A Data Science Approach to NFL Pre-Snap Evaluation

Ranier Gran - DS 504 Final Project

Overview

1. Project Introduction
2. Data Description
3. Exploratory Data Analysis
4. Data Preprocessing
5. Prediction Methodology
6. Results & Conclusions
7. Future Work

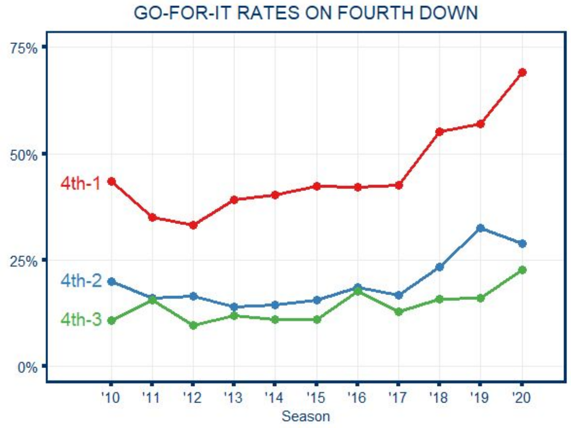
Project Introduction

How to make better football decisions?

- Data and analytics have been hot topics for the sports industry.

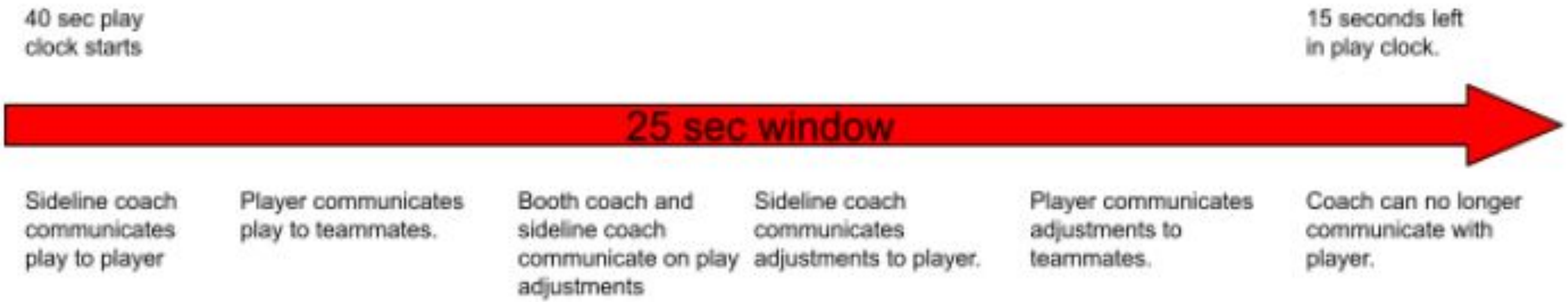
- The NFL is a game of inches
  - Coaches and players will do nearly anything for incremental improvements to their team
    - The Bad: Filming Sidelines, Banging on Trash Cans
    - The Good: Data Driven Decisions

- EX: With the help of analytics, teams have been attempting 4th down conversions at a much higher rate



Opportunities to make in game adjustments fueled by data

- Coaches are able to communicate with a headset to single player on offense and defense
- With a limited play clock and window for communication, decisions and adjustments must be made fast
- If we are able to predict a play’s performance with ML, we can make faster adjustments.



Other Data Driven Opportunities

- Pre-Game Planning
  - Identify which types of coverages a team leans toward
  - Practice defensive looks opposing QB has performed worse against
- Player Scouting
  - Learn player's strengths and weaknesses to find the best players for your team
  - Exploit opposing player's weaknesses more effectively

Data Description

Data

- For our analysis we used the Kaggle NFL Big Data Bowl 2021
  - https://www.kaggle.com/c/nfl-big-data-bowl-2021/data
  - 4 sets of data tables
    - Games
    - Players
    - Plays
    - Tracking

Games

- Games data structure:

gameId	gameDate	gameTimeEastern	homeTeamAbbr	visitorTeamAbbr	week
2018090600	09/06/2018	20:20:00	PHI	ATL	1
2018090901	09/09/2018	13:00:00	CLE	PIT	1
2018090902	09/09/2018	13:00:00	IND	CIN	1
2018090903	09/09/2018	13:00:00	MIA	TEN	1
2018090900	09/09/2018	13:00:00	BAL	BUF	1

- 253 games
- This will be a reference table only
- Game time and week number could be used as additional information
- Team abbreviations will be used for joins

