Predicting Rookie Quarterback Success for Fantasy Football

By Alex Jonauskas

General Model Approach

Model Type - Decision Tree, why?

- Pros of decision tree model
  - Handles nonlinear and categorical data with ease
  - Easily Interpretable
  - Can scale to random forest model with relative ease
- Cons of decision tree model
  - Prone to overfitting
  - Bias towards dominant data

To prevent overfitting I applied stratified k-fold cross validation at varying max depths and min sample sizes as well as different amounts of folds. To help handle the bias towards the dominant class (the misses) I use the highest F1-score when choosing the max depth and or min sample size. I specifically used stratified k-fold since it handles imbalanced data sets better.

Defining Success at the quarterback position

The general approach to define success (hits vs misses) was to consider any player whose,

- average score in any season was above a high threshold at least once,
- or average score was above a middling score at least twice,

to be a hit. The most valuable players are those who have shown they have a high scoring ceiling and/or those who consistently score above average.

In the following slide I show the 5th best and 10th best scoring QBs year after year.


The 5th and 10th Best QB's PPR/G Since 2000
PPG (PPR)
26
24
22
20
18
16
2000
2005
2010
2015
2020
Year

Defining QB Hits and Misses

Since QB scoring has been on an upward trend for the past 20 years we will define a hit as any QB who:

- Scores above the lower trendline in the corresponding year at least twice
- OR scores above the upper trendline in the corresponding year at least once
- AND played at least 12 games in that season

Note: A QB must have been drafted in 2019 or earlier to be considered since they need to play a few seasons before we can determine if they are a hit or bust.

Defining QB Hits and Misses

Here are all 41 hits at the QB position:

'daunte-culpepper-1', 'donovan-mcnabb-1', 'jeff-garcia-1', 'mark-brunell-1', 'peyton-manning-1', 'elvis-grbac-1', 'aaron-brooks-1', 'brett-favre-1', 'steve-mcnair-1', 'tom-brady-1', 'trent-green-1', 'michael-vick-1', 'matt-hasselbeck-1', 'marc-bulger-1', 'drew-brees-1', 'carson-palmer-1', 'tony-romo-1', 'ben-roethlisberger-1', 'aaron-rodgers-1', 'jay-cutler-1', 'philip-rivers-1', 'cam-newton-1', 'eli-manning-1', 'matt-ryan-1', 'matthew-stafford-1', 'andrew-luck-1', 'robert-griffin-iii-1', 'russell-wilson-1', 'andy-dalton-1', 'nick-foles-1', 'ryan-tannehill-1', 'blake-bortles-1', 'kirk-cousins-1', 'dak-prescott-1', 'carson-wentz-1', 'deshaun-watson-1', 'patrick-mahomes-1', 'lamar-jackson-1', 'josh-allen-7', 'justin-herbert-1', 'kyler-murray-1'

Defining QB Hits and Misses

Almost hits (Exactly 1 top ten season and not top 5):

kordell-stewart-1, drew-bledsoe-1, chad-pennington-1, brad-johnson-2, tommy-maddox-1, jake-delhomme-1, jake-plummer-1, derek-anderson-1, david-garrard-1, matt-schaub-1, kyle-orton-1, marcus-mariota-1, ryan-fitzpatrick-1, tyrod-taylor-1, derek-carr-1, alex-smith-3, jared-goff-1, mitch-trubisky-1, jameis-winston-1

