.9 60.4 53.6
NYJ 51.9 61.5 53.0 56.7 56.7 48.8 57.9 52.9 56.0 57.6 41.3 51.1 57.3 52.4 52.0 52.1 59.0 58.6 57.3 58.4 60.8 57.6 63.7 62.9
PHI 55.9 61.5 57.4 55.9 57.2 62.3 65.2 58.2 60.6 61.1 62.8 59.7 57.0 62.0 54.1 58.8 60.5 59.9 56.9 62.3 59.2 62.1 50.7 51.1
PIT 54.4 47.8 46.1 55.7 56.5 40.3 43.6 55.7 50.9 54.9 58.6 52.6 58.6 60.0 61.8 62.0 63.1 60.1 60.7 68.4 58.8 67.0 64.3 55.4
SEA 58.7 57.8 52.2 59.1 56.8 56.5 50.1 53.8 60.5 55.2 62.6 62.3 56.1 46.4 48.1 49.4 54.3 60.4 60.3 47.8 55.1 60.6 57.2 60.1
SF 59.2 59.5 51.8 57.3 52.4 59.9 50.9 52.9 62.0 59.2 61.2 57.9 51.9 49.4 48.5 54.2 60.5 54.3 62.0 58.4 49.8 58.8 52.1 52.2
TB 50.7 50.1 61.4 57.9 59.8 59.5 54.7 58.7 55.3 57.4 58.2 55.4 64.8 59.3 57.7 62.2 55.9 58.6 62.6 63.8 63.2 63.4 67.4 68.9
TEN 53.2 47.8 54.6 51.8 51.8 60.3 61.3 50.4 48.4 49.6 50.0 55.9 62.5 61.0 55.6 61.9 62.2 53.4 56.3 52.2 51.9 50.3 51.8 51.6
WAS 55.4 57.4 49.0 58.2 58.1 54.4 50.4 50.5 53.9 54.2 60.0 65.5 61.9 48.6 59.2 60.4 59.6 63.1 59.7 57.6 60.1 63.1 55.7 53.5
1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022
65
60
55
50
45

Analysis

Naive Bayes Results

Encoded	Coach/Team	Balanced	Accuracy	AUC
No	No	No	64.90%	0.694
Yes	No	No	66.77%	0.722
Yes	Yes	No	66.71%	0.724
Yes	Yes	Yes	*66.71%*	*0.725*

Random Forest Results

Encoded	Coach/Team	Balanced	Accuracy	AUC
No	No	No	70.88%	0.781
Yes	No	No	70.70%	0.780
Yes	Yes	No	***71.22%***	***0.786***
Yes	Yes	Yes	70.98%	0.785

Neural Network Results

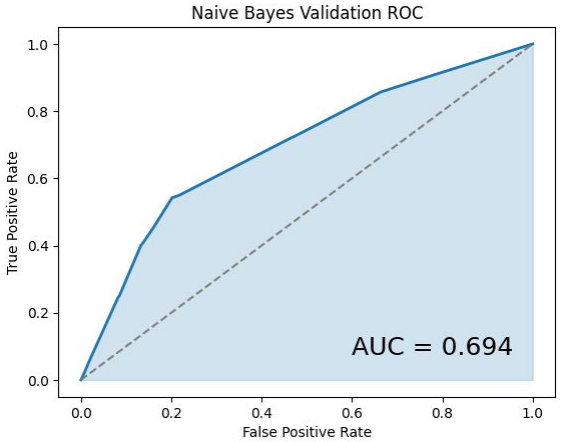
Hidden Layers	Layer Nodes	Activation	Accuracy	Loss
3	256	relu	***71.90%***	***0.5471***
2	256	relu	71.73%	0.5498
1	32	relu	71.39%	0.5593
1	128	tanh	70.70%	0.5692
3	32	relu	71.57%	0.5594

