NBA Winning Prediction Model

Presented By: Nicholas Kissoon - 100742790

Motivation

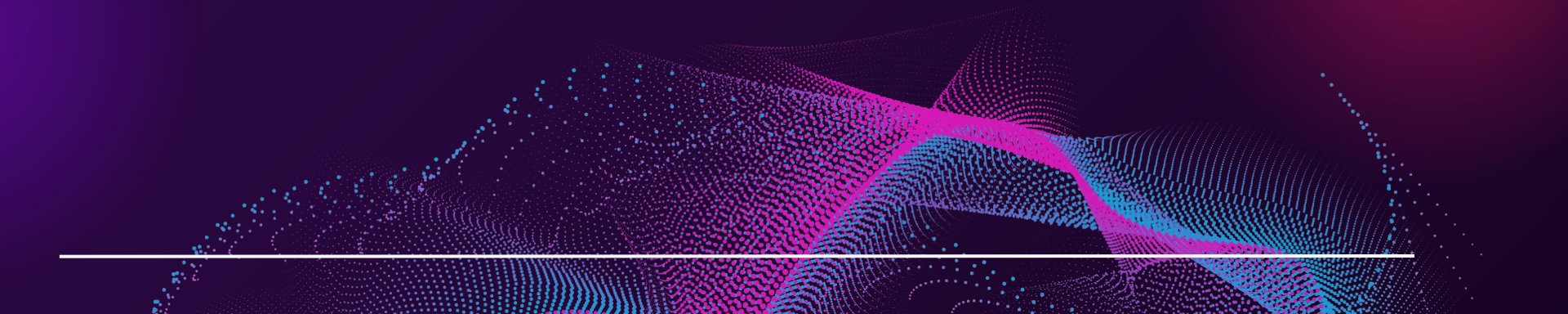
- Inspired by Professor Pu’s famous words, paraphrased of course, how can one make something profitable, or at the very least good for your resume.
- I watch the NBA on a pretty consistent basis, being aware of injuries, latest trade news and of course the box score
- Want to have a personal connection to this project
  - Passion
- While also creating a useful tool that I could see myself using for years to come

Introduction

- National Basketball Association (NBA) is the biggest professional basketball league in the world
- Consists of 30 teams: 29 in the United States and 1 in Canada
- Regular season runs from October to April, with each team playing 82 games
- As of 2020, NBA players are the world’s highest-paid athletes by average annual salary
- A lot of money comes is involved in the NBA and who says we can’t have a piece of that pie.
- The goal of this NBA ML Model is to be able to predict winning teams of games throughout the season.
- This in turn can be fully developed and fine tuned to then provide a service to help betters make informed betting decisions
- All of this can be achieved with the use of the nba_api in python

1 Classifier Models

Multi-layer Perceptron (MLP)



Why a Multi-layer perceptron?

This was the eventually selected approach due to the importance of feature selection and how it handles non-linear relationships within the data. As mentioned in my proposal feature selection plays a very large role, however the ability to handle complex, nonlinear is key here as that helps with variables such as player performance, team dynamics and historical data/statistics. It was also selected due to it being easy to implement and to get preliminary results to see if the data is useful.

Results

Best parameters: {'activation': 'tanh', 'alpha': 0.0001, 'hidden_layer_sizes': (100, 50), 'learning_rate': 'constant'}
Accuracy: 0.82

	precision	recall	f1-score	support
0	0.78	0.78	0.78	59
1	0.84	0.84	0.84	83
accuracy			0.82	142
macro avg	0.81	0.81	0.81	142
weighted avg	0.82	0.82	0.82	142

Feature Importance:

	feature	importance
2	FG_PCT	0.073944
1	FGA	0.072535
15	TOV	0.071127
10	DREB	0.059155
11	REB	0.037324
5	FG3_PCT	0.034507
13	STL	0.032394
16	PF	0.022535
17	PTS	0.021127
3	FG3M	0.019718
12	AST	0.016901
8	FT_PCT	0.014085
9	OREB	0.012676
4	FG3A	0.011268
6	FTM	0.001408
0	FGM	0.001408
7	FTA	-0.002113
14	BLK	-0.008451