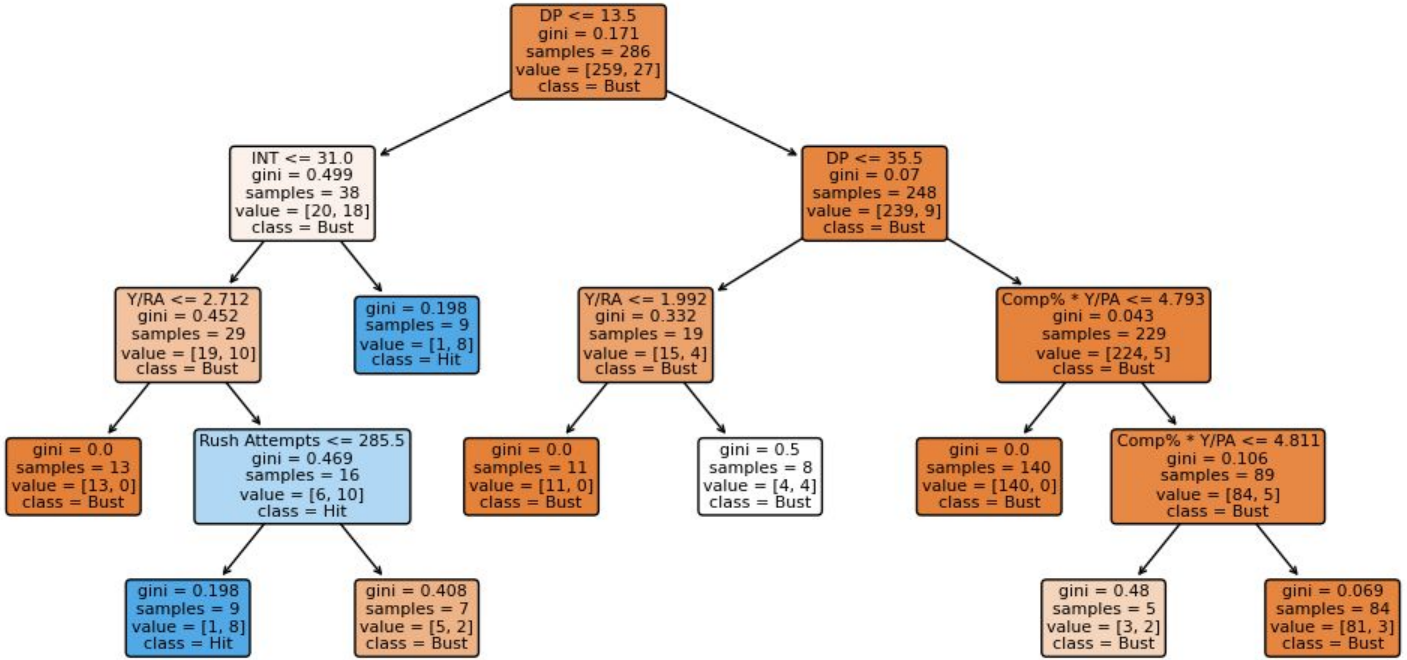
Some of these players have had some value and some good seasons but they are all pretty mediocre. Note: Goff will likely be considered a hit after adding the 2023 season data.

Testing the QB Model

The model was made with the sk-learn library in Python with the following parameters:

- Gini Index as the sorting algorithm (better with biased data than entropy)
- Minimum sample size of 5 for each leaf node
- Maximum depth of 4 for the model

