NFL Super Bowl Predictions

By: Carl Bent, Deron Porter, Chin Hsu, Otha Richardson II

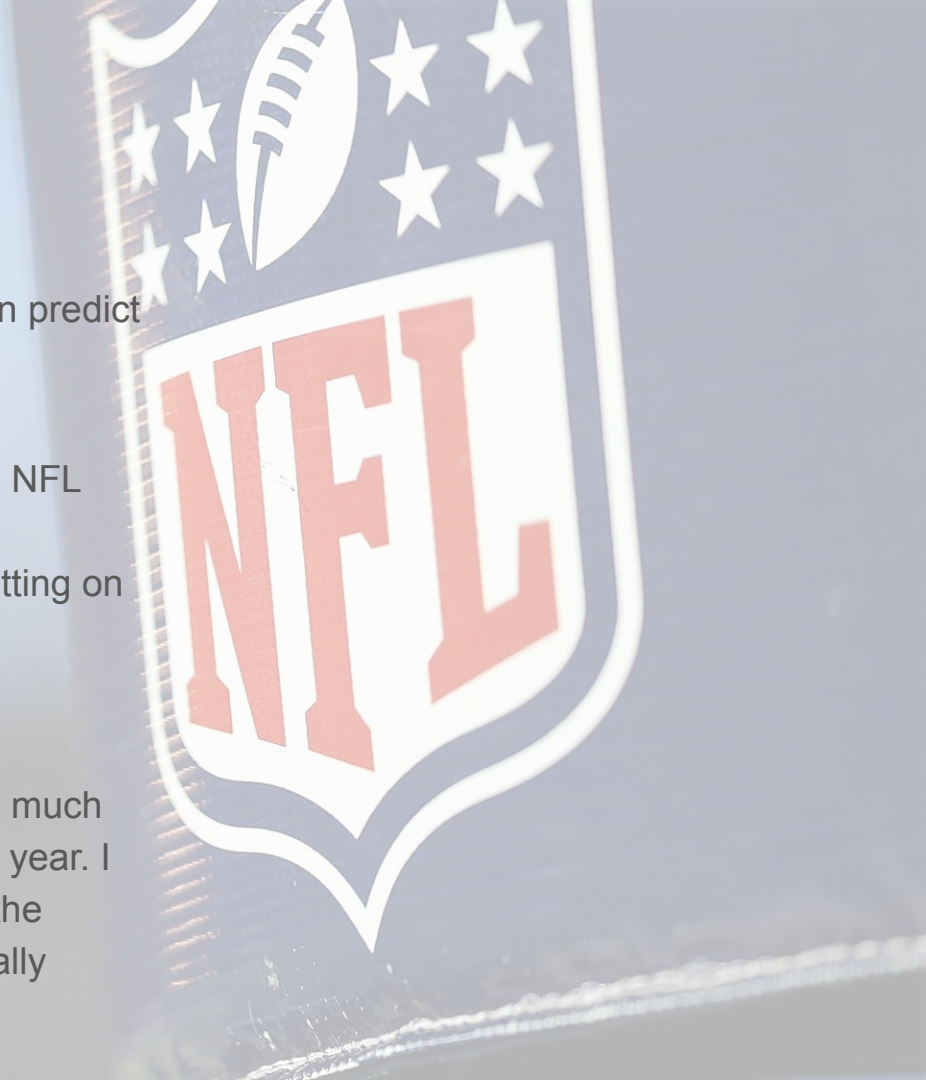
Executive Summary

The football industry is worth well over 15 billion dollars. Financially Technology has interrupted this industry through the utilization of machine learning and statistics. These statistical predictions are mostly used in fantasy football.

We have created a machine learning model that eliminates the time needed to analyze data. You can sit back, relax, and enjoy the game, without worrying about your return on investment.

Concept

- Description:
  - A machine learning based algorithm that can predict the winner of the 2023 NFL Super Bowl.
- Motivation for development?
  - There is an increasing demand for accurate NFL data.
  - To help decrease the risk obtained when betting on the odds of an NFL team winning.
  - Easy to find open source data.
- User story
  - “As a fantasy football investor I have lost so much money betting on the the wrong team every year. I need an algorithm that will help me predict the winner of the 2023 Super Bowl, so I can finally make ROI.“



Data Techniques

- Determined based on these main statistics
  - Number of wins per team
  - Net Points
  - Previous Super Bowl Wins
- Reasoning for data selection
  - Too much data to choose from with limited time. This is our sample prototype.
  - In the future, we will add:
    - Additional statistics for each current player in the NFL
    - The number of yards ran per team
    - Interceptions per team
    - Touchdowns etc…
- Collection, exploration and cleaning process
  - Converted NFL statistics from a CSV file, to a pandas dataframe.

