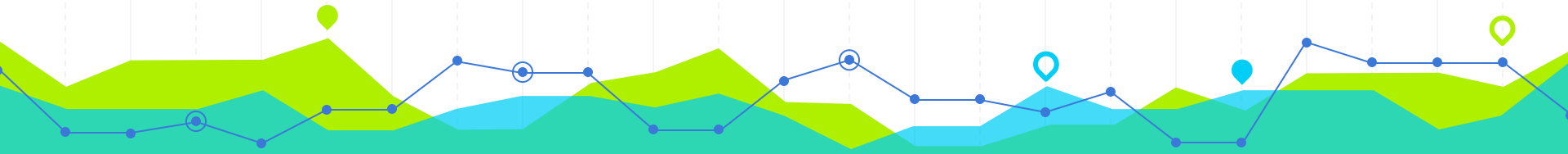
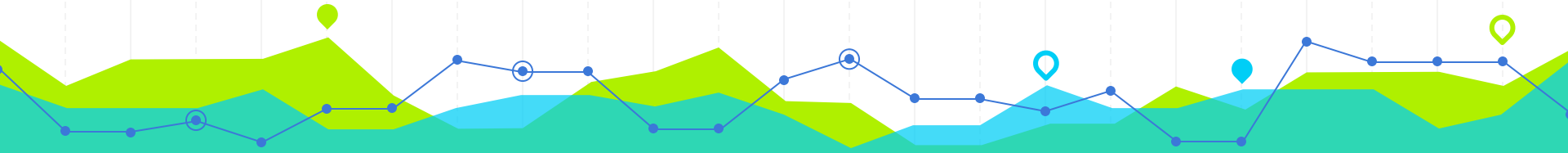


FIFA Position Recommendation

Akash Bhatnagar, Chung-Yeh Yang, Kimberly Low, Micah Lee

AGENDA

1. Introduction: How do we come up with our data?
2. Data Description: What does our original data look like?
3. Data Preprocessing: How do we clean the data?
4. Our Model: How do we run the prediction model through the data?
5. Implications: What are our findings and how can the model be used in the real world?

