Evaluating NHL Players

Will Gibbs and Parker Gibbons

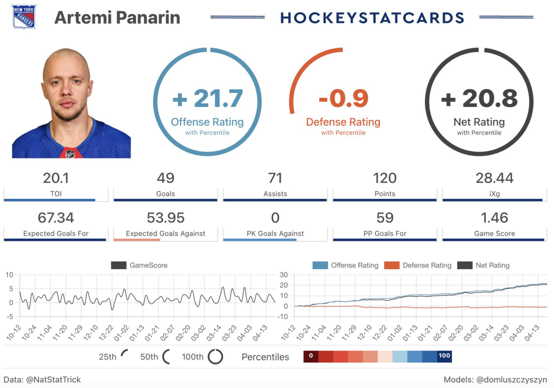
Project Motivation/Past Literature

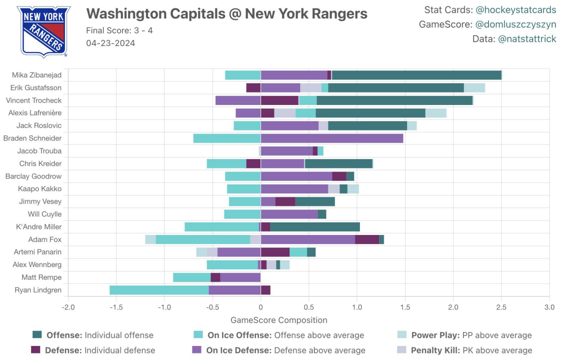
- **"An Expected Goals Model for Evaluating NHL Teams and Players" by Macdonald (2012)**
  - Used MLR and ridge regression to estimate a player's contribution to his team's expected goals per 60 minutes
  - Only based on shooting, did not take position into account
- **"Quantifying the Contribution of NHL Player Types to Team Performance" by Chan et al (2012)**
  - Used k-means clustering to classify players
  - Used a regression model to determine a quantitative relationship between different player types and their effect on a team's performance
  - Only used basic performance statistics
  - Concluded goalies contribute most to team performance
- **"An Analysis of Usage of a Multi-Criteria Approach in an Athlete Evaluation: An Evidence of NHL Attackers" by Vavrek (2021)**
  - Used TOPSIS methods to evaluate NHL forwards
- **Salary Cap: 83.5 million**

- **Is there a better way to evaluate players and their contribution to team performance?**

Game Score

- Player evaluation metric with the objective to measure single game player productivity
- Weights each of the actions of each player to give us a single number representing their overall performance in that game
- Formula not publicly available