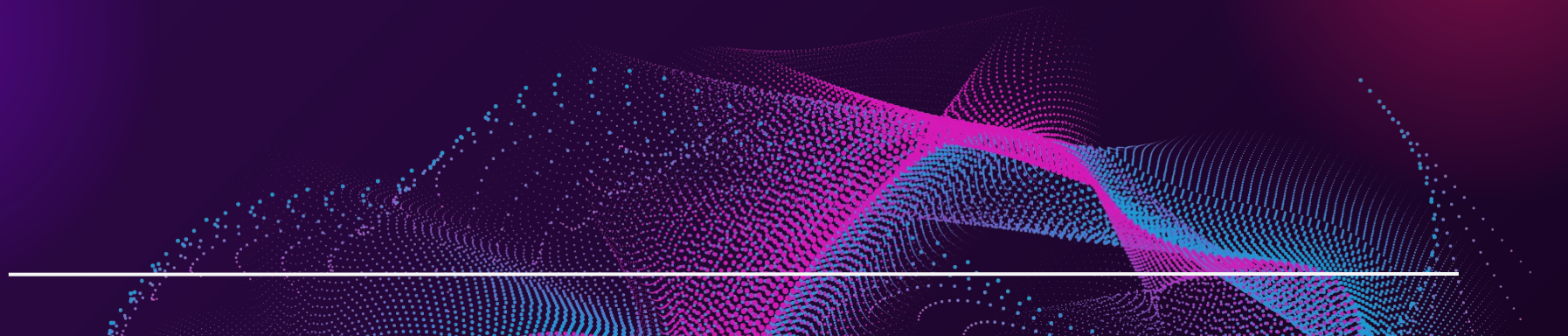
Predictions for new games (2023-24 season):

	GAME_DATE	MATCHUP	Predicted_Outcome	Win_Probability
0	APR 14, 2024	GSW vs. UTA	W	0.991287
1	APR 12, 2024	GSW vs. NOP	W	0.999994
2	APR 11, 2024	GSW @ POR	W	0.999990
3	APR 09, 2024	GSW @ LAL	W	0.999954
4	APR 07, 2024	GSW vs. UTA	W	1.000000
..	...	...	...	...
77	NOV 01, 2023	GSW vs. SAC	L	0.000098
78	OCT 30, 2023	GSW @ NOP	W	1.000000
79	OCT 29, 2023	GSW @ HOU	W	0.999983
80	OCT 27, 2023	GSW @ SAC	W	0.999956
81	OCT 24, 2023	GSW vs. PHX	L	0.000022

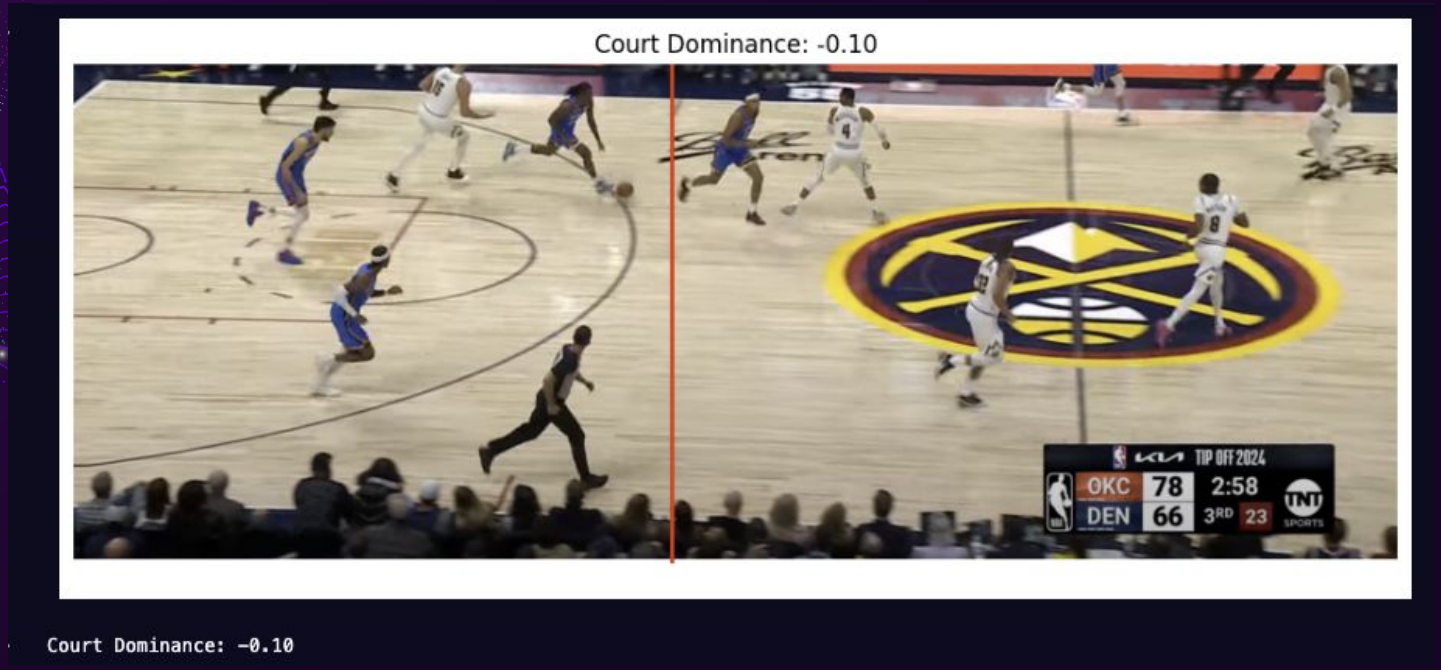
[82 rows x 4 columns]