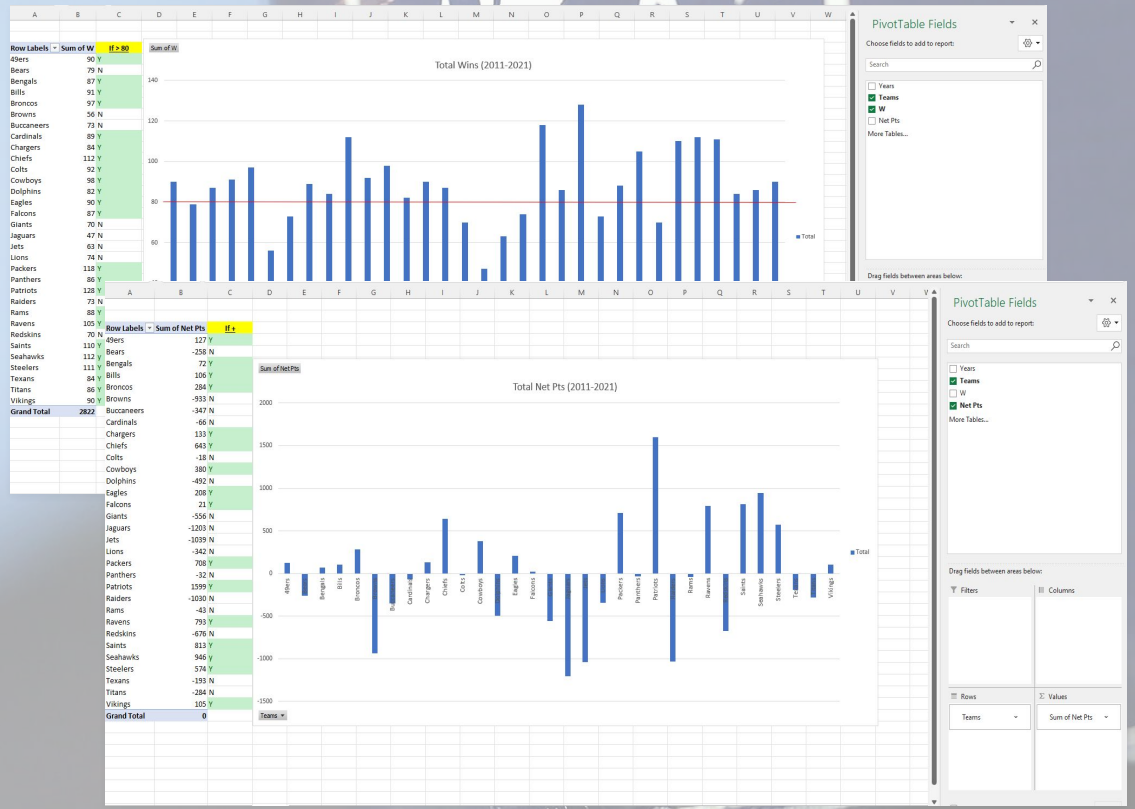
Approach

- Technologies used
  - Languages & Libraries: Python, Pandas, NumPy, Tensorflow, Sklearn, HvPlot, Pydotplus, and IPython
  - Platform: Amazon Web Services (Sagemaker, Lex, and Lambda), Jupyter lab, and Github.
- Breakdown of tasks and roles
  - Otha Richardson II- Data Research, Decision Tree, and logistic regression.
  - Chin Hsu- Data Research, logistic regression, and clustering
  - Deron Porter- Data Research, and deep learning
  - Carl Bent- Data Research, AWS (Lex, and Lambda) and Powerpoint
- Challenges
  - It became difficult to include more data in our model because of timing.
  - Finding a way to display the results in a user friendly way.
- Successes
  - It was really easy to find open source data

Preliminary Analysis - MS Excel (2011-2021)