Players

- Players initial data structure (7 columns, 1303 players):

nflId	height	weight	birthDate	collegeName	position	displayName
2539334	72	190	1990-09-10	Washington	CB	Desmond Trufant
2539653	70	186	1988-11-01	Southeastern Louisiana	CB	Robert Alford
2543850	69	186	1991-12-18	Purdue	SS	Ricardo Allen
2555162	73	227	1994-11-04	Louisiana State	MLB	Deion Jones
2555255	75	232	1993-07-01	Minnesota	OLB	De'Vondre Campbell

- Height had some string values (ex: “6-2”)
- Birthdate had multiple formats (ex: YYYY-MM-DD vs MM/DD/YYYY)
- collegeName and displayName will not be used
- One-Hot-Encoded position
- Updated Players data structure (29 columns):

nflId	height	weight	birthDate	collegeName	position	displayName	Pos_CB	Pos_DB	Pos_DE	...	Pos_NT	Pos_OLB	Pos_P	Pos_QB	Pos_RB	Pos_S	Pos_SS	Pos_TE	Pos_WR	age
2539334	6.000000	190	1990-09-10	Washington	CB	Desmond Trufant	1	0	0	...	0	0	0	0	0	0	0	0	0	28.305556
2539653	5.833333	186	1988-11-01	Southeastern Louisiana	CB	Robert Alford	1	0	0	...	0	0	0	0	0	0	0	0	0	30.163889
2543850	5.750000	186	1991-12-18	Purdue	SS	Ricardo Allen	0	0	0	...	0	0	0	0	0	0	1	0	0	27.033333
2555162	6.083333	227	1994-11-04	Louisiana State	MLB	Deion Jones	0	0	0	...	0	0	0	0	0	0	0	0	0	24.155556
2555255	6.250000	232	1993-07-01	Minnesota	OLB	De'Vondre Campbell	0	0	0	...	0	1	0	0	0	0	0	0	0	25.497222

Plays

- Plays initial data structure (27 columns, 19239 plays):

gameId	playId	playDescription	quarter	down	yardsToGo	possessionTeam	playType	yardlineSide	yardlineNumber	...	preSnapHomeScore	gameClock	absoluteYardlineNumber	penaltyCodes	penaltyJerseyNumbers	passResult	offensePlayResult	playResult	epa	isDefensivePI
2018090600	75	(15:00) M.Ryan pass short right to J.Jones pus...	1	1	15	ATL	play_type_pass	ATL	20	...	0.0	15:00:00	90.0	NaN	NaN	C	10	10	0.261827	False
2018090600	146	(13:10) M.Ryan pass incomplete short right to ...	1	1	10	ATL	play_type_pass	PHI	39	...	0.0	13:10:00	49.0	NaN	NaN	I	0	0	-0.372360	False
2018090600	168	(13:05) (Shotgun) M.Ryan pass incomplete short...	1	2	10	ATL	play_type_pass	PHI	39	...	0.0	13:05:00	49.0	NaN	NaN	I	0	0	-0.702779	False
2018090600	190	(13:01) (Shotgun) M.Ryan pass deep left to J.J...	1	3	10	ATL	play_type_pass	PHI	39	...	0.0	13:01:00	49.0	NaN	NaN	C	33	33	3.047530	False
2018090600	256	(10:59) (Shotgun) M.Ryan pass incomplete short...	1	3	1	ATL	play_type_pass	PHI	1	...	0.0	10:59:00	11.0	NaN	NaN	I	0	0	-0.842272	False

- gameClock and quarter fields combined to measure time left in game
- Personnel columns and offensive formation column one hot encoded

Plays cont.

- Several columns from Plays table removed

Column Removed	Reason for Removal
playDescription	String too long and variable for one-hot-encoding. Information reveals play result. Impossible to know before play for prediction. (Target Leakage)
playType	Information reveals play result. Impossible to know before play for prediction. (Target Leakage)
yardlineSide	Covered by another variable (absoluteYardlineNumber)
yardlineNumber	Covered by another variable (absoluteYardlineNumber)
typeDropback	Impossible to know before play for prediction. (Target Leakage)
penaltyCodes	Information reveals play result. Impossible to know before play for prediction. (Target Leakage)
penaltyJerseyNumbers	Information reveals play result. Impossible to know before play for prediction. (Target Leakage)