2 Computer Vision (CV)

OpenCV (CV2)



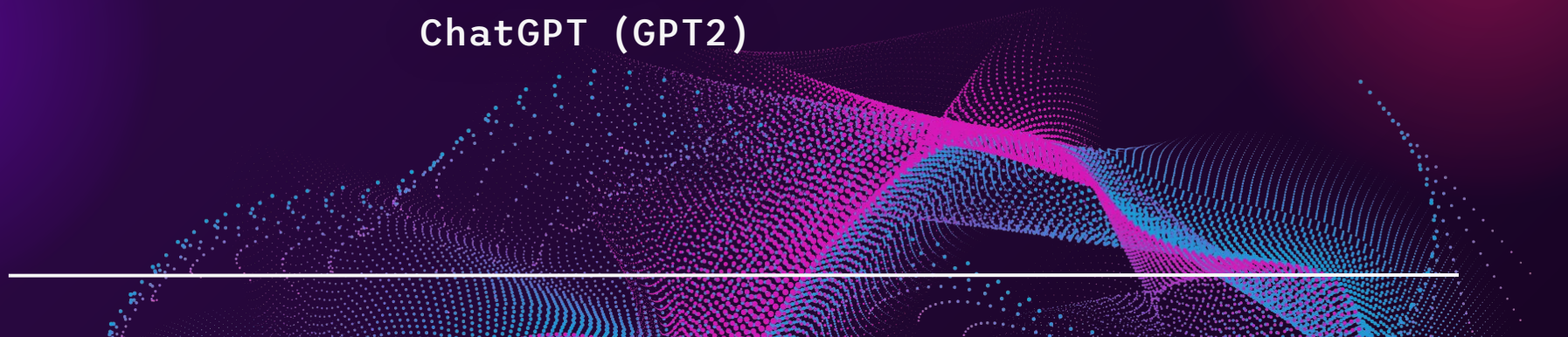
Results


Court Dominance: -0.10
TIP OFF 2024
OKC 78
DEN 66
2:58
3RD 23
TNT SPORTS
Court Dominance: -0.10

3

Natural Language Processing (NLP)

ChatGPT (GPT2)