		Net Pts (+)			Wins (+80)
1	49ers	127		49ers	90
2	Bengals	72		Bengals	87
3	Bills	106		Bills	91
4	Broncos	284		Broncos	97
5	Chargers	133		Chargers	84
6	Chiefs	643		Chiefs	112
7	Cowboys	380		Cowboys	98
8	Eagles	208		Eagles	90
9	Falcons	21		Falcons	87
10	Packers	708		Packers	118
11	Patriots	1599		Patriots	128
12	Ravens	793		Ravens	105
13	Saints	813		Saints	110
14	Seahawks	946		Seahawks	112
15	Steelers	574		Steelers	111
16	Vikings	105		Vikings	90
17				Cardinals	89
18				Colts	92
19				Dolphins	82
20				Panthers	86
21				Texans	84
22				Titans	86
23				Rams	88
24	Buccaneers	-347		Buccaneers	73
25	Giants	-556		Giants	70

'Net Points' is the biggest predictor for wins.



https://github.com/kingdee60/Project_2_SuperBowlPrediction/blob/main/Resources/Super%20Bowl%20Analysis.xlsx

Preliminary Analysis - Current Data & SuperBowl History

		Net Pts	Wins	Super Bowl Titles
1	Patriots	1599	128	2015 Super Bowl Champions 2017 Super Bowl Champions 2019 Super Bowl Champions
2	Packers	708	118	2011 Super Bowl Champions
3	Chiefs	643	112	2020 Super Bowl Champions
4	Seahawks	946	112	2014 Super Bowl Champions
5	Steelers	574	111	
6	Saints	813	110	
7	Ravens	793	105	2013 Super Bowl Champions
8	Cowboys	380	98	
9	Broncos	284	97	2016 Super Bowl Champions
10	Colts		92	
11	Bills	106	91	
12	49ers	127	90	
13	Eagles	208	90	2018 Super Bowl Champions
14	Vikings	105	90	
15	Cardinals		89	
16	Rams		88	2022 Super Bowl Champions
17	Bengals	72	87	
18	Falcons	21	87	
19	Panthers		86	
20	Titans		86	
21	Chargers	133	84	
22	Texans		84	
23	Dolphins		82	
24	Buccaneers	-347	73	2021 Super Bowl Champions
25	Giants	-556	70	2012 Super Bowl Champions

AFC Standings * - division winner, + - wild card Share & Export ▾ Glossary

Tm	W	L	T	W-L%	PF	PA	PD	MoV	SoS	SRS	OSRS	DSRS
AFC East												
Buffalo Bills	3	1	0	.750	114	58	56	14.0	4.1	18.1	5.2	12.9
Miami Dolphins	3	1	0	.750	98	91	7	1.8	9.1	10.9	7.1	3.8
New York Jets	2	2	0	.500	76	101	-25	-6.3	3.7	-2.6	-3.5	0.9
New England Patriots	1	3	0	.250	74	98	-24	-6.0	5.5	-0.5	-2.8	2.3
AFC North												
Baltimore Ravens	2	2	0	.500	119	100	19	4.8	6.5	11.2	12.7	-1.5
Cincinnati Bengals	2	2	0	.500	91	70	21	5.3	2.4	7.6	2.9	4.7
Cleveland Browns	2	2	0	.500	105	95	10	2.5	-3.6	-1.1	3.8	-4.9
Pittsburgh Steelers	1	3	0	.250	74	90	-16	-4.0	0.9	-3.1	-2.7	-0.4
AFC South												
Jacksonville Jaguars	2	2	0	.500	105	67	38	9.5	0.8	10.3	9.1	1.1
Tennessee Titans	2	2	0	.500	75	101	-26	-6.5	-0.1	-6.6	-0.5	-6.1
Indianapolis Colts	1	2	1	.375	57	85	-28	-7.0	-0.9	-7.9	-14.0	6.1
Houston Texans	0	3	1	.125	73	93	-20	-5.0	-7.4	-12.4	-1.3	-11.1
AFC West												
Kansas City Chiefs	3	1	0	.750	129	96	33	8.3	-3.1	5.1	14.2	-9.1
Denver Broncos	2	2	0	.500	66	68	-2	-0.5	-8.4	-8.9	-11.7	2.8
Los Angeles Chargers	2	2	0	.500	92	108	-16	-4.0	-1.2	-5.2	-6.0	0.8
Las Vegas Raiders	1	3	0	.250	96	100	-4	-1.0	-6.6	-7.6	1.2	-8.8