- Updated Plays data structure (135 columns):

gameId	playId	down	yardsToGo	possessionTeam	defendersInTheBox	numberOfPassRushers	preSnapVisitorScore	preSnapHomeScore	absoluteYardlineNumber	...	personnelD_5 DL, 2 LB, 4 DB	personnelD_5 DL, 3 LB, 2 DB, 1 OL	personnelD_5 DL, 3 LB, 3 DB	personnelD_5 DL, 4 LB, 2 DB	personnelD_5 DL, 5 LB, 1 DB	personnelD_6 DL, 1 LB, 4 DB	personnelD_6 DL, 2 LB, 3 DB	personnelD_6 DL, 3 LB, 2 DB	personnelD_6 DL, 4 LB, 1 DB	personnelD_7 DL, 3 LB, 1 DB
2018090600	75	1	15	ATL	7.0	4.0	0.0	0.0	90.0	...	0	0	0	0	0	0	0	0	0	0
2018090600	146	1	10	ATL	7.0	4.0	0.0	0.0	49.0	...	0	0	0	0	0	0	0	0	0	0
2018090600	168	2	10	ATL	6.0	4.0	0.0	0.0	49.0	...	0	0	0	0	0	0	0	0	0	0
2018090600	190	3	10	ATL	6.0	5.0	0.0	0.0	49.0	...	0	0	0	0	0	0	0	0	0	0
2018090600	256	3	1	ATL	8.0	6.0	0.0	0.0	11.0	...	0	0	0	0	0	0	0	1	0	0

Tracking Data

- Tracking initial data structure (20 columns, 18309388 time stamped tracking events):

time	x	y	s	a	dis	o	dir	event	nflId	displayName	jerseyNumber	position	frameId	team	gameId	playId	playDirection	route
2018-09-07T01:07:14.599Z	91.73	26.67	0.00	0.01	0.02	289.57	240.93	None	310.0	Matt Ryan	2.0	QB	1	away	2018090600	75	left	NaN
2018-09-07T01:07:14.599Z	88.89	36.47	0.01	0.01	0.01	105.63	66.66	None	79848.0	Malcolm Jenkins	27.0	SS	1	home	2018090600	75	left	NaN
2018-09-07T01:07:14.599Z	91.35	44.16	0.02	0.03	0.01	290.45	16.86	None	2495454.0	Julio Jones	11.0	WR	1	away	2018090600	75	left	HITCH
2018-09-07T01:07:14.599Z	86.31	22.01	0.09	0.42	0.01	70.12	168.91	None	2495613.0	Corey Graham	24.0	FS	1	home	2018090600	75	left	NaN
2018-09-07T01:07:14.599Z	90.78	36.15	0.00	0.00	0.00	257.61	193.97	None	2533040.0	Mohamed Sanu	12.0	WR	1	away	2018090600	75	left	HITCH

- Average player speed (s) and acceleration (a) calculated over all tracking data, joined to player table.

nflId	height	weight	birthDate	collegeName	position	displayName	age	s	a
2539334	6.000000	190	1990-09-10	Washington	CB	Desmond Trufant	28.305556	3.073965	1.891416
2539653	5.833333	186	1988-11-01	Southeastern Louisiana	CB	Robert Alford	30.163889	3.214184	1.916010
2543850	5.750000	186	1991-12-18	Purdue	SS	Ricardo Allen	27.033333	3.038352	1.928427
2555162	6.083333	227	1994-11-04	Louisiana State	MLB	Deion Jones	24.155556	2.794986	1.860145
2555255	6.250000	232	1993-07-01	Minnesota	OLB	De'Vondre Campbell	25.497222	2.985769	1.969998

Exploratory Data Analysis

Player Scouting, who to target?

- Players with highest speed are DBs and WRs (this makes sense!)
- Players with lowest speed are Kickers and interior linemen (this also makes sense!)
- Sorted by acceleration we see defensive ends coming off the edge like Myles Garrett.
  - Providing additional help on his side of the offensive line could help avoid QB pressures.

displayName	position	s	a
A.J. Moore	DB	6.509091	1.979899
Johnny Holton	WR	5.692356	2.300604
Dallin Leavitt	DB	5.623514	1.624797
Malachi Dupre	WR	4.960952	2.195952
Gehrig Dieter	WR	4.940460	2.492068
...	...	...	...
Kameron Canaday	LS	0.736042	0.863958
Bilal Nichols	DT	0.662642	0.914906
Cameron Heyward	DT	0.618542	0.576875
Don Muhlbach	LS	0.490000	0.614909
Matt Prater	K	0.452909	0.791455

displayName	position	s	a
Kasim Edebali	OLB	4.432393	2.923590
Myles Garrett	DE	2.707925	2.768679
Solomon Thomas	DE	4.157455	2.766182
Carl Nassib	OLB	3.885397	2.689048
Jehu Chesson	WR	3.489721	2.678212
...	...	...	...
Sam Martin	P	1.090364	0.747091
Don Muhlbach	LS	0.490000	0.614909
Cameron Heyward	DT	0.618542	0.576875
Hunter Bradley	LS	0.969649	0.562281
Thomas Morstead	P	1.065909	0.311061

Team Scouting, how to attack?

