From an early age I have been fascinated by science fiction, robotics and sports. Whether it be watching Transformers, Star Wars or the NBA; I recall being in awe by the concept of the "impossible becoming possible".

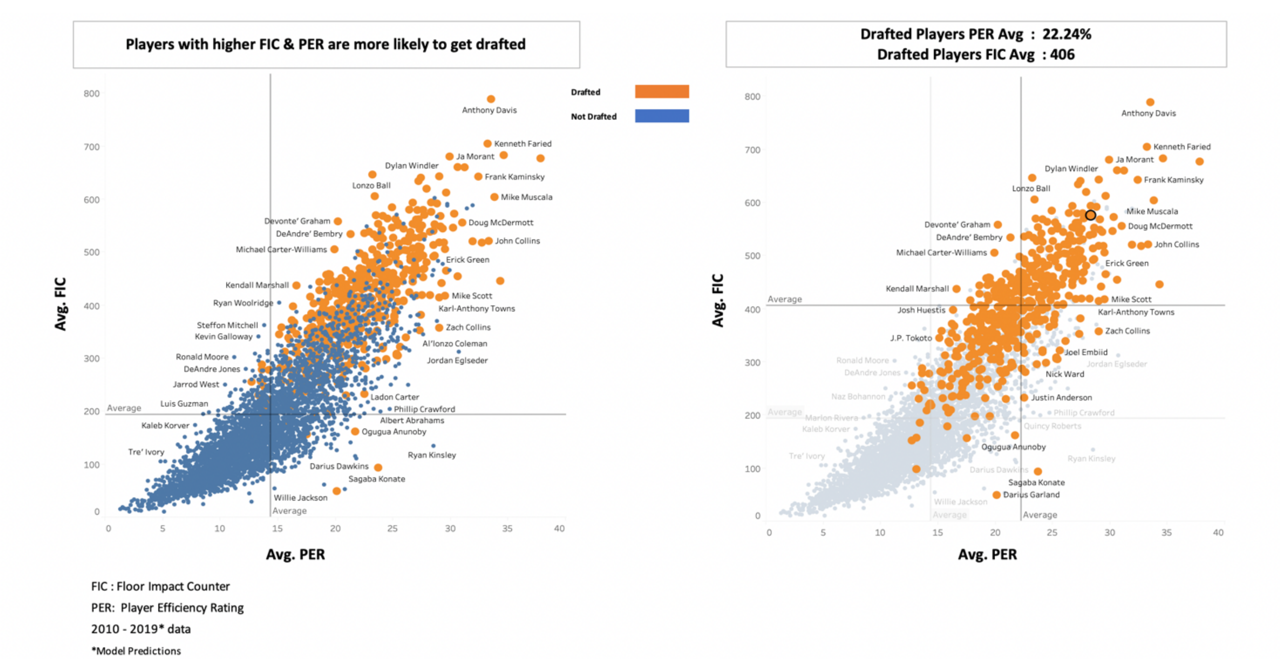
Seeing first-hand the impact of using cognitive technologies to drive more effective results, and overall positive impact on business and communities have exponentially heightened my interest in gaining more knowledge in technology, artificial intelligence, big data and machine learning.

Three months ago, I decided to join BrainStation to peruse a diploma in Data Science. Which turned out to be one of the most amazing experiences both professionally and personally. Where I was able to apply my learnings into my passion and interests.

Watching the results of the draft on Thursday June 20th was different for me this year than any other year. I was not just watching to see whose NBA prospects' dream will become possible. I was anticipating the results of my predictive model.

<u>Predicting the top 50 NCAA players entering the 2019 NBA Draft</u>

The NBA draft selection is a multi-million-dollar pivotal decision for every NBA team. They determine which player will bring the most success to their organization. This year's draft included big names like Zion Williamson as the #1 draft favourite. But, where will the rest of the exceptional players rank? We predicted the top 50 prospects entering the 2019 draft. The focus was on player's advanced stats in the most recent NCAA college basketball season.



Predictive Model Approach: I used Jupyter notebook and Python throughout the process.

Step 1. Data Collection and Scraping

I collected and scraped 10 years of NCAA players' advanced stats historical data and draft results from https://basketball.realgm.com. One condition I used when I scrapped the data was Qualified players only. Which means, 'a player must be on pace to play 15 minutes per game'. This effected any injured players during the NCAA season like top prospects Bol Bol and Darius Garland.