Results

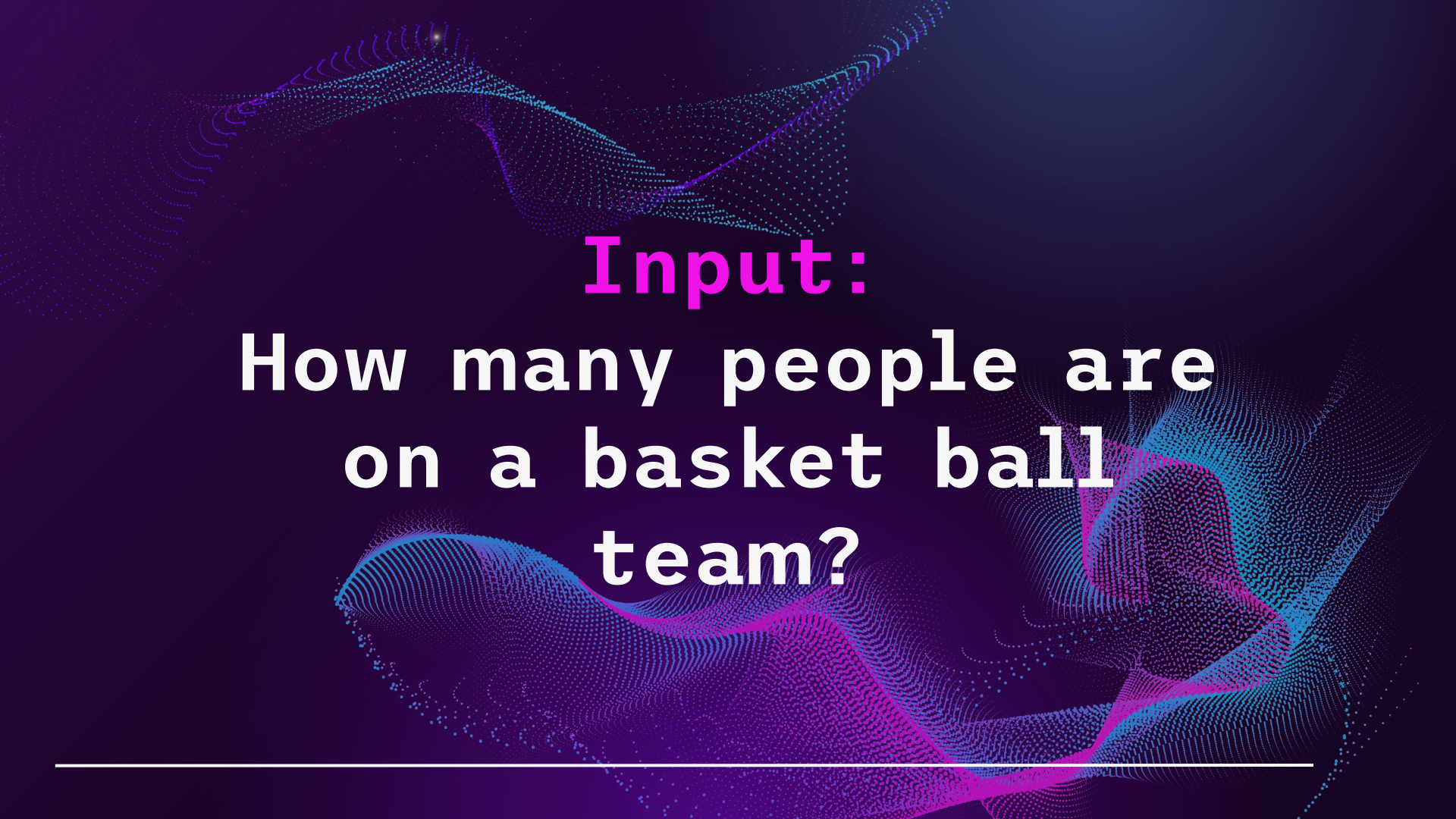
Input

How many people are on a basket ball team?

Output

It's hard to say. The average player is averaging around 20 points per game. So if you look at the NBA, it's very different. You're going to have a bunch of guys that are playing a lot more minutes, but it's a different league.

So when you look at the average NBA team, it's not that you're going to have a lot of players who are on a basket ball team, but


Input:
How many people are on a basket ball team?

Output:

It's hard to say. The average player is averaging around 20 points per game. So if you look at the NBA, it's very different. You're going to have a bunch of guys that are playing a lot more minutes, but it's a different league.

So when you look at the average NBA team, it's not that you're going to have a lot of players who are on a basket ball team, but

Conclusion

- We examined how this model can return viable information for bettors and sports analysts alike through the use of machine learning classification (Multi-layer Perceptron [MLP]), computer vision (CV) and natural language processing (NLP)
- Our classifier achieved an accuracy rate of 0.82, ranking the important features and even predicted scores for a new season.
- We were able to calculate the court dominance which can help provide insight to those betting as well as the visually challenged so long as this prompt is said allowed to them
- While doing its best to stay relevant and on topic we see here how GPT2 can answer some of the users questions.
- I thank you for taking the time to explore this project, feel free to clone the Github repository (https://github.com/Kxssoon/CSCI4052Project) and play around with it yourself to see how you can make it further beneficial for you.

Stay Ballin.