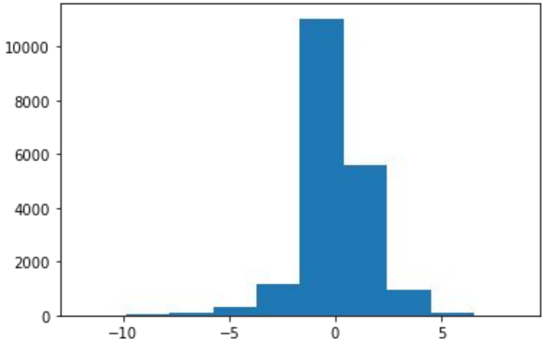
- Understanding how teams typically play defensively can improve how you attack them on offense.
- Teams who stack the box like the Arizona Cardinals are susceptible to throws over the top but are harder to run the ball against.
- Teams with less defenders in the box like the Detroit Lions can leave open cushions for short and intermediate pass plays but are more conservative against the deep ball.

defendingTeam	defendersInTheBox	numberOfPassRushers
DET	5.571942	4.003711
PIT	5.676236	4.417349
LAC	5.711033	4.023423
CHI	5.729290	4.074018
IND	5.748311	4.041451
KC	5.932024	4.218354

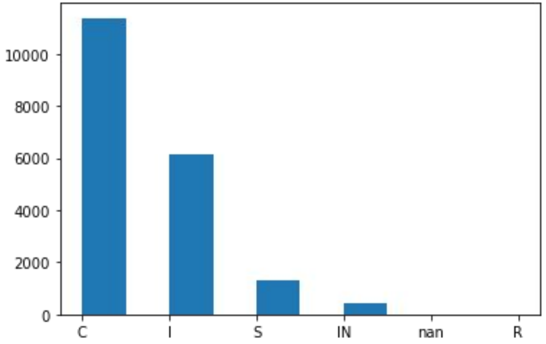
defendingTeam	defendersInTheBox	numberOfPassRushers
ARI	6.433333	4.422868
CAR	6.307958	4.369449
MIN	6.267717	4.172764
MIA	6.206093	4.262082
NO	6.199067	4.317590
CLE	6.197059	4.338906

Advanced Analytics, What to predict?

- Possible Target Variables Identified:
  - 'passResult', 'offensePlayResult', 'playResult', 'epa', 'isDefensivePI'
- In game, scoring is obviously key, focusing on epa (Expected Points Added) can allow us to judge how successful a play will be offensively or defensively.

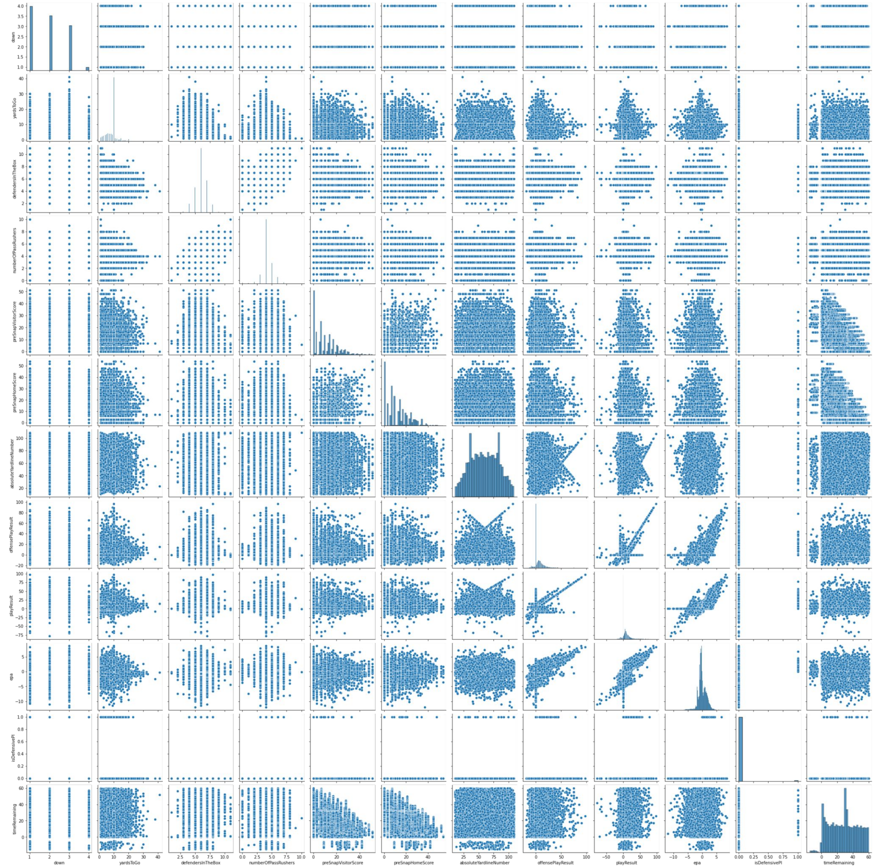


- From a classification standpoint, pass result could also be a reasonable target. Recognizing when a play would result in a catch (C), incomplete pass (I), sack (S), or interception (IN) before the snap could lead to important adjustments especially in clutch areas of the game.
  - Giving up a sack on 3rd down is much worse than throwing an incomplete pass



- Play Result predictors could be used for post play predictions to help judge the impact after a play.