Step 2. Preprocessing and Cleaning

I used Python NumPy and Pandas packages to understand the data and build it into a consistent format

Step 3. Supervised Learning Model Building

With 22 variables, including 18 advanced stats. I started my feature selection process by first, checking the coefficient of variation of each numerical. Second, from sklearn.metrics I checked adjusted_mutual_info_score for each variable.
And ended up with 9 main variables (eFG%, ORB%, DRB%, FIC , PER, class, position, draft pick, and draft results) to start building the predictive model.

Building of the predictive model using Scikit learn:

- Target: draft results (Drafted or Not Drafted)
- Checked different classification models to predict the draft results with Logistic Regression giving the best results
- Used the coefficients of the Logistic Regression model to assess the variables of the model:
- The below 3 variables has the highest impact on player getting drafted
    - FIC , Position , PER%
- The below 3 variables shows a negative impact of the players selection process
    - Colyear , DRB%, ORB%

Step 4. Evaluating and scoring the predictive model

- Calculated the AUC curve score: 92.78%
- Determined where the threshold yields at least 70% tpr
- Used threshold to calculate probabilities of player's getting drafted
- Calculated recall, and precision scores and confusion matrix

Step 5. Predicted 2019 Draft Picks

- Displayed results of players getting drafted or not drafted
- Calculated the probability of each player getting drafted and determined the rank of each player

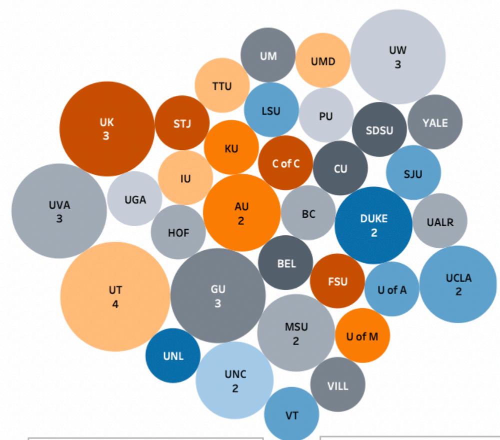
Results

- 82% accuracy on players draft prediction
- 58% accuracy on players' rank prediction based on round 1 and round 2 selection