Capitals @ Rangers

3 - 4

04-23-2024

Madison Square Garden

Subscribe to Capitals | HockeyStatCards.com | Subscribe to Rangers

PLAYER	TEAM	TOI	G	A	IXG	BLK	PEN	FO	XGF	XGA	GF	GA	GAME SCORE	AVG
Tom Wilson	Capitals	16.23	1	1	0.31	0	1	0	0.2	0.08	0	0	2.38	0.35
Mika Zibanejad	Rangers	18.75	1	1	0.57	1	0	-1	0.24	0.12	0	0	2.14	0.90
Erik Gustafsson	Rangers	15.75	0	2	0.01	1	-1	0	0.4	0.7	0.92	0	2.00	0.50
Vincent Trocheck	Rangers	22.67	1	1	0.31	0	2	9	0.29	0.67	0.92	1.06	1.73	1.10
Dylan McIlrath	Capitals	11.30	0	0	0.01	0	0	0	0.43	0.03	1.06	0	1.66	1.18
Alexis Lafrenière	Rangers	14.23	0	2	0.08	0	1	-1	0.26	0.49	0.92	1.06	1.46	0.80
Connor McMichael	Capitals	17.00	1	0	0.39	0	-1	-3	0.59	0.55	1.06	0	1.36	0.22
Dylan Strome	Capitals	14.87	1	0	0.23	0	0	2	0.2	0.09	0	0	1.35	0.77
Alexander Alexeyev	Capitals	17.27	0	0	0.00	3	0	0	0.65	0.15	0	0	1.28	-0.03
Trevor van Riemsdyk	Capitals	20.42	0	0	0.05	1	1	0	0.78	0.27	0	0	1.28	0.15
Jack Roslovic	Rangers	9.55	1	0	0.13	0	0	0	0.27	0.09	0	0	1.24	0.45
Hendrix Lapierre	Capitals	10.97	0	1	0.14	0	0	1	0.58	0.06	0	0.92	1.11	0.27
Alex Ovechkin	Capitals	17.07	0	0	0.08	0	0	0	0.59	0.37	1.06	0	1.05	0.45
Max Pacioretty	Capitals	12.63	0	1	0.34	0	0	0	0.58	0.08	0	0.92	1.01	0.31
Lucas Johansen	Capitals	11.93	0	0	0.03	0	1	0	0.43	0.13	0	0	0.96	-0.30
T.J. Oshie	Capitals	15.28	0	0	0.01	0	-1	0	0.59	0.26	1.06	0	0.89	0.23
Braden Schneider	Rangers	16.07	0	0	0.00	0	0	0	0.39	0.13	0	0	0.78	0.10
Barclay Goodrow	Rangers	13.58	0	0	0.15	2	0	5	0.31	0.12	0	0	0.60	-0.00
Chris Kreider	Rangers	16.82	0	1	0.28	0	-1	0	0.04	0.18	0	0	0.60	1.02
Kaapo Kakko	Rangers	11.75	0	0	0.14	0	0	0	0.47	0.2	0	0	0.55	0.32
Jacob Trouba	Rangers	15.35	0	0	0.00	2	-1	0	0.36	0.61	0.92	0	0.53	0.17
Aliaksei Protas	Capitals	14.23	0	0	0.00	1	0	1	0.2	0.1	0	0	0.47	0.40
Jimmy Vesey	Rangers	11.83	0	0	0.74	1	1	1	0.31	0.42	0	0	0.44	0.28
Sonny Milano	Capitals	10.98	0	0	0.10	0	0	0	0.58	0.1	0	0.92	0.39	0.54
Will Cuylle	Rangers	10.30	0	0	0.16	0	0	0	0.29	0.2	0	0	0.29	0.28
John Carlson	Capitals	26.88	0	1	0.04	1	-1	0	0.41	0.68	1.06	0.92	0.28	0.33
K'Andre Miller	Rangers	18.80	1	0	0.32	2	0	0	0.46	0.39	0	1.06	0.23	0.17
Adam Fox	Rangers	23.37	0	0	0.09	5	0	0	0.16	0.46	0	0	0.19	1.26
Artemi Panarin	Rangers	20.65	0	0	0.16	0	2	0	0.35	0.67	0.92	1.06	0.01	1.46
Beck Malenstyn	Capitals	13.53	0	0	0.00	4	0	0	0.09	0.44	0	0	-0.25	0.14
Nicolas Aube-Kubel	Capitals	7.52	0	0	0.00	2	-1	0	0	0.28	0	0	-0.26	0.24
Alex Wennberg	Rangers	15.57	0	0	0.07	2	0	-4	0.16	0.62	0	0	-0.37	0.46
Nic Dowd	Capitals	14.72	0	0	0.00	2	-1	-10	0.09	0.54	0	0	-0.71	0.21
Matt Rempe	Rangers	7.57	0	0	0.00	1	-1	0	0.02	0.61	0	0	-0.91	0.24
Igor Shesterkin	Rangers		3							1.86			-1.14	0.20
Martin Fehervary	Capitals	21.02	0	0	0.12	4	-1	0	0.3	0.7	0	0.92	-1.23	0.01
Ryan Lindgren	Rangers	17.70	0	0	0.00	2	0	0	0.16	0.62	0	1.06	-1.47	0.24
Charlie Lindgren	Capitals		4							2.46			-1.54	0.17

GameScore Algorithm:@domluszczyszyn   Data:@natstattrick

Our Approach

- **Research Question:** How can we alternatively evaluate NHL players?
- First: Use k-means clustering to classify players at each position (forwards, defensemen, goalies)
- Second: Random Forests: Creating our own game score using our own metrics
- Third: Use TOPSIS method to evaluate players based on our own metric, taking the clusters into account
- Fourth: Team Case Studies: Where do the good and bad teams get value from? How much money to they spend to get this value?