NFC Standings * - division winner, + - wild card Share & Export ▾ Glossary

Tm	W	L	T	W-L%	PF	PA	PD	MoV	SoS	SRS	OSRS	DSRS
NFC East												
Philadelphia Eagles	4	0	0	1.000	115	71	44	11.0	4.4	15.4	6.5	8.9
Dallas Cowboys	3	1	0	.750	71	62	9	2.3	2.1	4.4	0.1	4.3
New York Giants	3	1	0	.750	76	71	5	1.3	-4.2	-2.9	-3.7	0.8
Washington Commanders	1	3	0	.250	73	107	-34	-8.5	8.1	-0.4	-3.5	3.2
NFC North												
Minnesota Vikings	3	1	0	.750	86	80	6	1.5	3.9	5.4	-2.9	8.2
Green Bay Packers	3	1	0	.750	75	69	6	1.5	0.6	2.1	1.7	0.4
Chicago Bears	2	2	0	.500	64	77	-13	-3.3	-4.5	-7.8	-7.7	-0.1
Detroit Lions	1	3	0	.250	140	141	-1	-0.3	2.7	2.4	15.9	-13.5
NFC South												
Tampa Bay Buccaneers	2	2	0	.500	82	68	14	3.5	1.8	5.3	-3.9	9.2
Atlanta Falcons	2	2	0	.500	103	101	2	0.5	-3.6	-3.1	-1.3	-1.8
New Orleans Saints	1	3	0	.250	76	96	-20	-5.0	0.5	-4.5	0.8	-5.3
Carolina Panthers	1	3	0	.250	78	85	-7	-1.8	-3.8	-5.6	-5.2	-0.4
NFC West												
Los Angeles Rams	2	1	0	.667	61	70	-9	-3.0	2.9	-0.1	1.7	-1.8
Arizona Cardinals	2	2	0	.500	88	103	-15	-3.8	-2.0	-5.8	-5.0	-0.8
Seattle Seahawks	2	2	0	.500	95	115	-20	-5.0	-3.6	-8.6	-0.6	-8.0
San Francisco 49ers	1	2	0	.333	47	37	10	3.3	-8.3	-4.9	-8.0	3.1

** Added SuperBowl History (0 or 1), current standings considerations, and qualitative information (ie Tom Brady transfer to Buccaneers…Russell Wilson to Broncos…Rams 2022 SuperBowl Winners)

Data - Logistic Regression

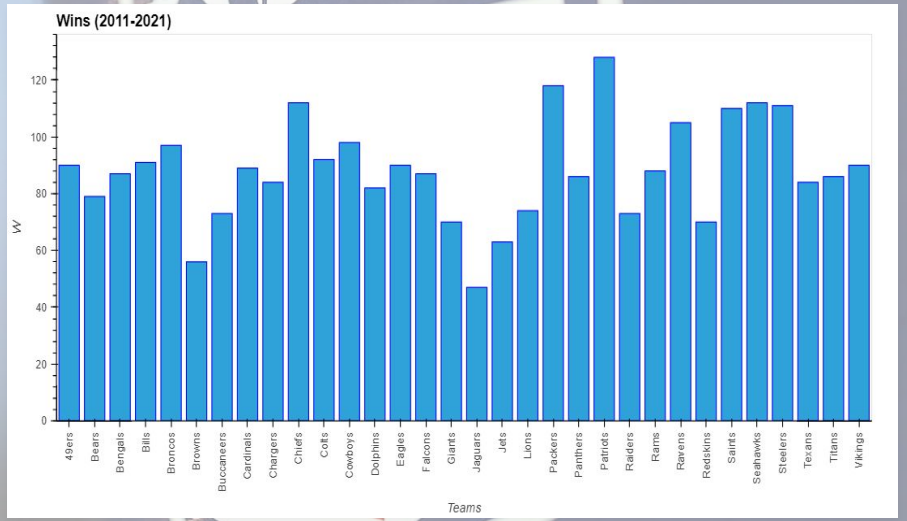
In [51]:
```
# Print the classification report for the model
from sklearn.metrics import classification_report

target_names = ["0", "1"]
print(classification_report(y_resampled, resampled_predictions, target_names=target_names))
```

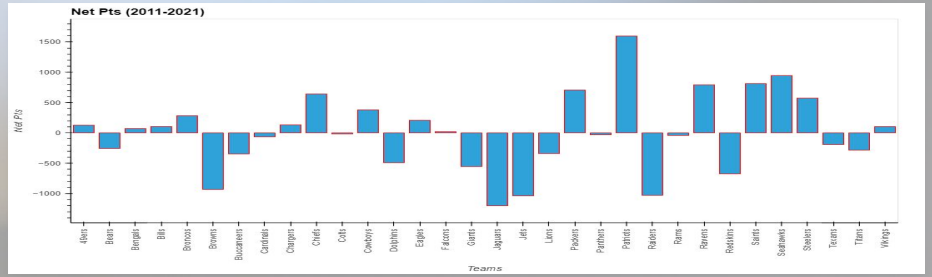
```
              precision    recall  f1-score   support

           0       0.81      0.76      0.79        17
           1       0.78      0.82      0.80        17

    accuracy                           0.79        34
   macro avg       0.80      0.79      0.79        34
weighted avg       0.80      0.79      0.79        34
```

