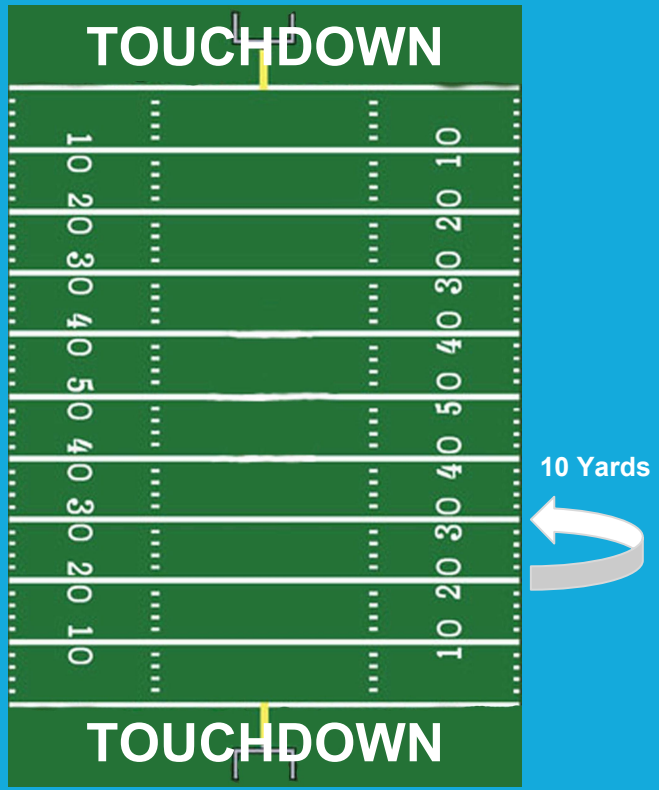
NFL Play Predictions

1007 Final Project Presentation

Ziv Schwartz, zs1349
Santiago Novoa, smn405
Esteban Navarro Garaiz, eng272
Charles Brillo-Sonnino, cbs488

Quick Overview of American Football

- Team on offense can either **run** or **pass** the ball to move it down the field
- Each team has 4 attempts to move the ball 10 yards known as **downs**
- If the team does not succeed, they give the ball to the opponent
- Generally, the team uses the fourth “down” to kick the ball away and push the opponent further from where they score
- Third down is the “make-or-break” moment for whether moving the ball down the field stalls or continues



Problem Focus : Third Down

= 2018 Division Leaders

Rk	Team	G	Pts/G	TotPts	Scrm Plys	Yds/G	Yds/P	1st/G	3rd Md	3rd Att	3rd Pct	4th Md	4th Att	4th Pct	Pen	Pen Yds	ToP/G	FUM	Lost	TO
1	Tampa Bay Buccaneers	13	25.5	332	877	430.1	6.4	25.2	75	154	49	8	10	80	99	815	31:07	20	7	-17
2	Indianapolis Colts	13	26.8	349	868	382.4	5.7	22.7	84	178	47	6	14	43	96	773	29:18	14	7	+1
3	Baltimore Ravens	13	24.7	321	928	369.6	5.2	23.4	88	188	47	10	18	56	94	765	32:16	19	7	-6
4	Kansas City Chiefs	13	36.2	471	829	437.5	6.9	24.8	67	145	46	10	11	91	115	960	29:56	13	4	+6
5	Atlanta Falcons	13	24.3	316	821	375.0	5.9	22.0	77	167	46	10	19	53	82	723	29:40	22	9	-4
6	New Orleans Saints	13	34.4	447	829	389.0	6.1	24.2	68	149	46	12	14	86	70	674	32:17	16	8	+8
7	Pittsburgh Steelers	13	28.2	367	860	408.1	6.2	23.8	73	161	45	6	10	60	92	837	30:44	17	7	-8
8	Los Angeles Rams	13	32.7	425	848	422.5	6.5	24.4	66	154	43	6	13	46	79	705	30:18	11	4	+10
9	New England Patriots	13	28	364	880	397.1	5.9	23.3	71	168	42	6	10	60	70	559	30:48	10	6	+6
10	Tennessee Titans	13	19.3	251	771	319.1	5.4	18.2	71	171	42	6	10	60	67	607	30:20	17	5	-4