Data

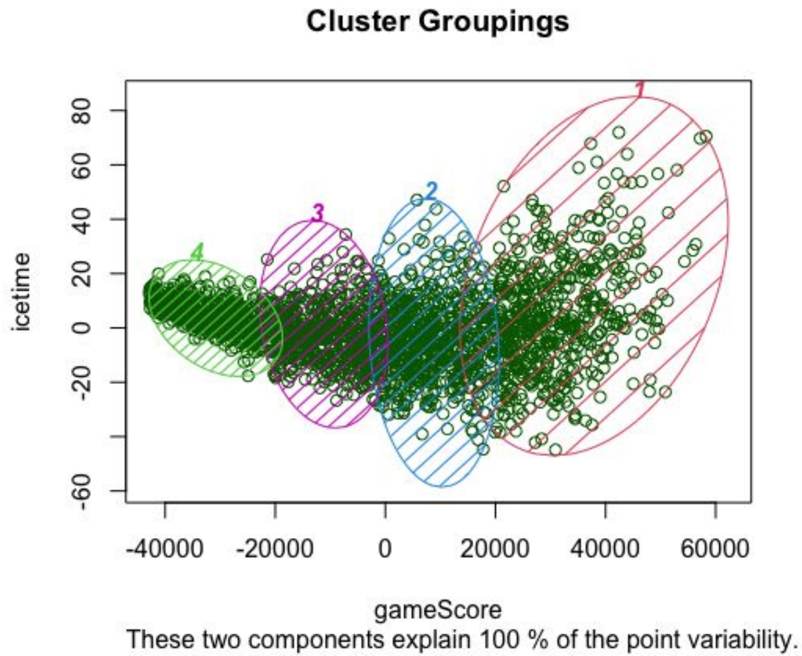
- **All situation data since 2019-2020 season**
- **Worked with correlations and MLR to wrangle 150 predictors down**
- **Forwards/Defensemen (20)**
  - Games, Icetime, Shifts, Game Score, xGoals, Goals, xRebounds, Rebounds, Assists, Points, Hits, Takeaways, Giveaways, Low danger goals above expected, Medium danger goals above expected, High danger goals above expected, Faceoff wins, PIM, PIM drawn, Shots blocked
- **Goalies (11)**
  - Games, Icetime, Goals saved above expected (game score), Goals, xGoals, Rebounds above expected, Play continued above expected, Low danger goals above expected, Medium danger goals above expected, High danger goals above expected, PIM

K-Means Clustering

- Unsupervised non-linear algorithm that clusters data based on their similarity to one another
- Uses a a pre-specified number of clusters
  - Used elbow method (we used 4 clusters for forwards/defensemen, 3 for goalies)
- We want to classify types/tiers of players at each position (forwards, defensemen, goalies)
  - Will allow us to later see how each types of player at each position contribute to team success

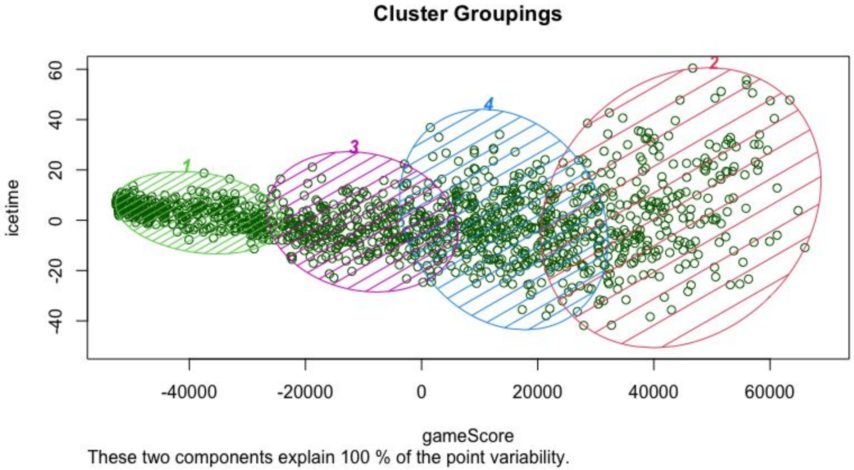
Clustering Forwards

Cluster	Number of Players
1- Top Six Player	162
2 - Middle Six Player	133
3- Defensive Forward	89
4 - Bottom Six Player	91