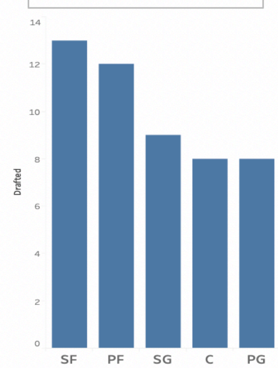
Top 50 Player's 2019 Draft Predictions

Player	Team_x	eFG%	ORB%	DRB%	FIC	PER	year	Class	Round	Position	Pick
Zion Williamson	DUKE	0.708	12.7	18.2	677	37.8	2019	Fr	Rd1	PF	1
Ja Morant	MSU	0.553	3.9	12.7	681	29.9	2019	Soph	Rd1	PG	2
Brandon Clarke	GU	0.693	13.9	19.7	683	34.6	2019	Jr	Rd1	PF	3
Grant Williams	UT	0.582	8.5	17.3	592	28.8	2019	Jr	Rd1	PF	4
R.J. Barrett	DUKE	0.506	4.8	17.2	536	22.5	2019	Fr	Rd1	SG	5
Rui Hachimura	GU	0.608	5.9	17	503	27.3	2019	Jr	Rd1	C	6
Dylan Windler	BEL	0.651	6.8	27.3	643	29	2019	Sr	Rd1	SF	7
Dedric Lawson	KU	0.525	11	23.2	562	27	2019	Jr	Rd1	PF	8
Jarrett Culver	TTU	0.505	5.7	17.2	489	24.1	2019	Soph	Rd1	SF	9
Chuma Okeke	AU	0.577	10.6	16.9	477	23	2019	Soph	Rd1	C	10
Justin Wright-Foreman	HOF	0.596	1.2	11.5	510	27.4	2019	Sr	Rd1	SG	11
Shamorie Ponds	STJ	0.521	3.1	10.3	501	25.4	2019	Jr	Rd1	PG	12
Bruno Fernando	UMD	0.612	12	27.5	503	25.6	2019	Soph	Rd1	C	13
De'Andre Hunter	UVA	0.579	5.6	13.3	411	23.1	2019	Soph	Rd1	PF	14
Daniel Gafford	U of A	0.66	11.2	22.7	452	26.7	2019	Soph	Rd1	C	15
Cameron Johnson	UNC	0.621	6.5	13.3	485	23.6	2019	Sr	Rd1	SF	16
Nickeil Alexander-Walker	VT	0.546	1.8	13	385	20.9	2019	Soph	Rd1	SF	17
P.J. Washington	UK	0.567	10.1	19.7	434	24.6	2019	Soph	Rd1	PF	18
Zach Norvell	GU	0.551	3.2	11.5	404	20.3	2019	Soph	Rd1	SG	19
Jared Harper	AU	0.505	1	8	426	19.3	2019	Jr	Rd1	PG	20
Ty Jerome	UVA	0.532	1.8	12.9	449	21.9	2019	Jr	Rd1	PG	21
Nicolas Claxton	UGA	0.49	8.3	20.7	404	20.2	2019	Soph	Rd1	C	22
Tyler Herro	UK	0.536	1.6	13.9	371	18.7	2019	Fr	Rd1	SG	23
Jarrell Brantley	C of C	0.559	6.6	24.2	483	26.9	2019	Sr	Rd1	PF	24
Jordan Bone	UT	0.527	1.9	8.5	404	18.1	2019	Jr	Rd1	SG	25
Tremont Waters	LSU	0.5	1.1	8.6	374	20.7	2019	Soph	Rd1	PG	26
Ignas Brazdeikis	UM	0.531	4.7	16.2	338	20.5	2019	Fr	Rd1	PF	27
Romeo Langford	IU	0.491	4.9	12.8	327	19	2019	Fr	Rd1	SF	28
Charlie Brown	SJU	0.496	6.4	13.2	362	20.9	2019	Soph	Rd1	SF	29
Coby White	UNC	0.516	1.3	11.4	341	18.6	2019	Fr	Rd1	PG	30
Quinndary Weatherspoon	MSU	0.574	4.7	11.6	405	22.8	2019	Sr	Rd2	SF	31
Jaylen Nowell	UW	0.561	3.5	14.9	383	19.4	2019	Soph	Rd2	SG	32
Carsen Edwards	PU	0.49	1.3	10.7	377	22.5	2019	Jr	Rd2	PG	33
Kyle Guy	UVA	0.585	2.7	12.5	392	19.7	2019	Jr	Rd2	SG	34
Amir Coffey	U of M	0.481	2.2	9.6	327	17	2019	Jr	Rd2	PF	35
Jaylen Hands	UCLA	0.499	1.3	10.5	358	17.8	2019	Soph	Rd2	PG	36
Justin James	UW	0.454	3.2	22.2	440	20.9	2019	Sr	Rd2	SF	37
Eric Paschall	VILL	0.528	4.2	15.9	377	18.2	2019	Sr	Rd2	C	38
Keldon Johnson	UK	0.521	5.5	16.4	329	17.9	2019	Fr	Rd2	SF	39
Ky Bowman	BC	0.486	2.3	18.2	403	19	2019	Jr	Rd2	SG	40
Jalen McDaniels	SDSU	0.492	7.3	22.7	374	19.9	2019	Soph	Rd2	PF	41
Jaxson Hayes	UT	0.728	8.8	16.4	306	25.3	2019	Fr	Rd2	PF	42
Matisse Thybulle	UW	0.5	3	8.9	358	18.7	2019	Sr	Rd2	SF	43
Martin Krampelj	CU	0.635	8.2	23.7	380	24.3	2019	Jr	Rd2	C	44
Isaiah Roby	UNL	0.497	6.6	17.8	370	18.7	2019	Jr	Rd2	PF	45
Rayjon Tucker	UALR	0.585	2.6	18.3	387	20.8	2019	Jr	Rd2	SG	46
Miye Oni	YALE	0.515	2.6	18.6	355	21.2	2019	Jr	Rd2	SF	47
Mfiondu Kabengele	FSU	0.538	11.5	20	353	27.3	2019	Soph	Rd2	SF	48
Admiral Schofield	UT	0.548	4.5	16.4	390	19.4	2019	Sr	Rd2	SF	49
Moses Brown	UCLA	0.607	15.6	21.1	329	21	2019	Fr	Rd2	C	50

Top 50 Player's Colleges



Top 50 Player's Position



Top 50 Player's Class