Offensive Scoring Rank

- → 10
- → 8
- → 1 ★
- → 2
- → 4
- → 3
- → 6

/32 teams

End of Week 14 teams by 3rd down conversion percentage

Goal: Predict play type (run or pass) on third down

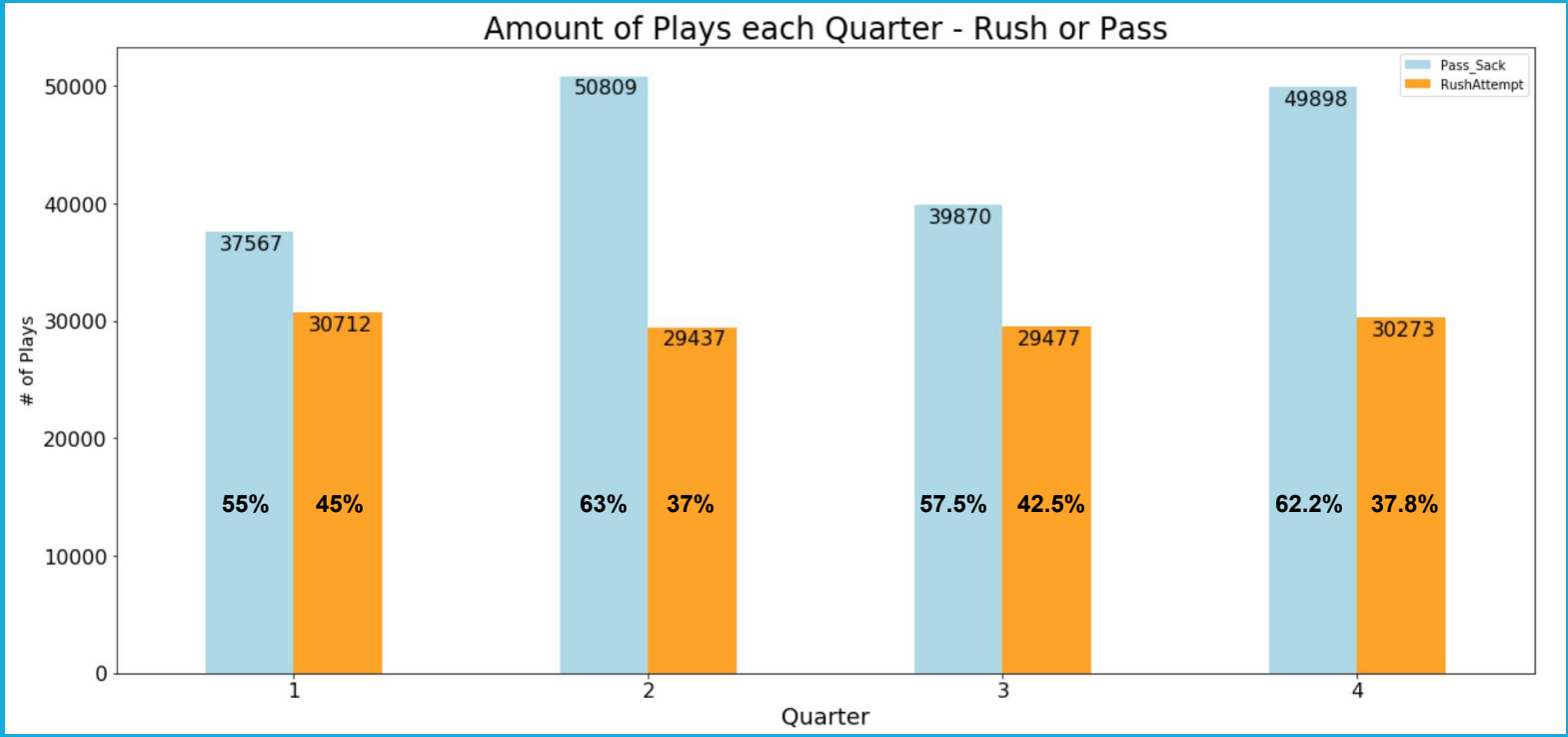
- Data compiled by Max Horowitz ：https://www.kaggle.com/maxhorowitz/nflplaybyplay2009to2016
- 407,688 instances with 102 features each
- Details every event from every game for 9 NFL seasons (2009-2017):
  - 256 games per season so 2304 games in total
  - 67,398 third downs
  - Many instances with “No Play”

Sample of Instances with “No Play”

- 10/29/2017, Detroit v. Pittsburgh, 3rd quarter, 3rd down : **'(11:03) (Shotgun) B.Roethlisberger pass incomplete short middle to J.James (Q.Diggs). PENALTY on DET-Q.Diggs, Unnecessary Roughness, 11 yards, enforced at DET 22 - No Play.'**
- 9/28/2017, Dallas v. Arizona, start of 2nd quarter : **'The game has been suspended. Field cleared temporarily due to impending lightning.'**
- 9/13/2009, Seattle v. St. Louis, 1st quarter : **Timeout #2 by STL at 06:51.**
- 1/1/2012, Baltimore v. Cincinnati, 3rd quarter : **END QUARTER 3**

Play Type by Quarter

Note: ‘Attempted’ Passing plays includes sacks. A sack is defined by a called passing play where the QB gets tackled before throwing the ball.