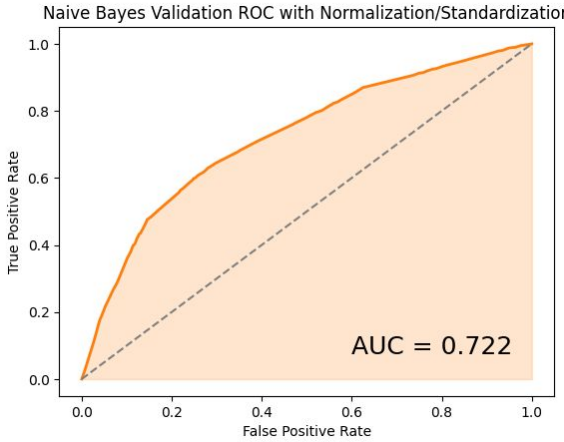
Analysis

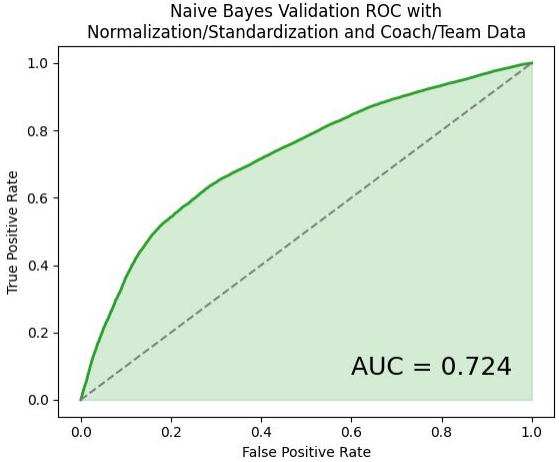
2022 Season Test Results

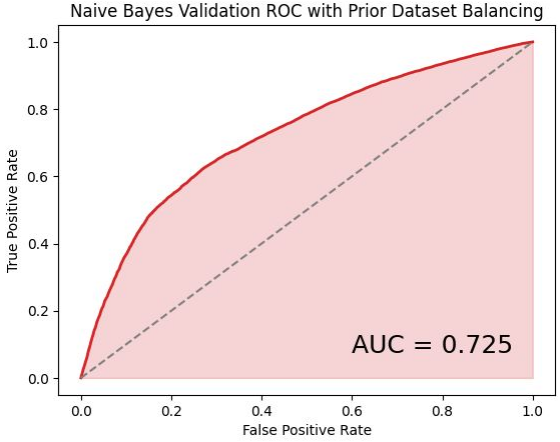
Algorithm	Accuracy
Naive Bayes	69.24%
Random Forest	71.60%
Neural Network	***72.20%***