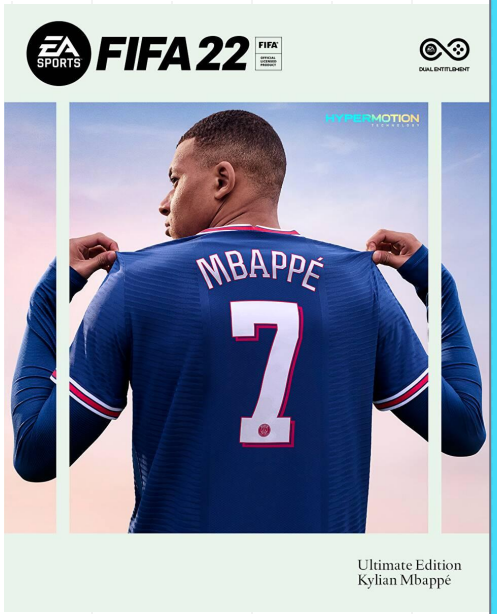


Introduction

1

Our Inspiration and Finding the Data Set

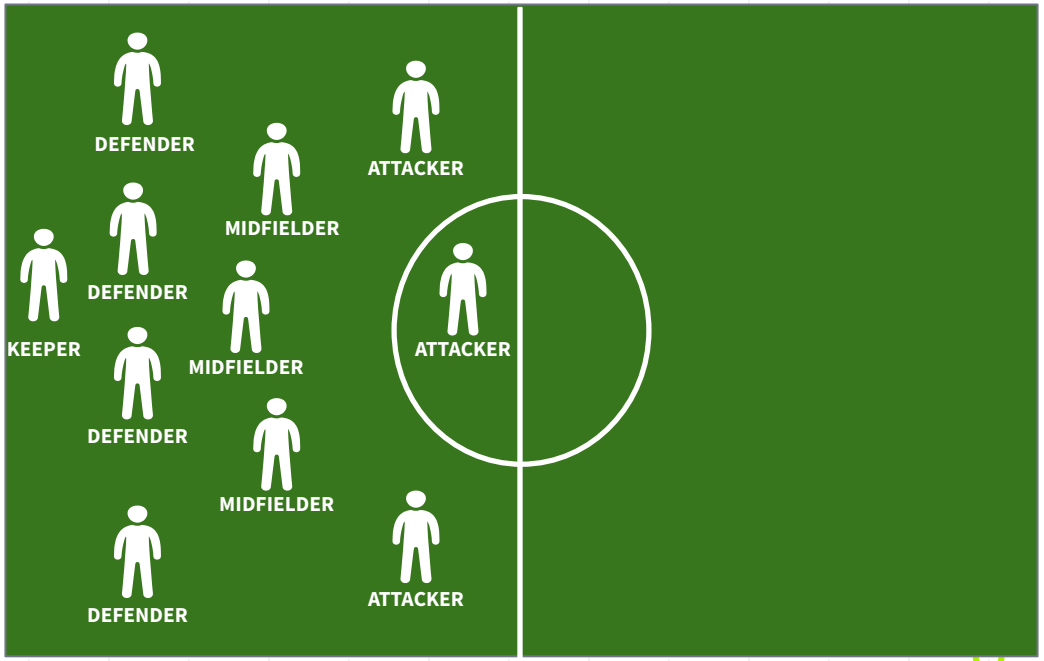
THE WORLD CUP INSPIRED US TO USE DATA FROM THE FIFA 22 GAME

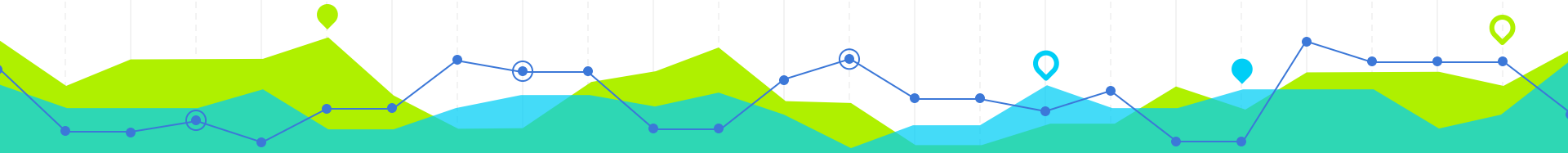


- ★ Found data on Github
- ★ Enjoy playing the video game
- ★ Input variables normalized
- ★ FIFA World Cup - relevant players on the biggest stage

THE MODEL PREDICTS A PLAYER'S POSITION BASED ON VARIOUS PHYSICAL AND SKILL ATTRIBUTES

- Uses physical attributes such as height and weight
- Combines with relevant skill/athletic attributes as rated by EA Sports
- 17 unique positions - further classified into four distinct positions:
  - Forward/attacker
  - Midfielder
  - Defender
  - Goalkeeper





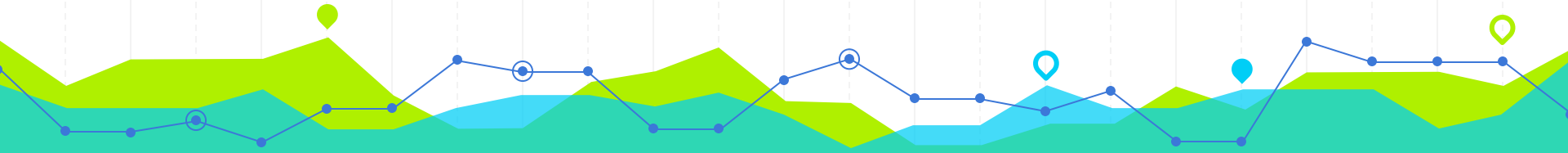
The Data

2

What do our original data look like?

19,329 records : all players recorded in FIFA 22

110 different attributes



Data preprocessing

Cleaning data with 110 variables

3

WHAT DO WE WANT IN THE MODEL?