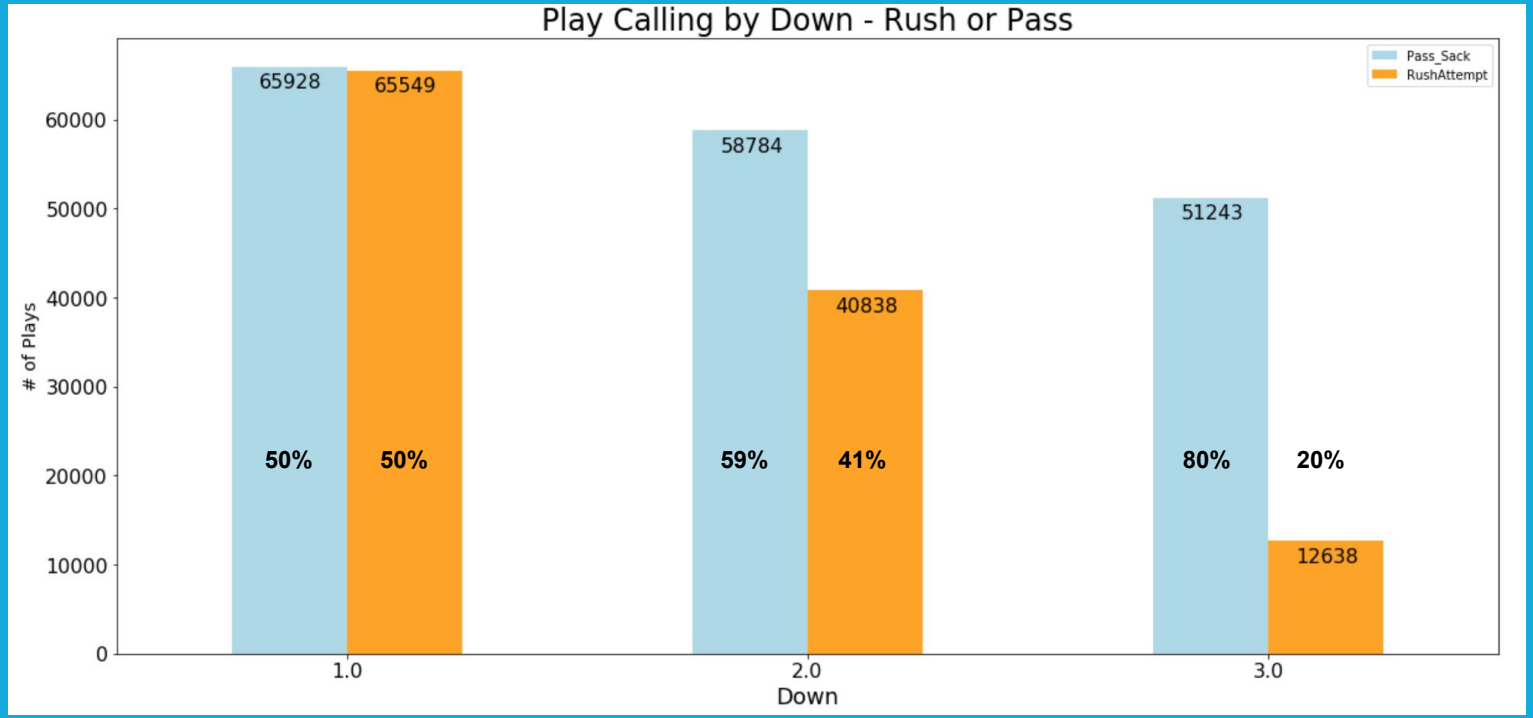
Passing plays are called more often , especially in the 2nd and 4th quarters

Plays Per Down

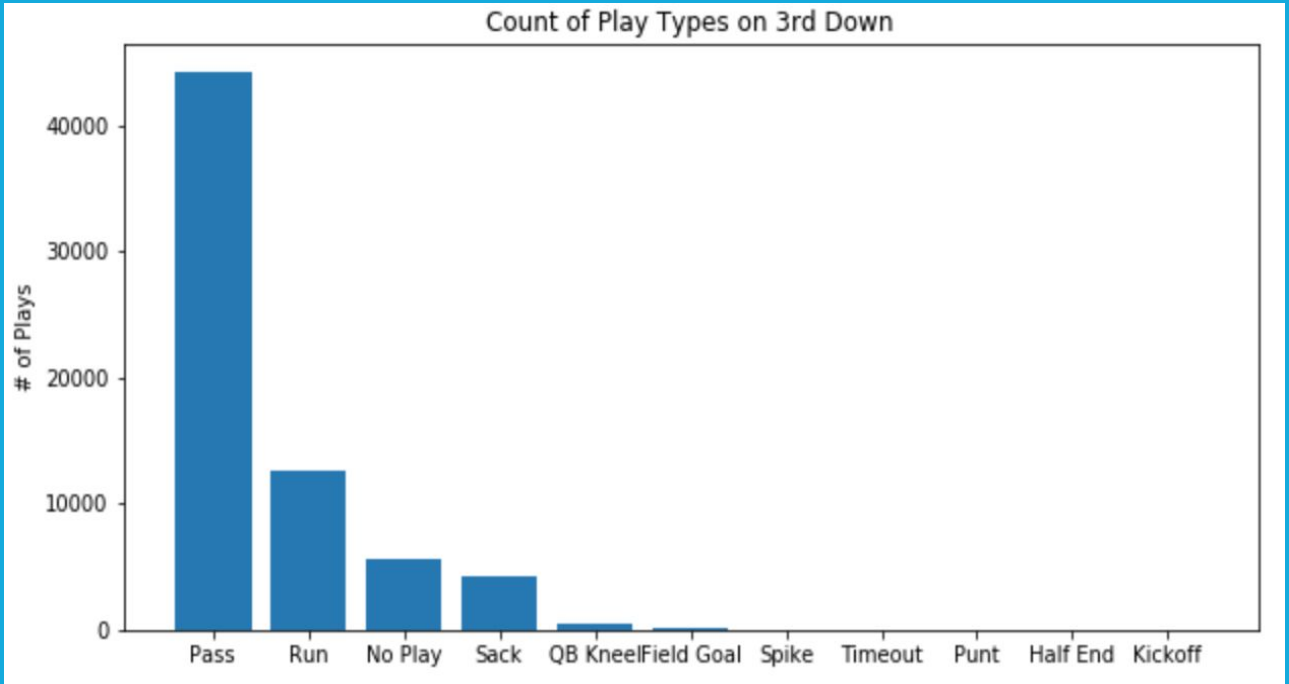
1st Down: Near 50/50 split on Pass and Run