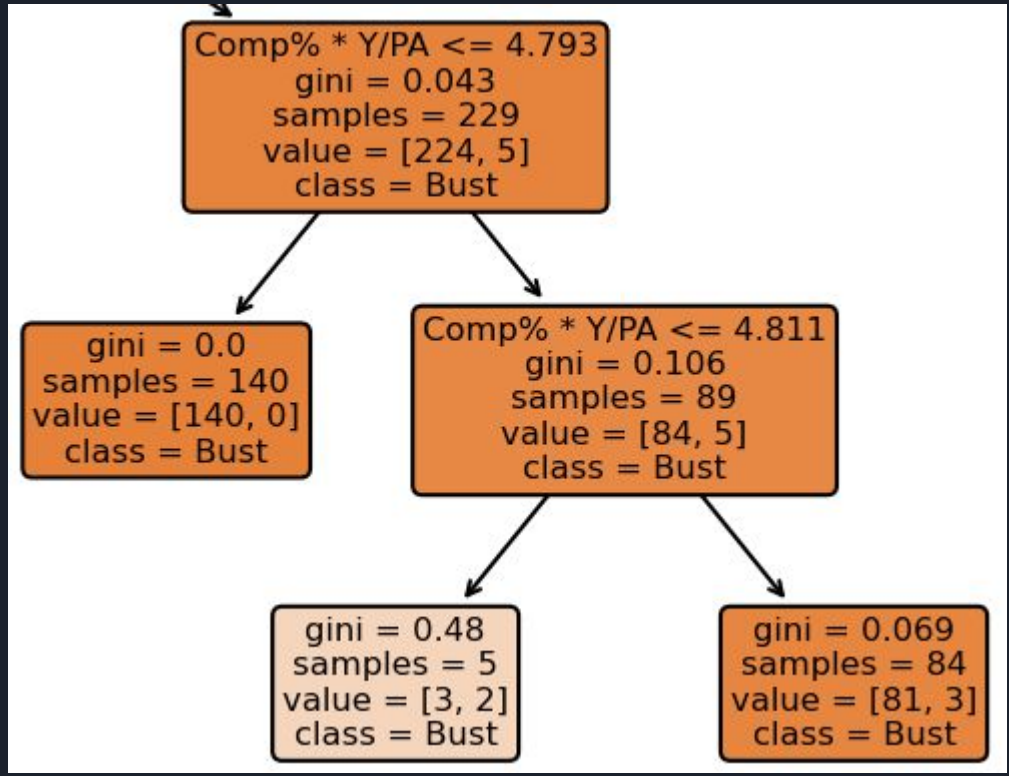
These parameters were decided on by using a stratified k-fold at varying depths, min sample sizes, and number of folds. At a depth of 4, min sample size of 5, and 3 folds I found an average accuracy of 90% with a standard deviation of 0.0274 as well as an average F1 score of 90% with a standard deviation of 0.016.


DP <= 13.5
gini = 0.171
samples = 286
value = [259, 27]
class = Bust
INT <= 31.0
gini = 0.499
samples = 38
value = [20, 18]
class = Bust
DP <= 35.5
gini = 0.07
samples = 248
value = [239, 9]
class = Bust
Y/RA <= 2.712
gini = 0.452
samples = 29
value = [19, 10]
class = Bust
gini = 0.198
samples = 9
value = [1, 8]
class = Hit
Y/RA <= 1.992
gini = 0.332
samples = 19
value = [15, 4]
class = Bust
Comp% * Y/PA <= 4.793
gini = 0.043
samples = 229
value = [224, 5]
class = Bust
gini = 0.0
samples = 13
value = [13, 0]
class = Bust
Rush Attempts <= 285.5
gini = 0.469
samples = 16
value = [6, 10]
class = Hit
gini = 0.0
samples = 11
value = [11, 0]
class = Bust
gini = 0.5
samples = 8
value = [4, 4]
class = Bust
gini = 0.0
samples = 140
value = [140, 0]
class = Bust
Comp% * Y/PA <= 4.811
gini = 0.106
samples = 89
value = [84, 5]
class = Bust
gini = 0.198
samples = 9
value = [1, 8]
class = Hit
gini = 0.408
samples = 7
value = [5, 2]
class = Bust
gini = 0.48
samples = 5
value = [3, 2]
class = Bust
gini = 0.069
samples = 84
value = [81, 3]
class = Bust

Additional Pruning

The last step on the right is likely a result of overfitting since there no reason for a small range between 4.793 and 4.811 for “Comp% * Y/PA” to have a spike in hits for players drafted after pick 35.

After playing with the aforementioned split I found there to be nothing of great note to split this node by and will simply remove the final split.



Model Observations/Analysis

Overall Accuracy: 95.5%

Positive Predictive Value:

59.3% (not including bin that is split 50/50)

74.1% (including bin that is split 50/50)

Negative Predictive Value:

99% (including 50/50 split)

97.7% (not including 50/50 split)

Feature List:

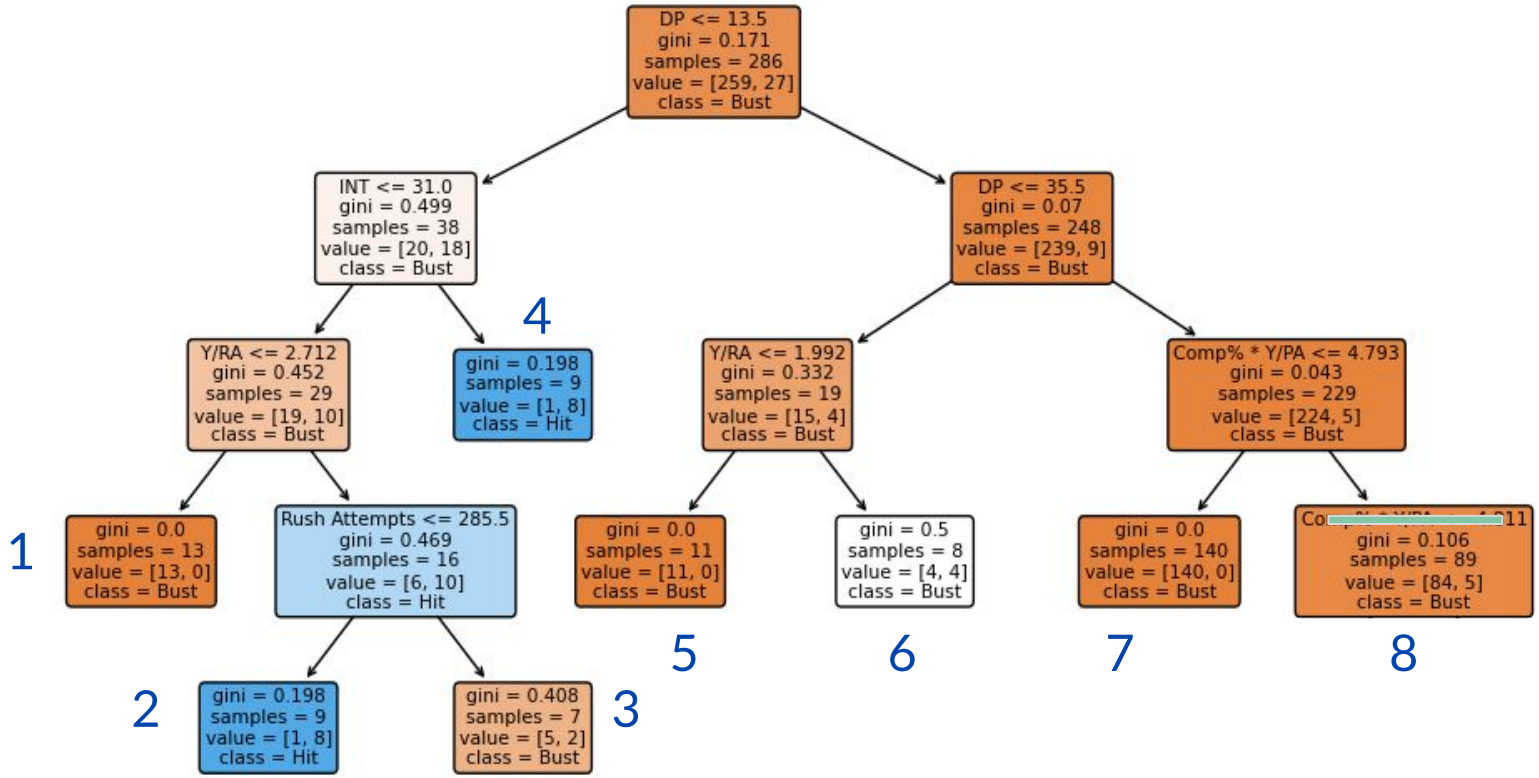
In the model:

- DP: Draft Pick
- INT: Total interceptions
- Y/RA: Yards per rush attempt
- Rush Attempts: Total rush attempts
- Comp% * Y/PA: Completion % times yards per pass attempt

Left out of the model:

- Total: Pass Attempts, Pass Yards, Pass TDs, Rush Yards, Rush TDs
- Comp: total completions
- Comp%: completion percentage
- PPR/G: Fantasy points per game
- Best PPG: PPG in best season
- DR: Draft Round
- Gs: total games played
- Y/PA: Yards per pass attempt
- Y/A: Yards per attempt
- INT %: interception percentage
- Snaps: Total snaps played