Data Visualizations

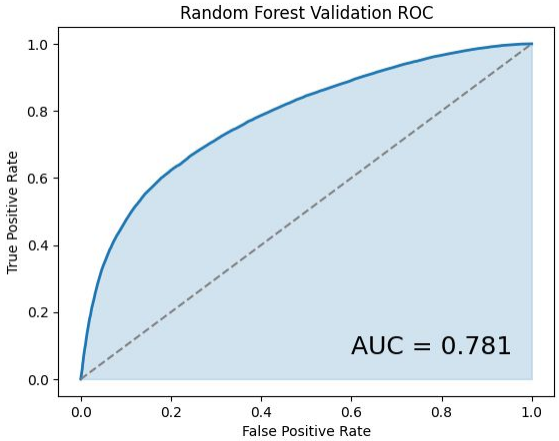


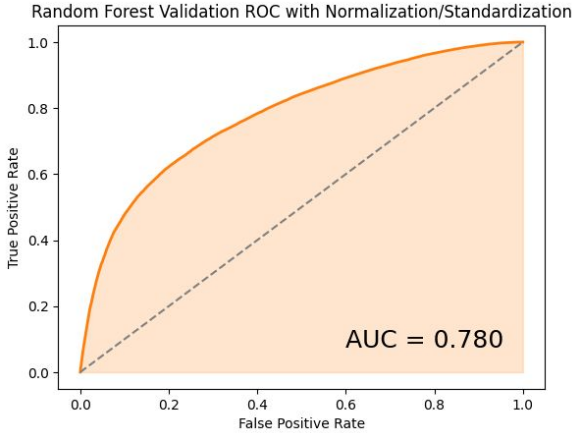


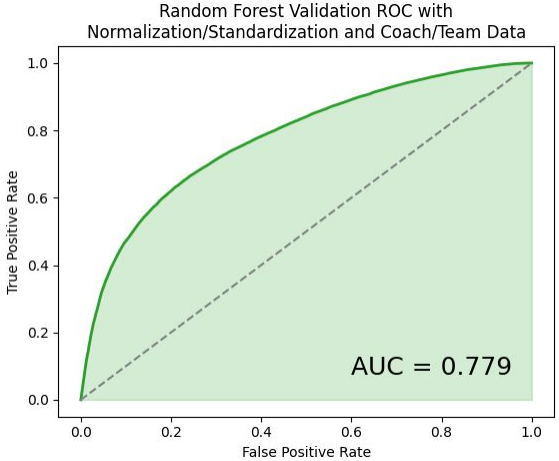


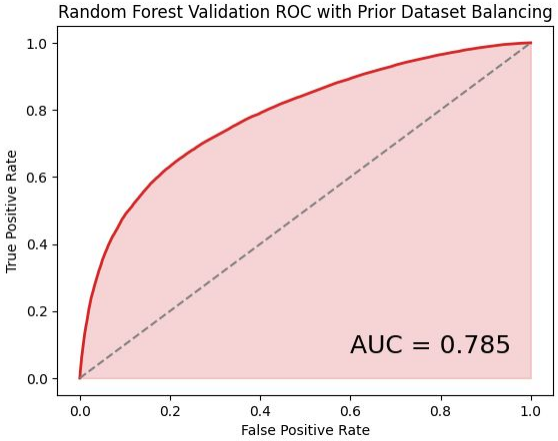


Data Visualizations

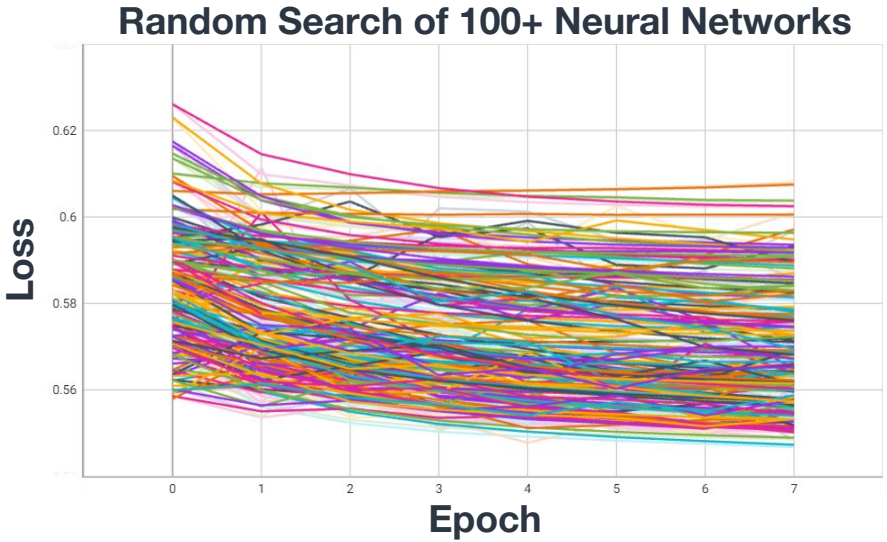


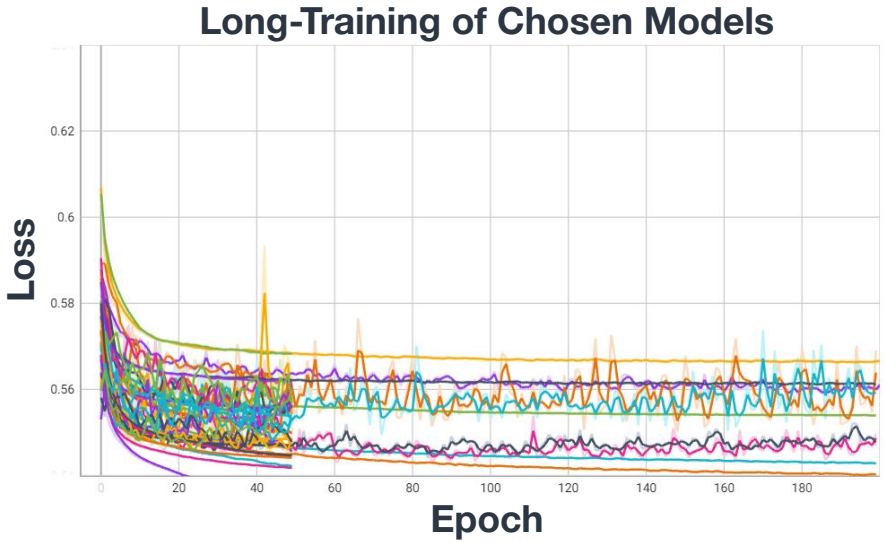