- The variable that must be included: position!
- Other variables to help predict players' position:


Height/Weight


Attacking


Skill


Movement

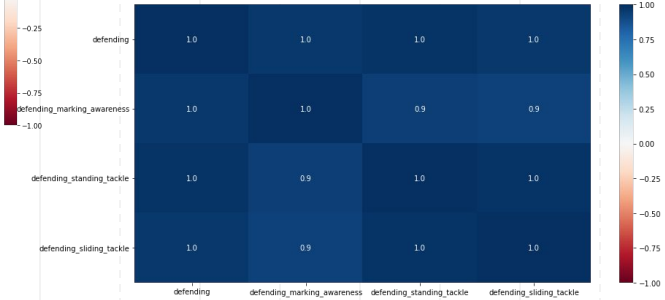

Power


Defending

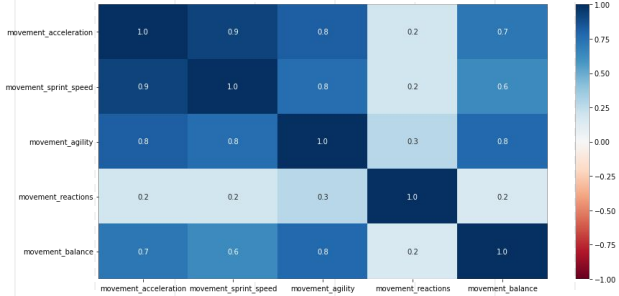
CORRELATIONS UNDER EACH CATEGORY

Attacking

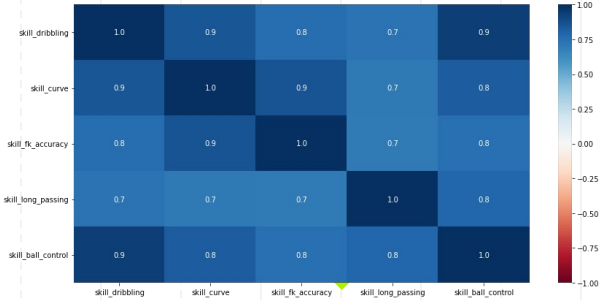
Defending



Movement



Skill



Power

Should we group them?

PCA OR COMPUTING AVERAGE?

	% of variance explained
0	0.395
1	0.274
2	0.098
3	0.054
4	0.024
5	0.021
6	0.018

	0	1	2	3	4	5	6
attacking_crossing	0.2	0.2	-0.2	0.0	0.2	0.3	0.4
attacking_finishing	0.3	0.1	0.2	-0.0	0.1	-0.2	-0.2
attacking_heading_accuracy	-0.1	0.1	0.4	-0.3	0.0	-0.0	0.3
attacking_short_passing	0.1	0.2	0.0	0.0	0.1	-0.2	0.2
attacking_volleys	0.3	0.1	0.2	0.0	-0.1	-0.1	-0.0
skill_dribbling	0.2	0.2	-0.1	-0.0	0.2	-0.1	0.2
skill_curve	0.2	0.2	-0.0	0.1	-0.1	0.4	0.2
skill_fk_accuracy	0.2	0.2	0.0	0.3	-0.4	0.5	-0.1
skill_long_passing	0.0	0.3	-0.0	0.1	0.1	-0.2	0.1
skill_ball_control	0.1	0.2	0.0	-0.0	0.2	-0.2	0.2
movement_acceleration	0.1	0.0	-0.3	-0.4	0.2	0.1	-0.0
movement_sprint_speed	0.1	0.0	-0.3	-0.4	0.3	0.2	0.0
movement_agility	0.2	0.1	-0.3	-0.2	-0.1	-0.1	-0.0
movement_reactions	0.0	0.2	0.1	-0.1	0.1	-0.1	0.1
movement_balance	0.1	0.1	-0.4	-0.1	-0.4	-0.3	-0.1
power_shot_power	0.2	0.2	0.2	0.0	0.0	-0.1	-0.2
power_jumping	-0.1	0.1	0.1	-0.5	-0.7	-0.0	0.2
power_stamina	0.0	0.2	-0.0	-0.3	0.1	0.1	-0.6
power_strength	-0.1	0.1	0.4	-0.3	0.2	0.3	-0.1
power_long_shots	0.3	0.2	0.2	0.1	-0.1	-0.1	-0.3
defending	-0.3	0.3	-0.0	-0.0	-0.0	-0.0	-0.0
defending_marking_awareness	-0.3	0.3	-0.0	0.0	-0.0	-0.0	-0.1
defending_standing_tackle	-0.3	0.4	-0.1	0.0	0.0	-0.0	-0.1
defending_sliding_tackle	-0.3	0.3	-0.1	0.0	0.0	-0.0	-0.0