Labeling Bins



Observing QB Bins: Bin 1 Training Data

cfb_id	DP	INT	Y/RA	Hit/Miss	cfb_id	DP	INT	Y/RA	Hit/Miss
baker-mayfield-1	1	30	2.68	0	mark-sanchez-1	5	16	0.47	0
sam-darnold-1	3	22	2.42	0	jamarcus-russell-1	1	21	0.57	0
josh-rosen-1	10	26	-1.41	0	matt-leinart-1	10	23	-0.53	0
jared-goff-1	1	30	-0.67	0	byron-leftwich-1	7	26	0.95	0
jameis-winston-1	1	28	1.96	0	david-carr-1	1	18	1.16	0
blaine-gabbert-1	10	18	2.07	0	joey-harrington-1	3	18	1.56	0
sam-bradford-1	1	16	0.47	0					

Interesting Note: The best qbs in this class (Goff, Mayfield, and Winston) were the closest to being above the INT threshold for hits, and Baker was close to the Y/RA threshold as well.

Note:

In the following slides I will present “training” and “test” data. I wanted to note that the training data was used in the model and the test data was qbs drafted in 2020 or later, and therefore left out of the model.

Observing QB Bins: Bin 1 Test Data

cfb_id	DP	INT	Y/RA
cj-stroud-1	2	12	1.7
bryce-young-1	1	10	1.52
justin-herbert-1	6	23	2.42

The only QBs to fall into bin 1 since 2020 are these 3. Unfortunately two of them are hits in a bin of all misses. I do think there are some reasonable explanations for both:

- Herbert has a better Y/RA than all qbs in the miss category except for Baker and will potentially lower that threshold in future iterations of the model.
- CJ Stroud seems to be one of the savviest passers to come out of college football and had such great wrs that he simply rarely felt the need to run. He showed that he could when needed in his final game vs Georgia in the Peach bowl where he averaged 2.8 yards per carry.

Observing QB Bins: Bin 2 Training Data

cfb_id	DP	INT	Y/RA	Rush Att.	Hit/Miss
kyler-murray-1	1	14	7.14	207	1
josh-allen-7	7	21	3.24	237	1
mitch-trubisky-1	2	10	3.66	120	0
carson-wentz-1	2	14	4.66	201	1
blake-bortles-1	3	19	2.91	194	1
andrew-luck-1	1	22	5.87	163	1
ryan-tannehill-1	8	21	3.21	115	1
cam-newton-1	1	7	5.56	285	1
michael-vick-1	1	6	5.93	104	1

Really great bin of QBs. I would consider 5, maybe 6 elite if I count Wentz pre-injury, with another 3 okay qbs in Tannehill, Mitch, and Bortles.

Observing QB Bins: Bin 2 Test Data

cfb_id	DP	INT	Y/RA	Rush Attempts
anthony-richardson-2	4	15	6.93	161
trey-lance-1	3	0	6.51	169
justin-fields-2	11	9	4.36	260
trevor-lawrence-1	1	17	4.08	231
zach-wilson-3	2	15	3.03	212
tua-tagovailoa-1	5	11	3.18	107
joe-burrow-1	1	11	3.18	258

Still a good bin with probably 5/7 hits and 2 guaranteed misses. There is nothing on this table that sticks out about Lance and Wilson but there were questions about both of them during the draft process.

Observing QB Bins: Bin 3 Training Data (No Test Data)

cfb_id	DP	INT	Y/RA	Rush Att.	Hit/Miss
daniel-jones-4	6	29	3.26	406	0
patrick-mahomes-1	10	29	2.74	308	1
marcus-mariota-1	2	14	6.64	337	0
robert-griffin-iii-1	2	17	4.27	528	1
christian-ponder-1	12	30	2.81	296	0
vince-young-1	3	28	6.84	457	0
alex-smith-3	1	8	3.75	286	0

If Alex Smith were a hit the threshold for Rush Att. would be ~322.5 to include Mahomes, Smith, and Ponder in the hit tier. Alex Smith was less than a ppg shy of,

- Having a top 5 season
- Having a 2nd top 10 season

On the other hand, RG3 has the most rush att., suggesting an elite rusher may overcome being a mediocre passer.