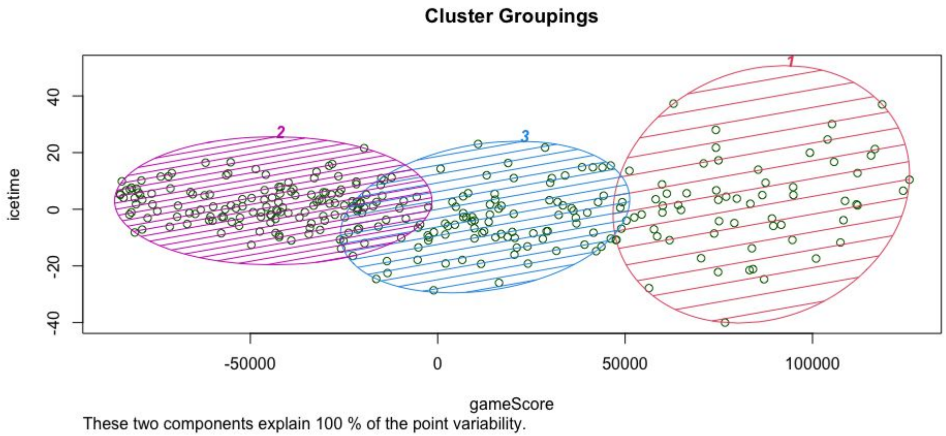
Clustering Defensemen

Cluster	Number of Players
1 - Bottom Pair Player	54
2 - Top Pair Player	75
3 - Pure Defensive	59
4 - Second Pair Player	74