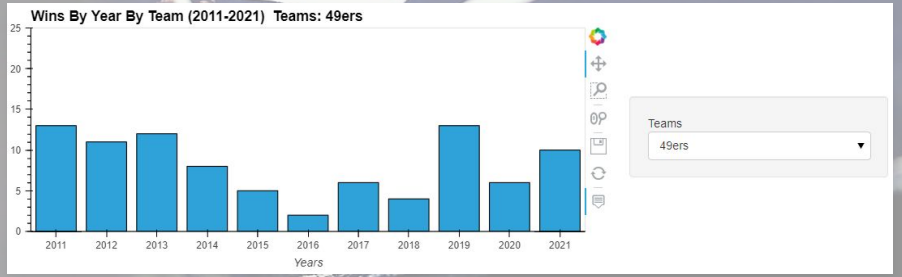
Step 4: Answer the following question

Question: How well does the logistic regression model, fit with oversampled data, predict both the `0` (losers) and `1` (winners) labels? 79%

Answer: The logistic regression model predicts with 78% precision for 'winners' and 81% precision for 'losers'. Overall accuracy is ~79% for categorized winners/losers in the test data.



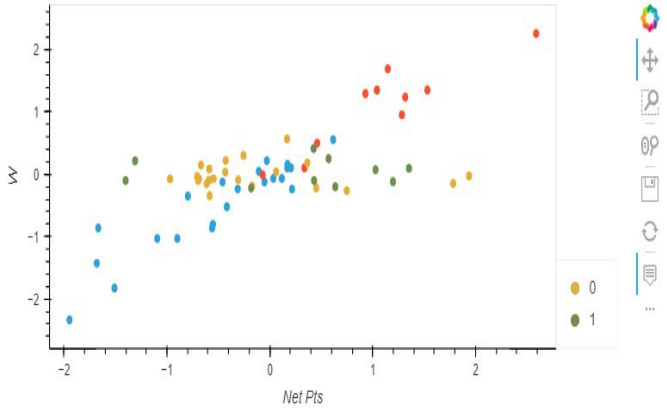


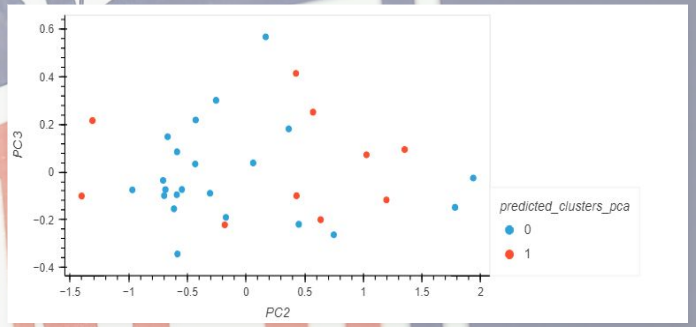


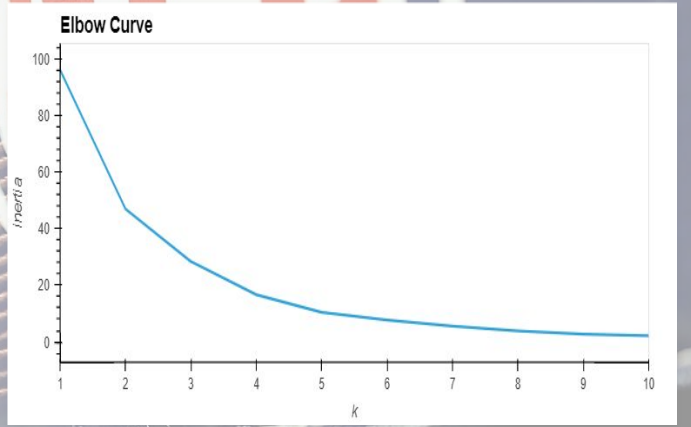
Data Continued- Clustering

```
In [39]: # Compoosite plot to contrast the clusters
composite_SuperBowlWinners3_predictions_plot = df_SuperBowl_predictions_plot * SuperBowlWinners3_pca_predictions_plot
composite_SuperBowlWinners3_predictions_plot
```

Out[39]:







Answer the following question: After visually analyzing the cluster analysis results, what is the impact of using fewer features to cluster the data using K-Means?