2nd Down: 60/40 split on Pass and Run

3rd Down: 80/20 split on Pass and Run



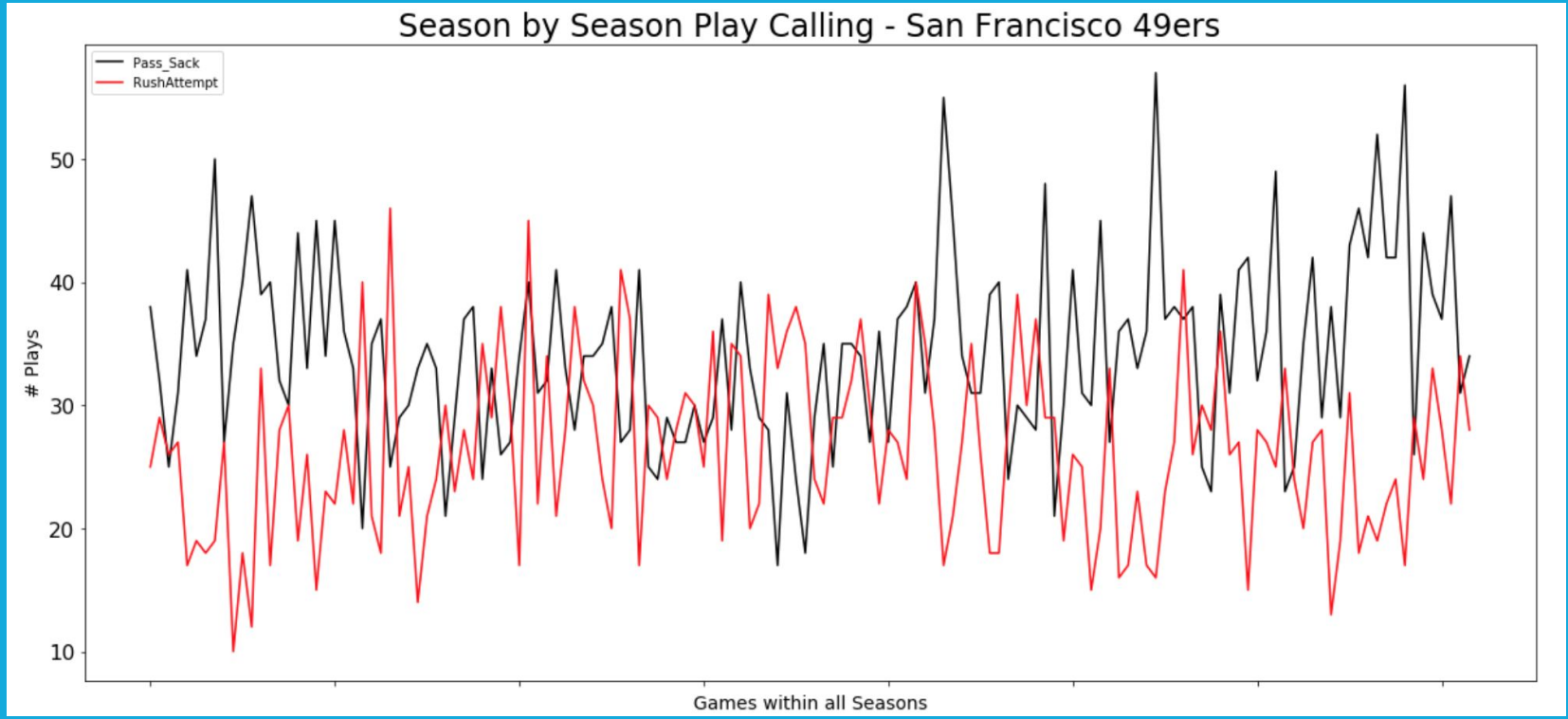
Third Down Play Calls



Pass	44229
Run	12638
No Play	5555
Sack	4200
QB Kneel	504
Field Goal	222
Spike	35
Timeout	8
Punt	5
Half End	1
Kickoff	1