Observing QB Bins: Bin 4 Training Data (No Test Data)

cfb_id	DP	INT	Hit/Miss
deshaun-watson-1	12	32	1
jake-locker-1	8	35	0
matthew-stafford-1	1	33	1
matt-ryan-1	3	37	1
jay-cutler-1	11	36	1
ben-roethlisberger-1	11	34	1
eli-manning-1	1	35	1
philip-rivers-1	4	34	1
carson-palmer-1	1	40	1

Another really great bin of hits. I was initially concerned with this bin since it was reached by sorting every qb drafted pick 13 or earlier that also threw over 31 interceptions in college. Interceptions are worth negative fantasy points and do not correlate positively to fantasy production.

However, there are two things to consider:

- A QB must play a lot of games and therefore be good enough to start to reach 31 interceptions in college
- In this bin, these QBs had to be taken pick 13 or higher. NFL teams must be certain a player is good to pick them early despite throwing so many INTs

Observing QB Bins: Bin 5, Training Data

This tier has a lot of okay QBs but none that are/were particularly good in fantasy football.

cfb_id	DP	Y/RA	Hit/Miss
dwayne-haskins-1	15	1.883495	0
teddy-bridgewater-1	32	0.752212	0
brandon-weeden-1	22	-3.84615	0
josh-freeman-1	17	1.602804	0
joe-flacco-1	18	0.496732	0
brady-quinn-1	22	0.716535	0
jason-campbell-1	25	1.232932	0
jp-losman-1	22	1.016878	0
kyle-boller-1	19	-0.29907	0
rex-grossman-1	22	-1.11765	0
patrick-ramsey-1	32	-0.7381	0

Observing QB Bins: Bin 5, Test Data

cfb_id	DP	Y/RA	Hit/Miss
kenny-pickett-1	20	1.920863	0
mac-jones-1	15	0.777778	0

These two fit right in, okay NFL qbs that are worse in fantasy football.

Observing QB Bins: Bin 6 Train & Test Data

Training Data				Test Data			
cfb_id	DP	Y/RA	Hit/Miss	cfb_id	DP	Y/RA	Hit/Miss
lamar-jackson-1	32	6.31	1	will-levis-1	33	2.38	0
paxton-lynch-1	26	2.39	0	jordan-love-2	26	2.37	0
johnny-manziel-1	22	6.29	0				
ej-manuel-1	16	2.78	0				
andy-dalton-1	35	3.90	1				
tim-tebow-1	25	4.26	0				
aaron-rodgers-1	24	2.1	1				
drew-brees-1	32	6	1				

This bin has some of the best, worst, and middling qbs ever drafted in the first round. Levis and Love fit this well because their range of outcomes is large.

Observing QB Bins: Bin 7 Train & Test Data

In this bin there are no QBs that are close to being hits out of the 140 in the training data or the 25 in the test data. The best are career backups/journeyman like Minshew, Brisset, and Josh McCown as well as short term starters Colin Kaepernick and Matt Schaub.

Observing QB Bins: Bin 8 Train & Test Data

Training Data Hits (5.6% Hit Rate)			
cfb_id	DP	Comp% * Y/PA	Hit/Miss
dak-prescott-1	135	5.04	1
russell-wilson-1	75	4.79	1
nick-foles-1	88	4.80	1
kirk-cousins-1	102	5.19	1
tony-romo-1	329	5.50	1

Relevant Test Data Players (4.1% Hit Rate)			
cfb_id	DP	Comp% * Y/PA	Hit/Miss
brock-purdy-1	262	5.62	0
jalen-hurts-1	53	5.90	1
sam-howell-1	144	5.88	0

If I had to guess, I would say Howell does not end up being a hit and Purdy does. That would mean the test data has 4.08% hits vs the training data's 5.62% hits. Note that of these 8 players listed, 6 are closer to pick 36 (highest pick in this bin) than the last pick in the draft.

Summarizing QB Model & Bins

My takeaways from the model:

- Being a high draft pick is most important
- Throwing a lot of interceptions in college and still being picked high is actually a good sign
- The second most important metric for qbs coming out of college seems to be being an efficient rusher