- **Question:** Which teams stand out as a result of using fewer features to cluster the data using K-Means? Team 21-Patriots
- **Answer:** Much of the information (or the maximum variance) of the original dataset is still retained. The outliers are more apparent. Team 21-Patriots is the most obvious. These others follow: Team 27-Seahawks, Team 26-Saints, Team 24-Ravens, Team 19-Packers, Team 9-Chiefs, Team 28-Steelers

Data Continued- Deep Learning

In [44]:
```
print("Original Model Results")

# Evaluate the model loss and accuracy metrics using the evaluate method and the test data
model_loss, model_accuracy = nn.evaluate(X_test_scaled,y_test,verbose=2)

# Display the model loss and accuracy results
print(f"Loss: {model_loss}, Accuracy: {model_accuracy}")
```

```
Original Model Results
1/1 - 0s - loss: 0.6347 - accuracy: 0.7500 - 17ms/epoch - 17ms/step
Loss: 0.6346674561500549, Accuracy: 0.75
```

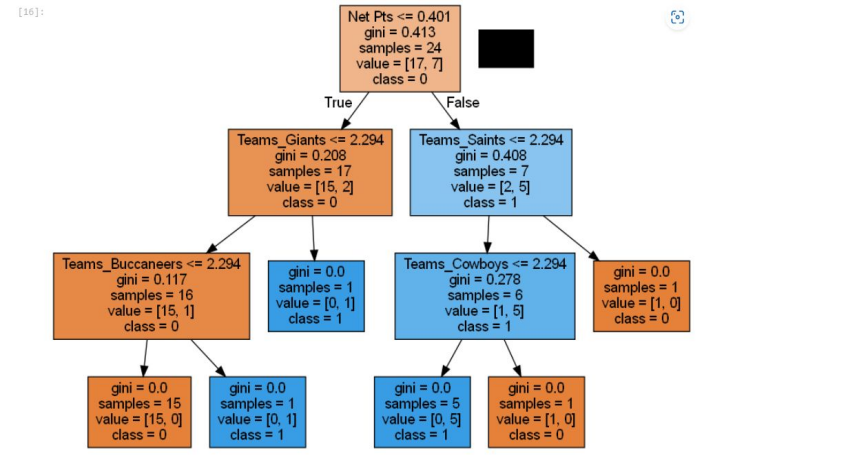
In [45]:
```
print("Alternative Model 1 Results")

# Evaluate the model loss and accuracy metrics using the evaluate method and the test data
model_loss, model_accuracy = nn_A1.evaluate(X_test_scaled,y_test,verbose=2)

# Display the model loss and accuracy results
print(f"Loss: {model_loss}, Accuracy: {model_accuracy}")
```

```
Alternative Model 1 Results
1/1 - 0s - loss: 0.5653 - accuracy: 0.8750 - 98ms/epoch - 98ms/step
Loss: 0.5653348565101624, Accuracy: 0.875
```

In [46]:
```
print("Alternative Model 2 Results")


# Evaluate the model loss and accuracy metrics using the evaluate method and the test data
model_loss, model_accuracy = nn_A2.evaluate(X_test_scaled,y_test,verbose=2)

# Display the model loss and accuracy results
print(f"Loss: {model_loss}, Accuracy: {model_accuracy}")
```

```
Alternative Model 2 Results
1/1 - 0s - loss: 0.6547 - accuracy: 0.3750 - 96ms/epoch - 96ms/step
Loss: 0.654725193977356, Accuracy: 0.375
```

Preliminary Analysis - Decision Tree



** Removed New England Patriots as the outlier. They are the ONLY NFL Team with multiple wins in the last decade.

1	Packers	708	118	2011 Super Bowl Champions
2	Chiefs	643	112	2020 Super Bowl Champions
3	Broncos	284	97	2016 Super Bowl Champions
4	Eagles	208	90	2018 Super Bowl Champions
5	Rams	-43	88	2022 Super Bowl Champions
6	Buccaneers	-347	73	2021 Super Bowl Champions
7	Cowboys	380	98	
8	Bills	106	91	
9	Vikings	105	90	
10	Bengals	72	87	