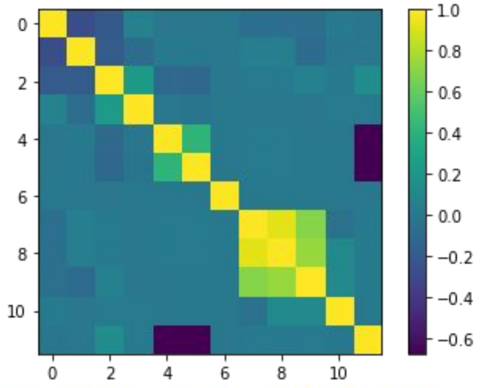
Variable Correlation, what leads to a big play?



- Generated Pairwise plot of non-player, non-one-hot encoded variables:
  'down',
  'yardsToGo',
  'defendersInTheBox',
  'numberOfPassRushers',
  'preSnapVisitorScore',
  'preSnapHomeScore',
  'absoluteYardlineNumber',
  'passResult',
  'offensePlayResult',
  'playResult',
  'epa',
  'isDefensivePI',
  'timeRemaining'
- From this we can see that not many variables are correlated. Some of the target variables are correlated to one another.

Variable Correlation cont.

- Looking at these from a numerical perspective, we confirm that the only variables that really have correlation are the predictor variables.
- EPA is driven by the pass result



	down	yardsToGo	defendersInTheBox	numberOfPassRushers	preSnapVisitorScore	preSnapHomeScore	absoluteYardlineNumber	offensePlayResult	playResult	epa	isDefensivePI	timeRemaining
down	1.000000	-0.276123	-0.198178	0.069329	-0.013084	-0.007228	0.002834	-0.063911	-0.066480	-0.056965	0.014154	-0.015483
yardsToGo	-0.276123	1.000000	-0.181236	-0.079169	0.010068	0.003595	-0.002560	0.040389	0.039519	-0.092148	-0.012971	-0.009049
defendersInTheBox	-0.198178	-0.181236	1.000000	0.237950	-0.097810	-0.115826	-0.008001	0.007371	0.009647	0.063781	0.009171	0.148415
numberOfPassRushers	0.069329	-0.079169	0.237950	1.000000	-0.012711	-0.021216	0.000634	-0.001797	-0.002532	-0.008584	0.014346	0.012173
preSnapVisitorScore	-0.013084	0.010068	-0.097810	-0.012711	1.000000	0.419709	0.004587	0.006263	0.005385	-0.005007	-0.006495	-0.675409
preSnapHomeScore	-0.007228	0.003595	-0.115826	-0.021216	0.419709	1.000000	-0.001514	0.011337	0.009910	0.004068	0.003077	-0.669392
absoluteYardlineNumber	0.002834	-0.002560	-0.008001	0.000634	0.004587	-0.001514	1.000000	0.002485	0.001031	0.002367	0.002519	-0.004421
offensePlayResult	-0.063911	0.040389	0.007371	-0.001797	0.006263	0.011337	0.002485	1.000000	0.931035	0.696615	-0.045012	0.013253
playResult	-0.066480	0.039519	0.009647	-0.002532	0.005385	0.009910	0.001031	0.931035	1.000000	0.744617	0.123340	0.014616
epa	-0.056965	-0.092148	0.063781	-0.008584	-0.005007	0.004068	0.002367	0.696615	0.744617	1.000000	0.115152	0.019624
isDefensivePI	0.014154	-0.012971	0.009171	0.014346	-0.006495	0.003077	0.002519	-0.045012	0.123340	0.115152	1.000000	0.001280
timeRemaining	-0.015483	-0.009049	0.148415	0.012173	-0.675409	-0.669392	-0.004421	0.013253	0.014616	0.019624	0.001280	1.000000