Almost no variables explaining PCA components more than 50%

Difficult to decide which variables to use!

ANY MISSING VALUES?

	player_positions	defending
5	GK	NaN
7	GK	NaN
8	GK	NaN
12	GK	NaN
18	GK	NaN
20	GK	NaN
21	GK	NaN
26	GK	NaN
31	GK	NaN
34	GK	NaN

Remove Goalkeepers
to compute average

New Variables

Player_positions

Height

Weight

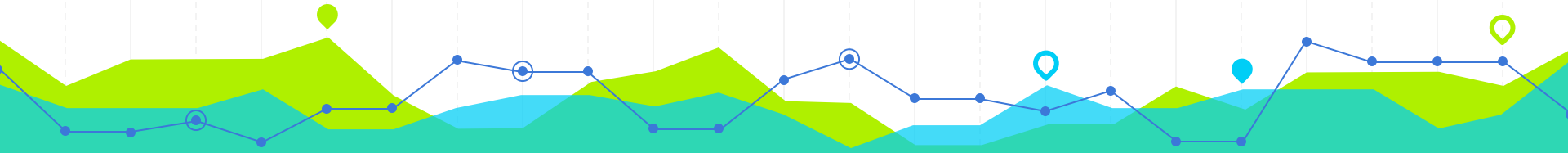
Attacking_average

Skill_average

Movement_average

Power_average

Defending_average



Our Model

4

How do we run the prediction model?

WHY WE CHOSE A DECISION TREE AS OUR MODEL

	Categorical Data	Multi-Class Outcome	Interpretable
Decision Tree	✔	✔	✔
kNN	✔	✔	✘
Neural Networks	✔	✔	✘
Logistic Regression	✔	✘	✘