		Net Pts	Wins	Super Bowl Championships
1	Patriots	1599	128	2015 Super Bowl Champions 2017 Super Bowl Champions 2019 Super Bowl Champions
2	Packers	708	118	2011 Super Bowl Champions
3	Chiefs	643	112	2020 Super Bowl Champions
4	Ravens	793	105	2013 Super Bowl Champions
5	Cowboys	380	98	
6	Broncos	284	97	2016 Super Bowl Champions
7	Bills	106	91	
8	Eagles	208	90	2018 Super Bowl Champions
9	Vikings	105	90	
10	Rams	-43	88	2022 Super Bowl Champions
11	Falcons	21	87	
12	Bengals	72	87	
13	Chargers	133	84	
14	Dolphins	-492	82	
15	Buccaneers	-347	73	2021 Super Bowl Champions
16	Giants	-556	70	2012 Super Bowl Champions

Patriots are the only team to win in multiple seasons in last 10 years

		Net Pts	Wins	Super Bowl Championships
1	Patriots	1599	128	2015 Super Bowl Champions 2017 Super Bowl Champions 2019 Super Bowl Champions

Repeat Champions

2	Packers	708	118	2011 Super Bowl Champions
3	Chiefs	643	112	2020 Super Bowl Champions
4	Ravens	793	105	2013 Super Bowl Champions
5	Broncos	284	97	2016 Super Bowl Champions
6	Eagles	208	90	2018 Super Bowl Champions
7	Rams	-43	88	2022 Super Bowl Champions
8	Buccaneers	-347	73	2021 Super Bowl Champions
9	Giants	-556	70	2012 Super Bowl Champions

One-Time Champions

10	Cowboys	380	98	
11	Bills	106	91	
12	Vikings	105	90	
13	Falcons	21	87	
14	Bengals	72	87	
15	Chargers	133	84	
16	Dolphins	-492	82	

Non- Champions

Based on 2021 Record

		Net Pts	Wins	Super Bowl Championships
1	Patriots	1599	128	2015 Super Bowl Champions 2017 Super Bowl Champions 2019 Super Bowl Champions
2	Packers	708	118	2011 Super Bowl Champions
3	Chiefs	643	112	2020 Super Bowl Champions
4	Broncos	284	97	2016 Super Bowl Champions
5	Eagles	208	90	2018 Super Bowl Champions
6	Rams	-43	88	2022 Super Bowl Champions
7	Buccaneers	-347	73	2021 Super Bowl Champions