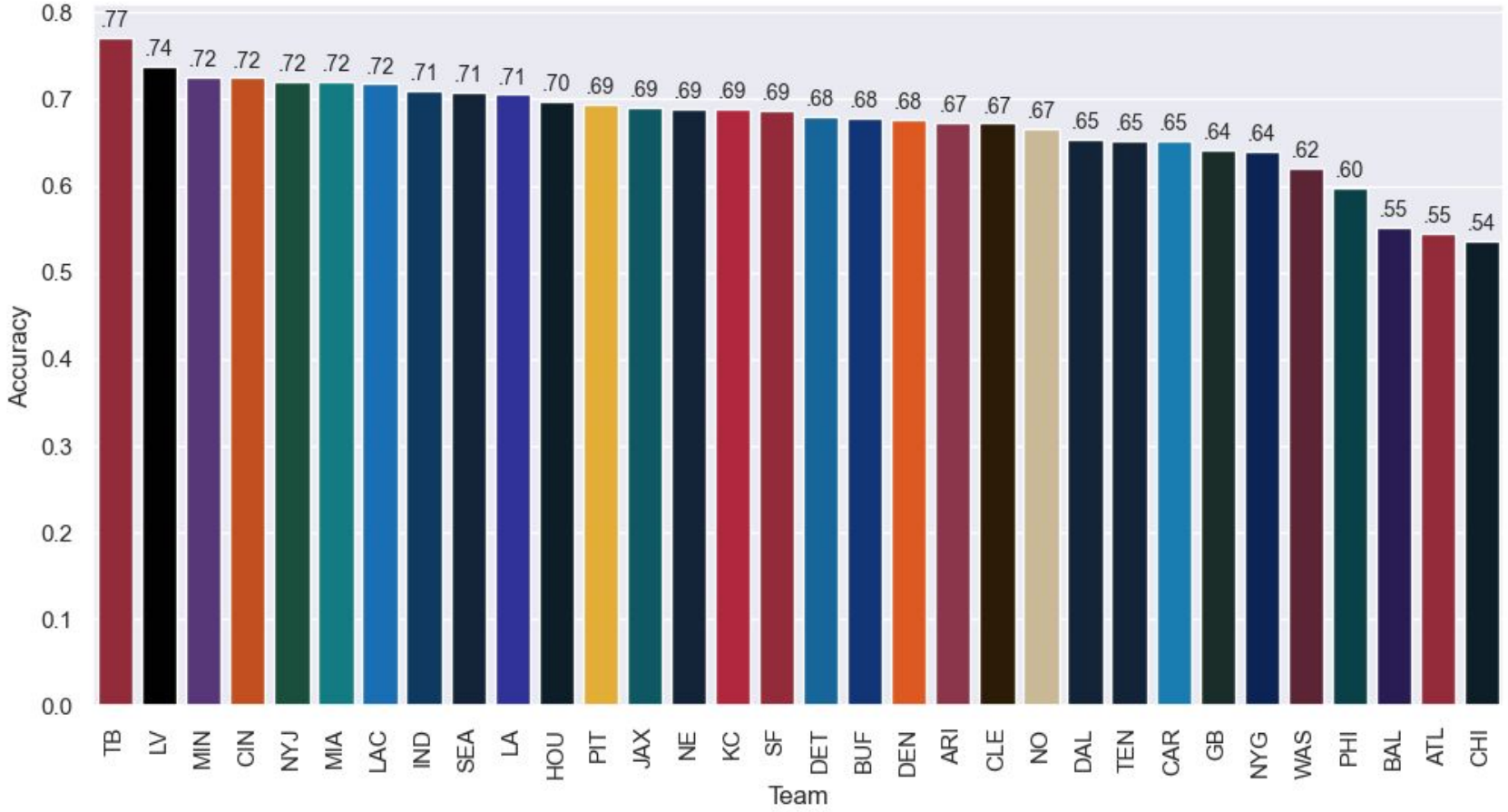


Data Visualizations

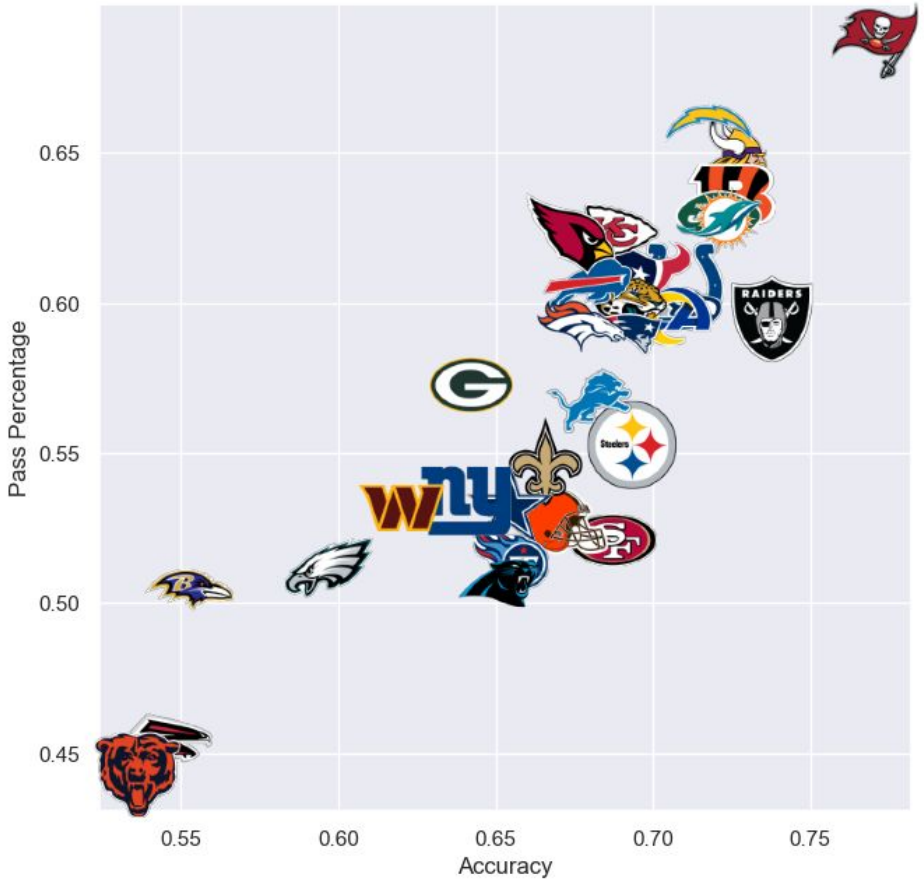


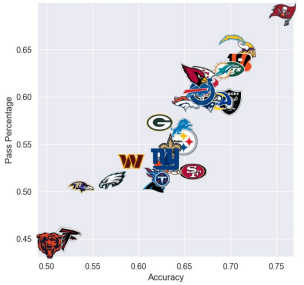


Data Visualizations

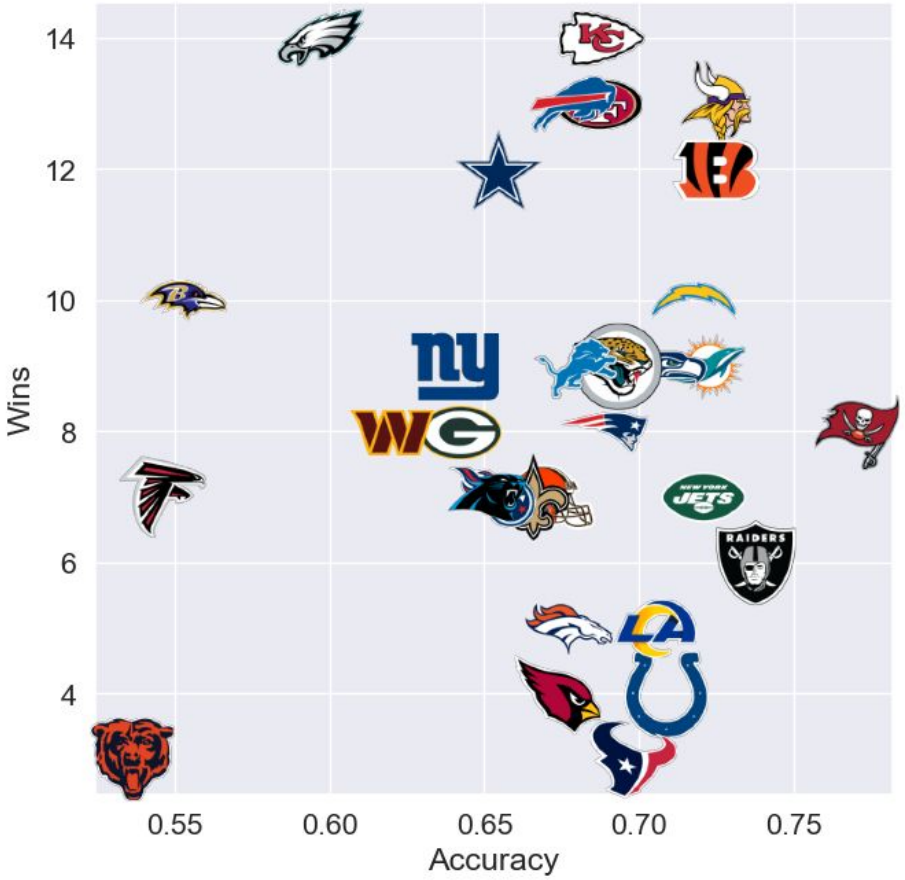


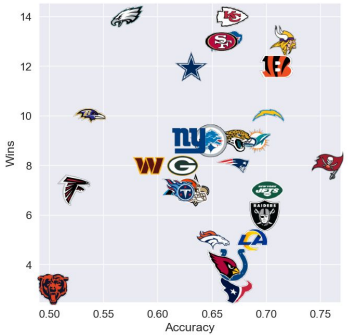
Data Visualizations





Data Visualizations





Data Visualizations

Recommendations and Next Steps

Gradient Boosting (*XGBoost*)

- Ensemble of decision trees like random forest
- New trees focus on previous trees mistakes
- Increased performance with unbalanced datasets

Recurrent Neural Network

- Each individual play sample includes previous plays
- LSTM allows selective retaining and releasing of previous data
- Previous research saw increased performance with play history

Further Investigation

- Imbalanced (pass-heavy) predictions
- Relationship between predictability and win percentage
- Game impact of mispredicted run plays vs. pass plays

Thank you!