Data Preprocessing

Full Data Structure

- For each play, all players involved are gathered by position with their height, weight, age, speed, and acceleration.
- Up to 22 players were tracked per play. (11 on defense, 11 on offense)
- Positions not tracked on a given play were labeled NA and later imputed with 0's
- 19239 plays, 145 columns

gameId	playId	down	yardsToGo	offenseFormation	personnelO	defendersInTheBox	numberOfPassRushers	personnelD	preSnapVisitorScore	...	K1weight	K1age	K1s	K1a	DT1	DT1height	DT1weight	DT1age	DT1s	DT1a
2018090600	75	1	15	I_FORM	2 RB, 1 TE, 2 WR	7.0	4.0	4 DL, 2 LB, 5 DB	0.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2018090600	146	1	10	SINGLEBACK	1 RB, 1 TE, 3 WR	7.0	4.0	4 DL, 2 LB, 5 DB	0.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2018090600	168	2	10	SHOTGUN	2 RB, 1 TE, 2 WR	6.0	4.0	4 DL, 2 LB, 5 DB	0.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2018090600	190	3	10	SHOTGUN	1 RB, 1 TE, 3 WR	6.0	5.0	4 DL, 1 LB, 6 DB	0.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2018090600	256	3	1	SHOTGUN	2 RB, 3 TE, 0 WR	8.0	6.0	6 DL, 3 LB, 2 DB	0.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

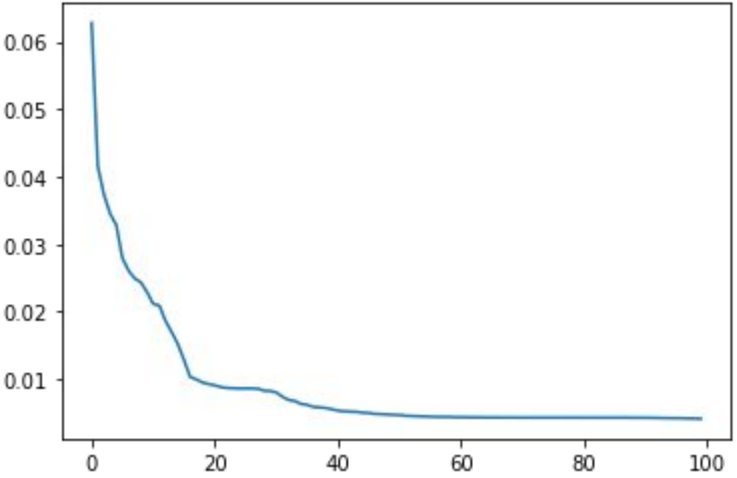
Final Full Data Structure

- Categorical Variables are one hot encoded
- Missing Values imputed with 0's
- All Variables are Z-transformed to normalize columns
- Unnecessary columns are removed
  - gameId, playId
- Predictor Variables are separated
- 19239 plays, 237 columns

	down	yardsToGo	defendersInTheBox	numberOfPassRushers	preSnapVisitorScore	preSnapHomeScore	absoluteYardlineNumber	timeRemaining	QB1height	QB1weight	...	personnelD_6 DL, 1 LB, 4 DB	personnelD_6 DL, 2 LB, 3 DB	personnelD_6 DL, 3 LB, 2 DB	personnelD_6 DL, 4 LB, 1 DB	personnelD_7 DL, 3 LB, 1 DB
0	-1.084262	1.528919	0.918898	-0.067085	-1.058652	-1.135487	1.253081	1.833042	0.242012	-0.425798	...	-0.010196	-0.012488	-0.024982	-0.024982	-0.00721
1	-1.084262	0.270776	0.918898	-0.067085	-1.058652	-1.135487	-0.348228	1.730990	0.242012	-0.425798	...	-0.010196	-0.012488	-0.024982	-0.024982	-0.00721
2	0.093876	0.270776	-0.015641	-0.067085	-1.058652	-1.135487	-0.348228	1.726352	0.242012	-0.425798	...	-0.010196	-0.012488	-0.024982	-0.024982	-0.00721
3	1.272015	0.270776	-0.015641	0.795075	-1.058652	-1.135487	-0.348228	1.722641	0.242012	-0.425798	...	-0.010196	-0.012488	-0.024982	-0.024982	-0.00721
4	1.272015	-1.993880	1.853438	1.657235	-1.058652	-1.135487	-1.832369	1.609456	0.242012	-0.425798	...	-0.010196	-0.012488	40.028115	-0.024982	-0.00721

Dimension Reduction

- With so many predictor columns we utilize PCA to reduce the dimensionality
- PCA Scree Plot shows us we can use closer to 30 principle component predictors
  - Value of 35 selected to explain most of the variance in the data for modeling



Prediction Methodology

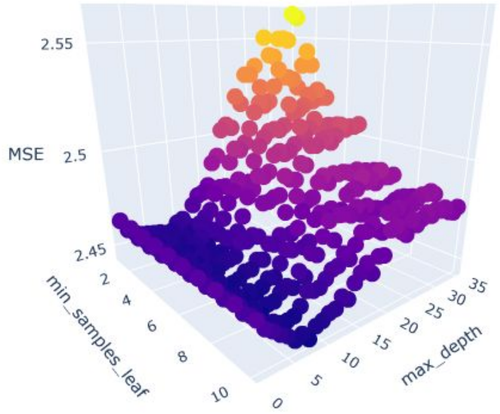
Four ways of Foresting

Only knowing pre-snap information, can we predict EPA?