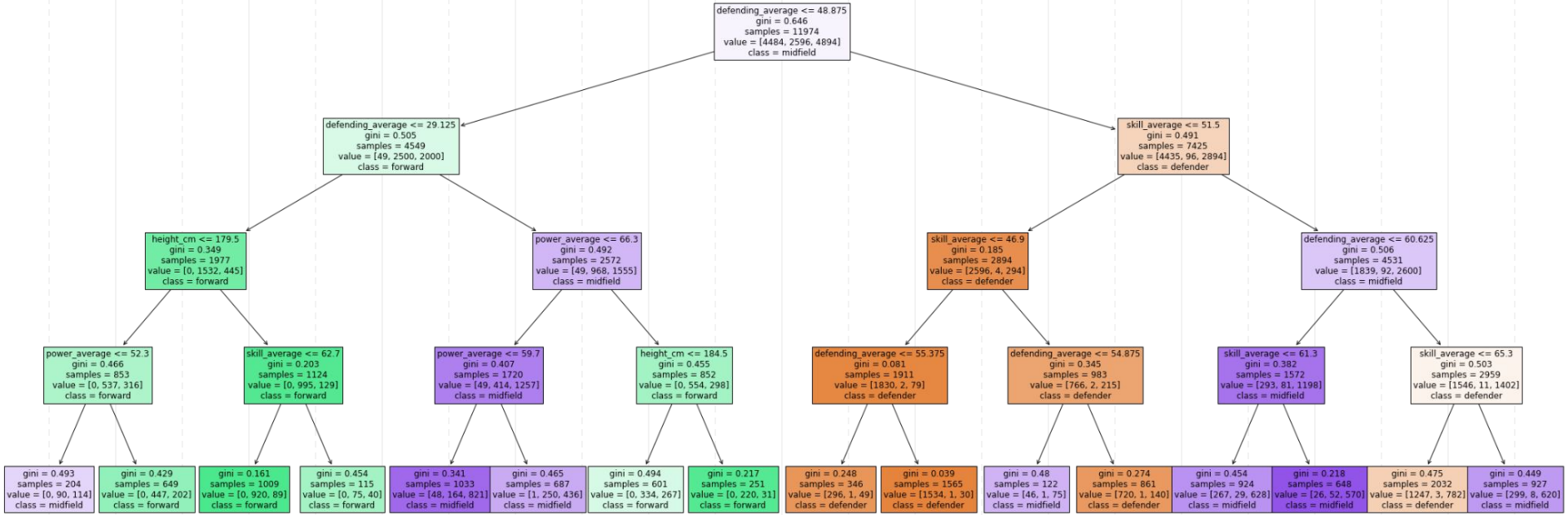
BUILDING AN OPTIMAL MODEL

Splitting data into test and train using a 70-30 split

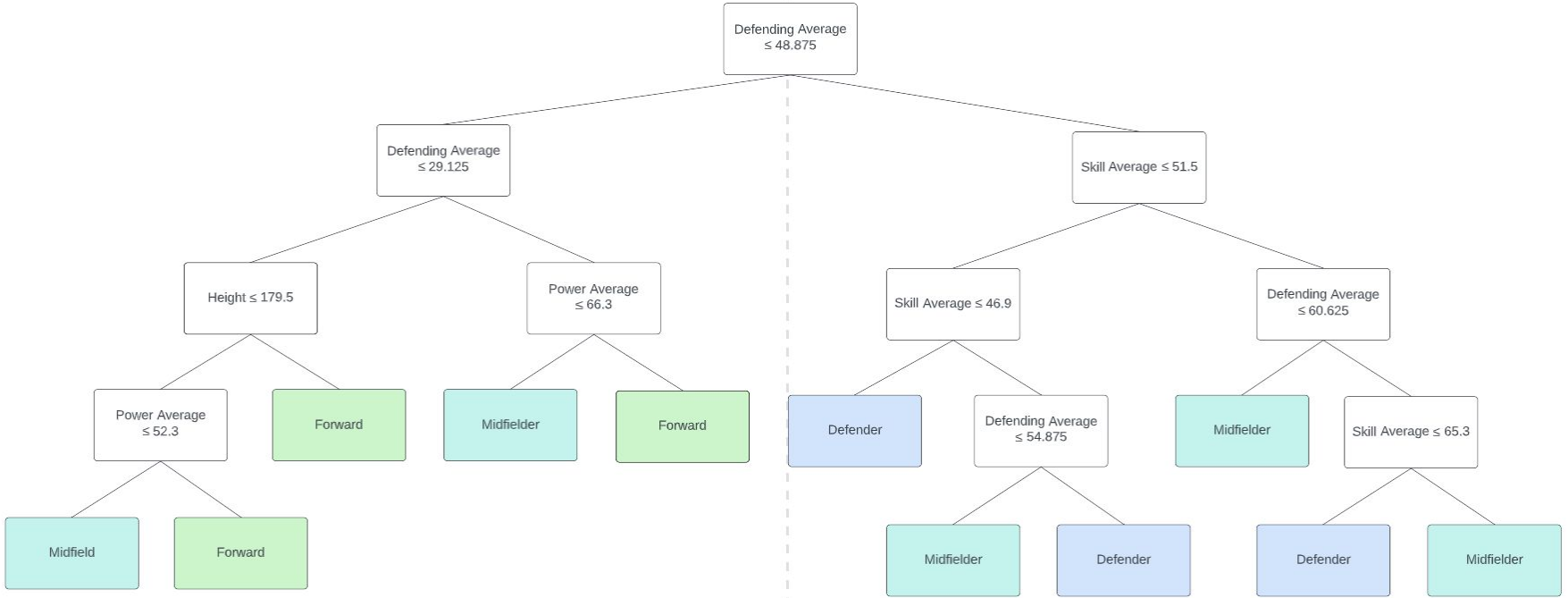
Running GridSearch to find optimal parameters