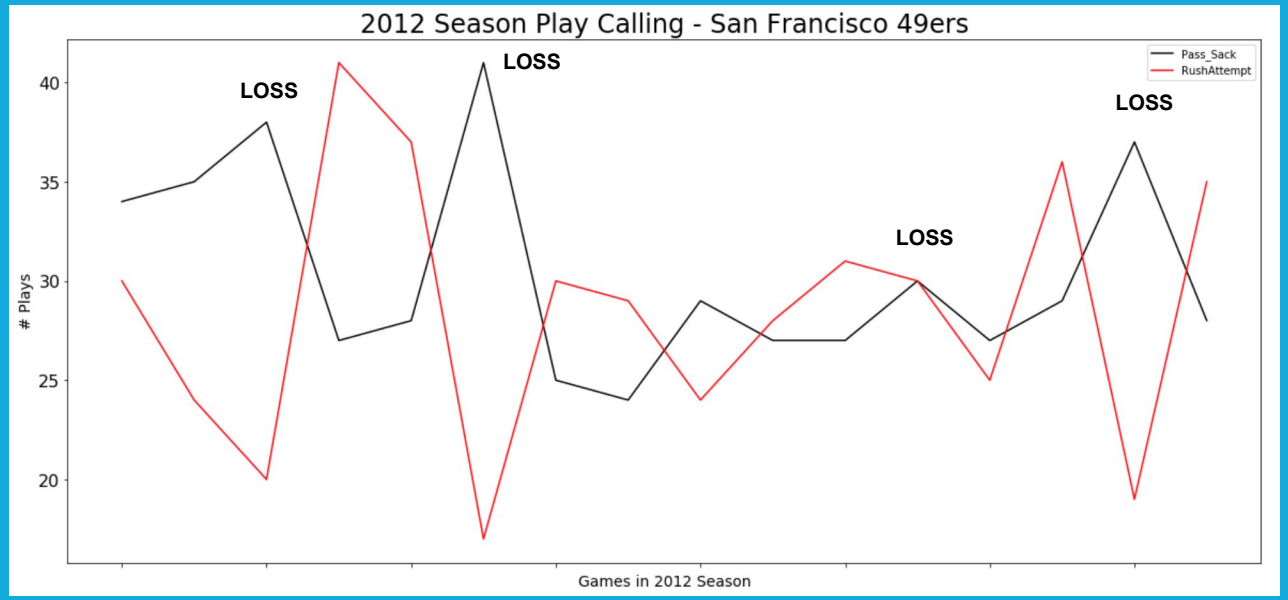
Kickoff Description: M.Schaub pass incomplete short right to C.Brown (V.Davis). Play was challenged by MIA, but play considered unreviewable.

All San Francisco 49ers Play Calls



Zoom: 49ers Play Calls - 2012 Season

Week	Date	Opponent	Result
1	September 9	at Green Bay Packers	**W** 30–22
2	September 16	Detroit Lions	**W** 27–19
3	September 23	at Minnesota Vikings	**L** 13–24
4	September 30	at New York Jets	**W** 34–0
5	October 7	Buffalo Bills	**W** 45–3
6	October 14	New York Giants	**L** 3–26
7	October 18	**Seattle Seahawks**	**W** 13–6
8	October 29	at **Arizona Cardinals**	**W** 24–3
9			*Bye*
10	November 11	**St. Louis Rams**	**T** 24–24 (OT)
11	November 19	Chicago Bears	**W** 32–7
12	November 25	at New Orleans Saints	**W** 31–21
13	December 2	at **St. Louis Rams**	**L** 13–16 (OT)
14	December 9	Miami Dolphins	**W** 27–13
15	December 16	at New England Patriots	**W** 41–34
16	December 23	at **Seattle Seahawks**	**L** 13–42
17	December 30	**Arizona Cardinals**	**W** 27–13



The 49ers lost 4 games during the 2012 Season: weeks 3, 6, 13, and 16. Within each of those games, the team attempts a significantly higher amount of passing plays, except for Week 13 where the game went to overtime.

3rd Down Offensive Means

```
r_off_agg = df3[(df3.PlayType == 'Run')]
p_off_agg = df3[(df3.PlayType == 'Pass')|(df3.PlayType == 'Sack')]
```

We can look specifically at the feature 'Yards Gained' to better understand how teams perform on average on 3rd down given the certain play-type they execute.

The table to the right contains the Rushing and Passing Means (in yards) on a 3rd downs during a specific game.