8	Cowboys	380	98	
9	Bills	106	91	
10	Vikings	105	90	
11	Bengals	72	87	
12	Titans	-284	86	
13	Dolphins	-492	82	

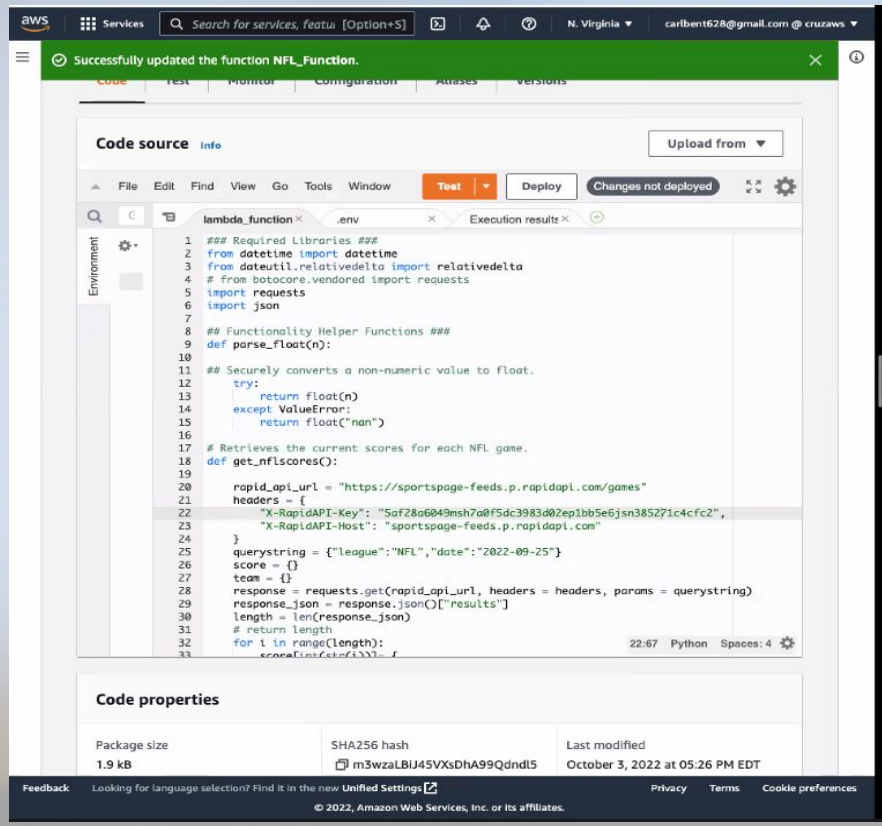
Based on 2022 Week 3

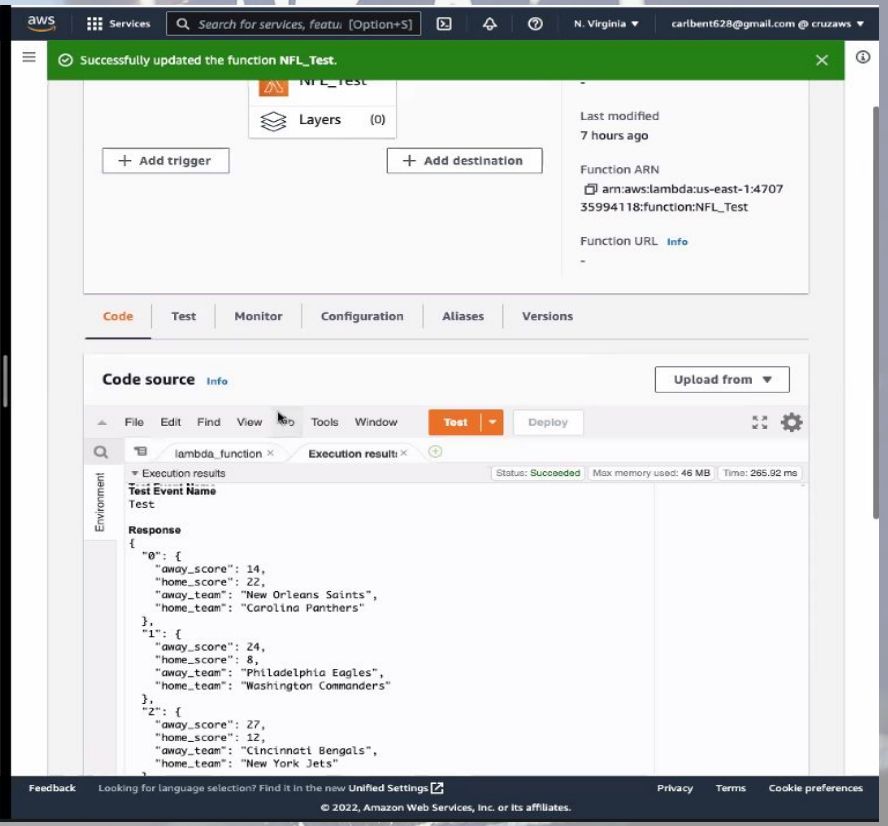
		Net Pts	Wins	Super Bowl Championships
1	Packers	708	118	2011 Super Bowl Champions
2	Chiefs	643	112	2020 Super Bowl Champions
3	Broncos	284	97	2016 Super Bowl Champions
4	Eagles	208	90	2018 Super Bowl Champions
5	Rams	-43	88	2022 Super Bowl Champions
6	Buccaneers	-347	73	2021 Super Bowl Champions

7	Cowboys	380	98	
8	Bills	106	91	
9	Vikings	105	90	
10	Bengals	72	87	

SuperBowl Predictions

NFL Weekly Scoring API (Sports Trader)





Project #2 Demo

Next Steps

- Additional questions that surfaced.
  - What will predictive analytics look like in 5-10 years from now?
  - Will the government add any limitations to this in the future?
- Additional topics to research.
  - Creating additional algorithms to add to our model using:
    - Receiving, passing, touchdowns, tackles, fumbles, interceptions, and more
    - Statistics for each current player in the NFL
- Plan for future development.
  - Based on the 2023 Super Bowl results we plan to:
    - If we receive **inaccurate** results
      - **Fix** our model to get better predictions
    - If we receive **accurate** results
      - **Expand** and go more in depth on our model

Links

- Deployed application.
  - Github Repository
- Interactive Poll
  - https://take.supersurvey.com/poll4503144xC17742Bc-141

Works Cited

- National Football League- (NFL)
  - https://www.nfl.com/stats/team-stats/
  - https://www.nfl.com/standings/
  - https://www.pro-football-reference.com/