Tweaking constraints: limiting max depth to 4

PYTHON OUTPUT DECISION TREE



A SIMPLIFIED DECISION TREE

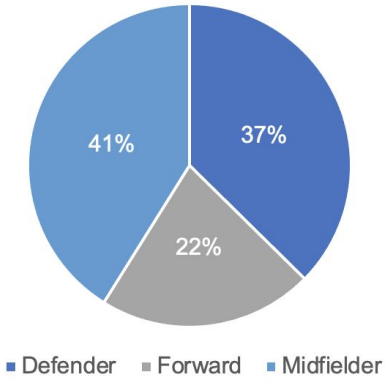


EVALUATING OUR MODEL

Actual	Predicted: DEFENDER	FORWARD	MIDFIELDER
Defender	1556	0	354
Forward	2	831	251
Midfielder	478	277	1384
Precision	76%	75%	70%

COMPARED TO BENCHMARKS

Original Composition of Data



33%

Accuracy using Random Classification

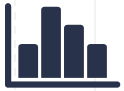
41%

Accuracy using Majority Classification

73%

Accuracy of Decision Tree

LIMITATIONS OF OUR MODEL



Using more accessible player attributes

No formula for each calculated player stat.

More accurately aggregating input variables

Low correlation between attributes in the power category.

Accounting for multiple positions per player

Model only accounts for one position but many players can do more than one.

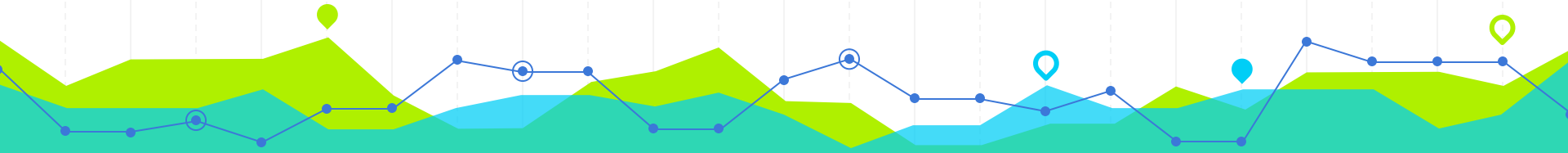
OTHER MODELS WE TRIED

K-Nearest Neighbors

- Had highest accuracy rate when K=27
- Accuracy rate of around 79%
- No insight to how classification happened

Neural Networks

- Used MLPClassifier function
- 10 nodes, logistic activation, lbfgs solver, and 1000 max iterations
- Accuracy rate of 80.3%
- No insight to how classification happened



Findings and Implications

5

How can the model be used in the real world?

KEY FINDINGS

Defensive stat matters the most when classifying players based on their average FIFA scores

Midfielders have the largest spread of their stats, with the most well rounded stats of the positions

There is a lot more that goes into choosing what position you play outside of individual stats

POSSIBLE IMPLICATIONS

FIFA - Customized Players

- In career mode in FIFA, you can create your own player
- You can level up certain skill sets to change the positions they play
- Predict the best positions for custom characters based on their unique skill set

	Height(CM)	Weight(KG)	Attacking	Skill	Movement	Power	Defending	Predicted Position
Akash	177	68	44	53	89	91	82	**Defender**
Oliver	180	90	85	98	92	93	70	**Midfielder**
Kimbo	175	70	69	73	88	67	55	**Midfielder**
Micah	178	56	90	42	79	92	34	**Forward**
Sami K	200	104	81	47	51	99	60	**Defender**

OTHER POSSIBLE IMPLICATIONS

Amateur Players

If players skills are translated into relative skill scores like those in FIFA, predictions on best matching position can be made.

FIFA - Create Your Best Team

Upgrade your favorite players' stats to fit a specific position (turn a defender into an attacker).

THANKS!

Any questions?