GameID	posteam	RushingMean	PassingMean
2009091000	PIT	0.000000	1.500000
2009091000	TEN	-1.333333	3.300000
2009091301	BAL	1.000000	8.750000
2009091301	KC	-0.500000	8.750000
2009091302	CAR	0.000000	2.230769

Feature Selection

Target Variable:

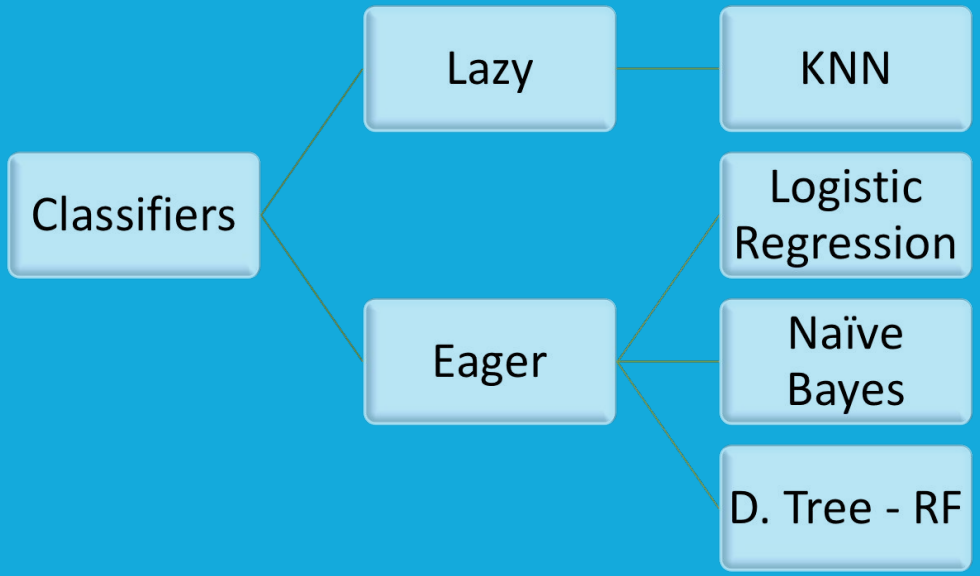
- **PlayType** - Type of Play. Pass+Sacks (1) or Run (0)

Features:

- **TimeSecs** - Time remaining in game in seconds
- **Yrdline100** - Distance to opponent's end-zone, ranges from 1-99
- **Ydstogo** - Yards to go for a first down
- **GoalToGo** - Binary. Goal down situation (1), else (0)
- **ScoreDiff** - The difference in score between the offensive and defensive teams (offensive.score - def.score). Shows if offensive team is ahead or behind.
- **Posteam_timeouts_pre** - Timeouts remaining for offensive team at the start of the play
- **Touchdown_Prob** - Probability of the possession team scoring a touchdown next
- **Field_Goal_Prob** - Probability of the possession team scoring a field goal next
- **Safety_Prob** - Probability of the possession team allowing a safety next
- **2Min** - Binary for 2 or less minutes remaining in half (1), else(0)
- **Run_avg2** - Moving average of net yards gained from running plays for team per game
- **Pass_avg2** - Moving average of net yards gained from passing plays for team per game

Predictive Modeling

Classification predictive modeling is the task of approximating a mapping function (f) from input variables (X) to discrete output variables (y).