- Utilizing our Principle Component Dataset we train a random forest regression model.
- Cross Validation:
  - Train:test split 66:33
  - Hyperparameter Tuning - min_samples_leaf, max_depth
- Error Metric: Mean Squared Error (MSE)
- The Baseline:
  - Null model (always predicting the mean EPA) 2.499 MSE
- Best Model:

min_samples_leaf	max_depth	Test MSE
8	7	2.446

- Model Improvement from Baseline: 2% Increase in Accuracy

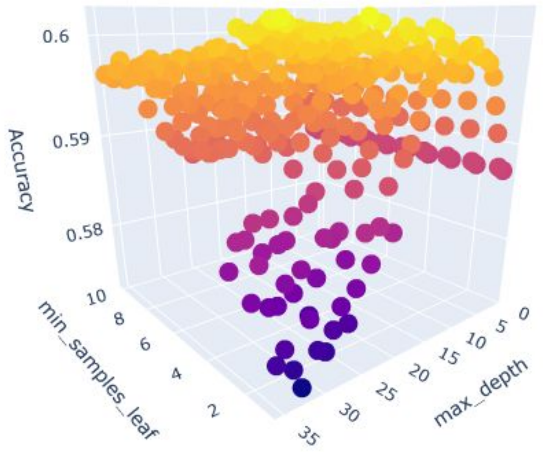


Only knowing pre-snap information, can we predict the pass result (C,I,S,INT)?

- Utilizing our Principle Component Dataset we train a random forest classification model.
- Cross Validation:
  - Train:test split 66:33
  - Hyperparameter Tuning - min_samples_leaf, max_depth
- Error Metric: Accuracy
- The Baseline:
  - Null model (always predicting a catch) 59.1% accuracy
- Best Model:

min_samples_leaf	max_depth	Test Accuracy
4	13	60.2%

- Model Improvement from Baseline: 1.1% Increase in Accuracy

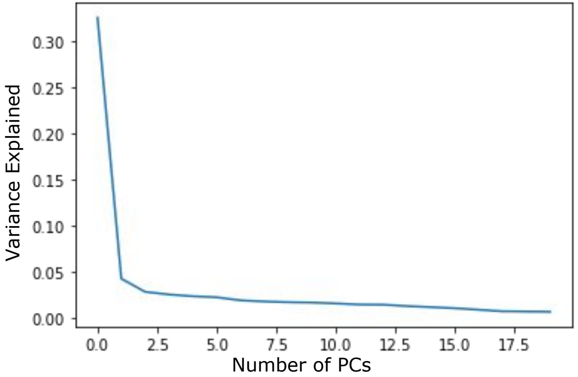


Knowing pre-snap information and play result, can we predict EPA?

- This time including play result information we recalculate the PCA. This results in a lower amount of PC’s used to explain the variance: 7
- Cross Validation:
  - Train:test split 66:33
  - Hyperparameter Tuning: kept fixed from previous regression
- Error Metric: Mean Squared Error (MSE)
- The Baseline:
  - Null model (always predicting the mean EPA) 2.499 MSE
- Best Model:

min_samples_leaf	max_depth	Test MSE
8	7	0.730

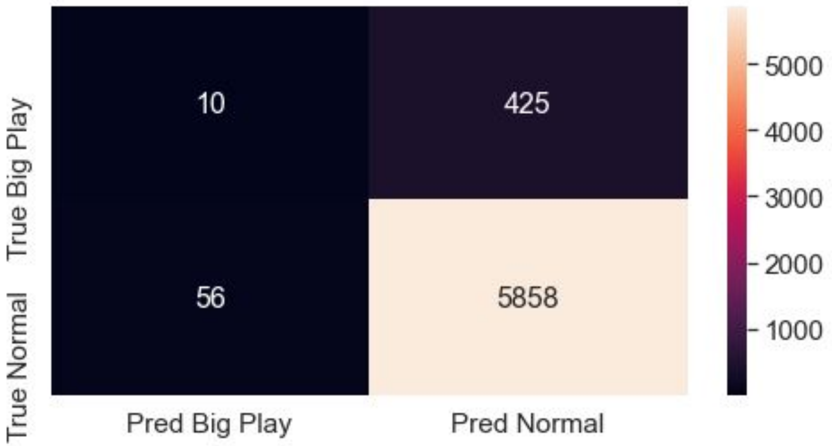
- Model Improvement from Baseline: 70% Increase in Accuracy



Big Plays are outliers, but important to defend, can we predict them?