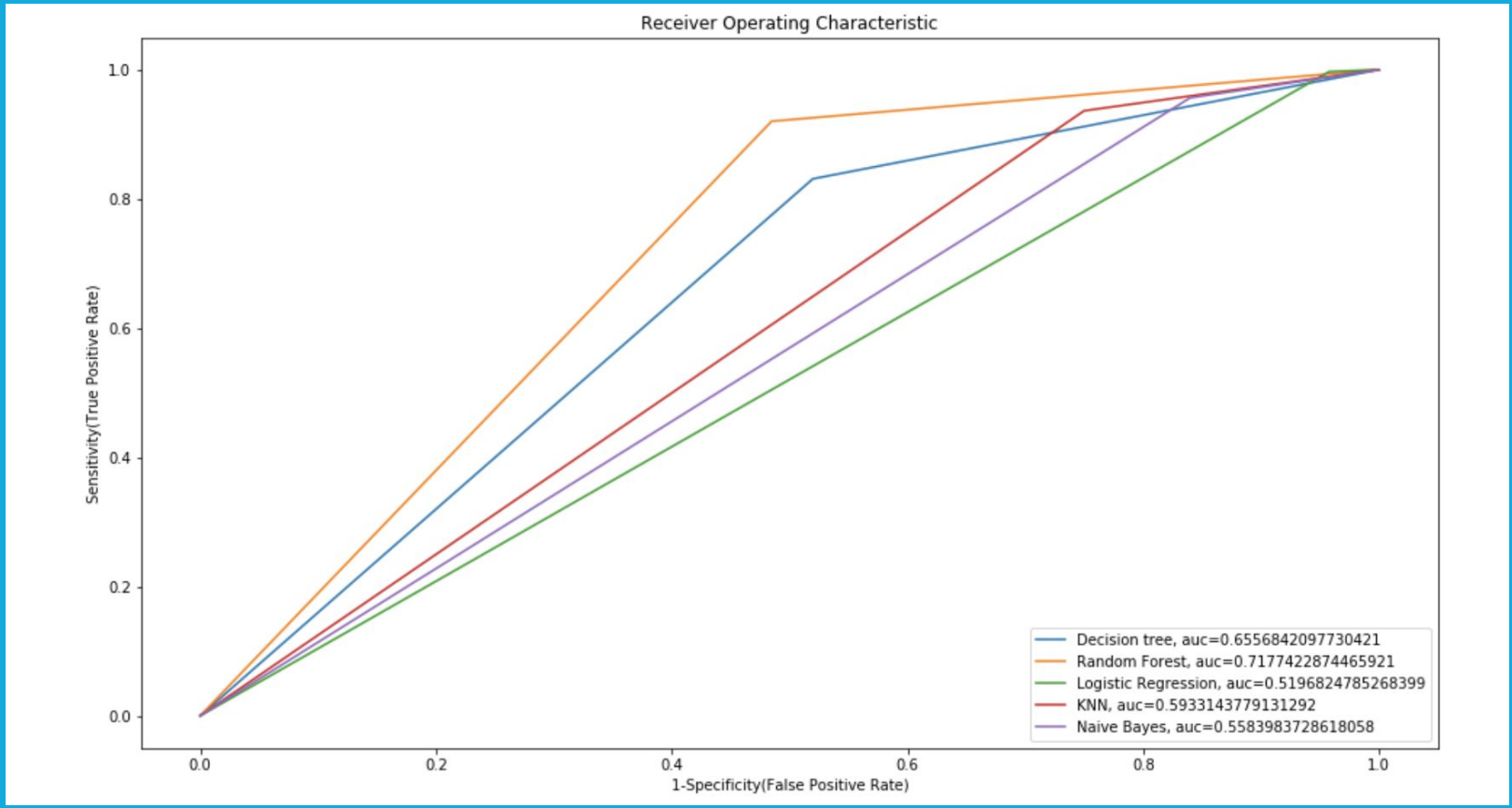


Model Evaluation

Steps for evaluating our classifiers:

1. Partition data-set into: Training and Test sets (80%-20%)
2. Select the classification techniques we want to test
3. Create the models and train them with our Training data
4. Use fitted model to predict the ‘3rd Down’ play on test data
5. Compare model predictions with truth to estimate which model has the best generalization performance.