- Back to our original PCA data without play result, we train an Isolation Forest to identify outliers.
- Big plays must be defined
  - EPA ≥ 2 and absoluteYardlineNumber > 40
  - These are plays that start at midfield or further and result in an almost guaranteed field goal at minimum
- Cross Validation:
  - Train:test split 66:33
- Test Results:

```
Accuracy: 0.924
Precision: 0.932
Recall: 0.991
f1 Score: 0.961
Area Under the Precision-Recall Curve (AUPRC): 0.932
```



Results and Conclusions

How do we make an impact on a team?

Comparing our models, what is helpful?

- We began with a goal of predicting the play result with only pre-snap information.
  - Our Regression technique on EPA only provided a 2% improvement from the mean
  - Our Classification technique on Pass Result only provided a 1% improvement from predicting a Catch on every play
- While the NFL is a game of inches, and these minimal improvements could be slightly helpful, they aren’t strong enough to push to make considerable change.

Model Type	min_samples_leaf	max_depth	Improvement from Baseline
Regression (EPA)	8	7	2%
Classification (Pass Result)	4	13	1.1%
Regression (EPA post play)	8	7	70%

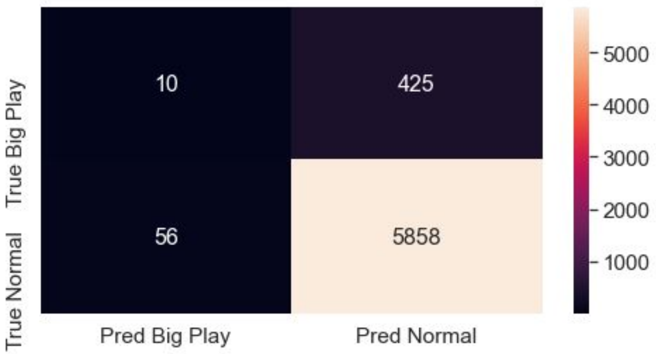
Comparing our models, what is helpful?

- Next we looked at a post-play model. Understanding EPA after a play can change how a team might respond on the next play.
  - This model provided a 70% improvement in accuracy as compared to the mean
  - Using this model, if a defending team gave up a play with an exceptionally high predicted EPA, they might look to change their strategy on the next play, or in the next time they find themselves in a similar position on the field.

Model Type	min_samples_leaf	max_depth	Improvement from Baseline
Regression (EPA)	8	7	2%
Classification (Pass Result)	4	13	1.1%
Regression (EPA post play)	8	7	70%

Comparing our models, what is helpful?

- Finally we looked at an outlier detection model to identify big plays
  - This model's accuracy is fairly high, however it does miss out on a considerable amount of big plays
- However, stopping any big play can be game changing.
  - A model like this could be implemented in clutch situations to adjust a formation, change an assignment, or even call a timeout to stop the play from happening at all.
- Many of the ones it does catch occur on 4th down and some are even trick plays like fake punts
  - (10:00) (Punt formation) K.Byard pass deep right to D.Cruikshank for 66 yards, TOUCHDOWN.
  - (4:12) C.Clark reported in as eligible. C.McCaffrey pass short right to C.Manhertz for 50 yards, TOUCHDOWN.



Conclusions

- Predicting play results in the NFL is not an easy task. If it was, every team in the league would be implementing it and strategies would change drastically.
- Modeling can provide some incremental gains and a 2% improvement in play recognition could mean the difference between winning and losing a game.
  - These models could be used by coaches as a supplement to their current strategies rather than a replacement
- Post Play models could be useful in adjusting strategies to improve on the next play with 70% accuracy increases.
- While only a small percentage of big plays can be predicted, stopping a single big play could change the tides of a game.
  - Implementing the outlier detection model would have limited risk while providing immediate impact

Future Work

Future Work

- Test out other modeling techniques

- Run with parallel processing to improve runtime
  - Project was completed using Python Jupyter Notebooks and local memory

- Test model predicting on live games

References

- Background information has been gathered from previous research as a part of humanities and arts capstone as well as other sources.
  - Capstone - Sports Analytics The Story Behind the Numbers: https://drive.google.com/file/d/13F2bQRG8Zz80xTZIVGL0ldU-rpSGrcgT/view?usp=sharing
  - Headset Rules: https://gethypedsports.com/do-football-players-have-speakers-microphones-in-their-helmet/#:~:text=There%20are%20a%20set%20of,what%20the%20coach%20is%20saying
  - 4th down statistics: https://footballscoop.com/news/nfl-teams-going-for-it-on-fourth-down-more-often

Thank you! Please feel free to reach out with any questions, comments, or concerns!

rxgran@wpi.edu
raniergran@gmail.com