Model Evaluation



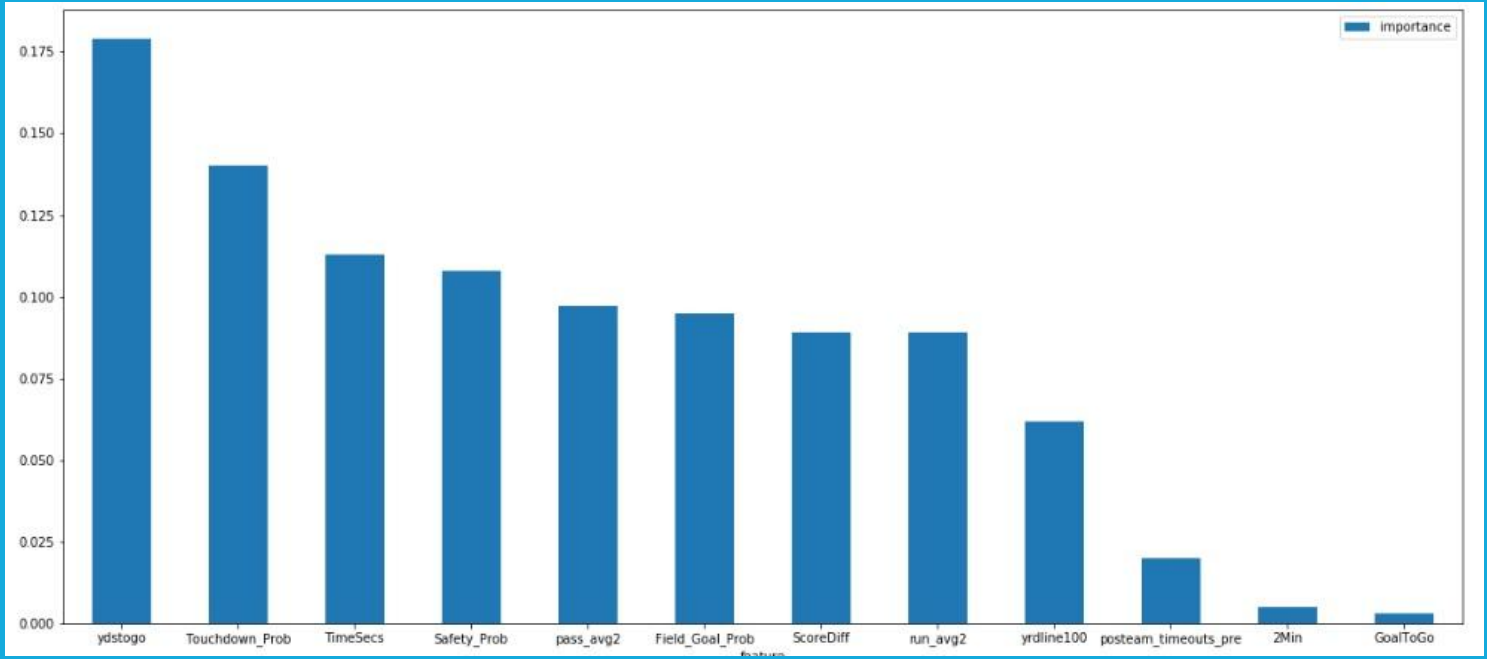
Decision Trees

- Easily interpretable.
- Can handle all sorts of data: binary (2Min, GoalToGo), probabilities (TD, FG, Safety) and numerical (Yrdline100, TimeSecs)
- High variance (sample size sensitive), low bias (can learn the data, if you let it grow long enough)
- NFL decision making is nuanced and scenario dependent: teams with leads run late in games, teams losing tend to pass more, etcetera.
- Perfect candidate for ensemble method. Random Forest !

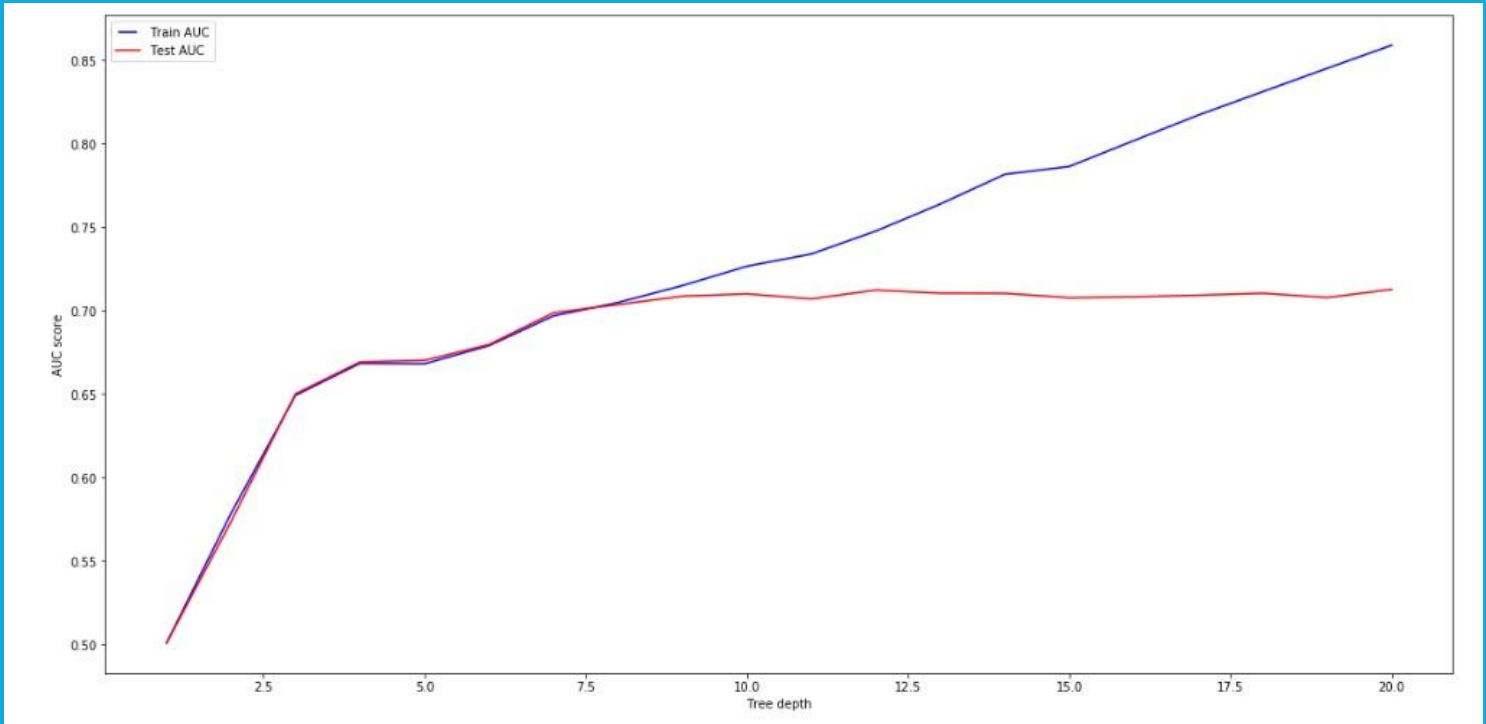
Random Forests

- Algorithm builds multiple decision trees, outputs the majority vote classification.
- Free lunch: decreases variance without increasing bias by bootstrap aggregation and sampling the features randomly at each split.
- Naturally ranks feature importance by computing out-of-sample error before and after permuting a feature.
- Major disadvantage: loses decision tree impretability.

Feature Importance



Tree Depth vs Train/Test AUC



Final Predictions

Accuracy:
84.21%

Precision:
87.83%

Recall:
92.21%

True / Predicted (Raw count and class %)	Predicted run	Predicted pass
Run	1314 (10.79%)	1238 (10.17%)
Pass	684 (5.61%)	8937 (73.41%)

Code repository for the project

https://github.com/smn405/1007-NFL-Project