Clustering Goalies

Cluster	Number of Players
1 - Elite	22
2 - Below-Average	27
3 - Average	27



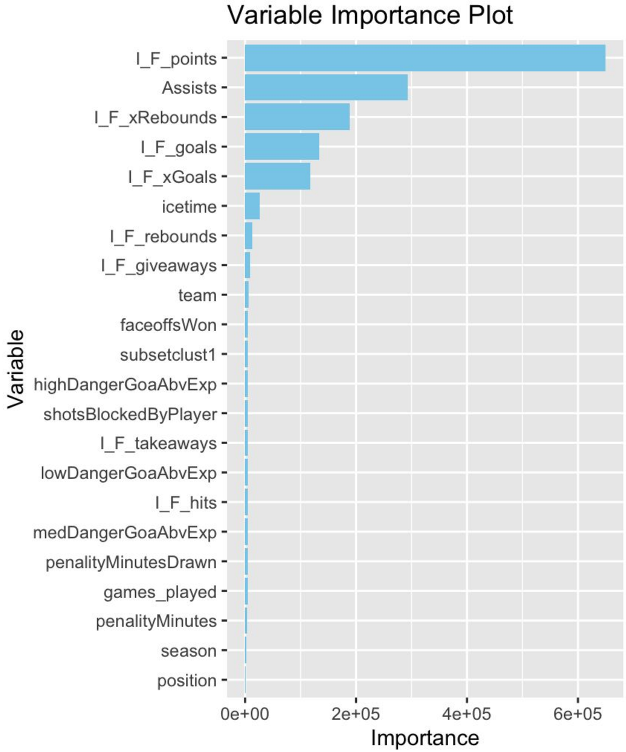
Random Forests

Forwards Random Forest

Explained 93.83% of variance

Mtry value of 10

300 trees

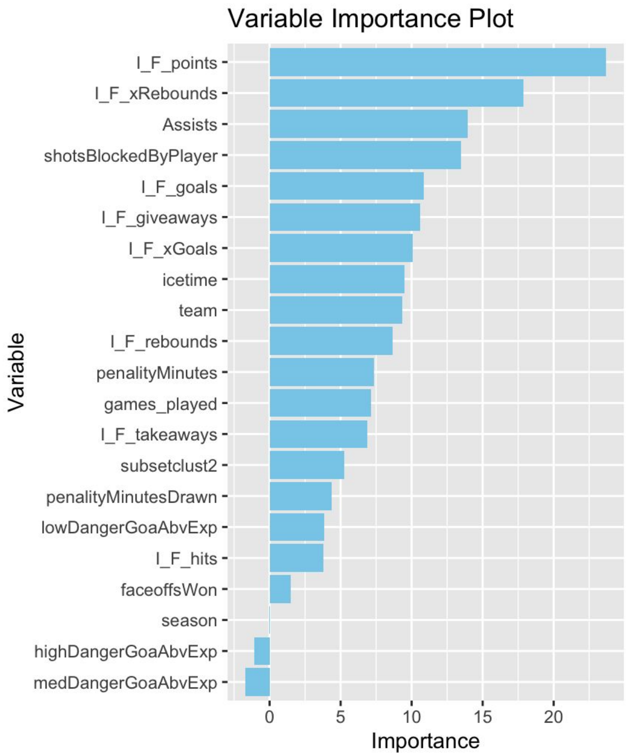


Defensemen Random Forest

Explained 86.57% of variance

Mtry value of 10

300 trees

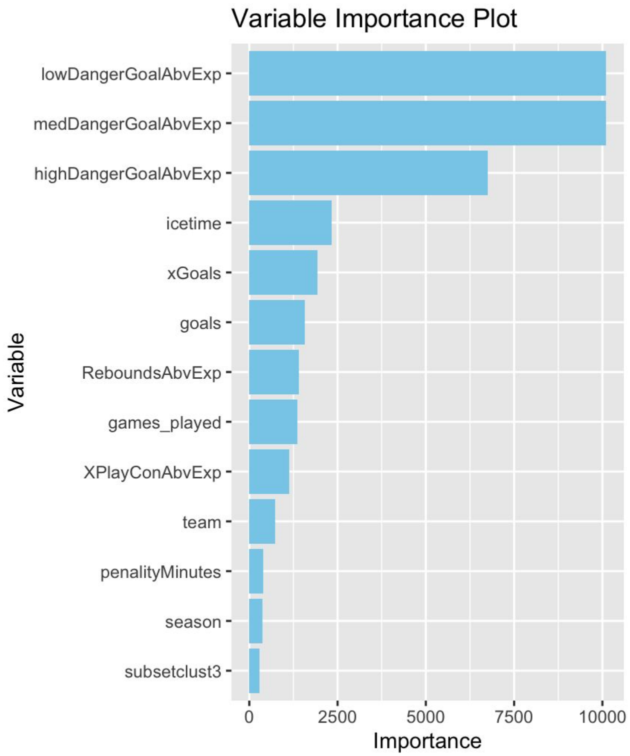


Goalie Random Forest

Explained 84.27% of variance

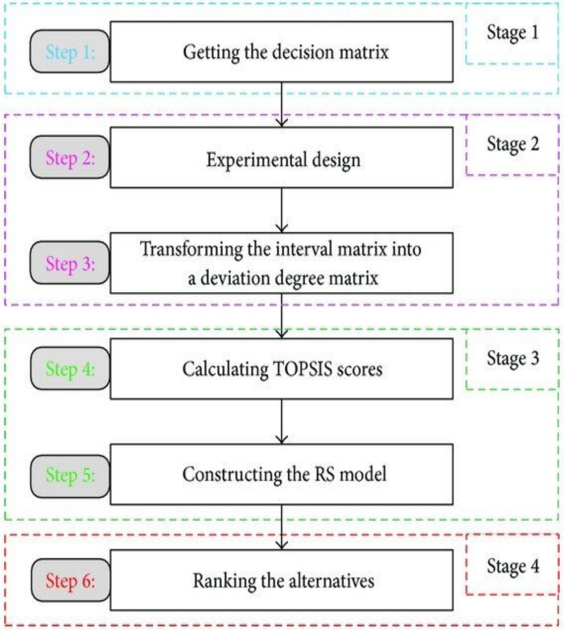
Mtry value of 4

100 trees



TOPSIS Method

What is the TOPSIS Method?



Different from other methods because of how it calculates next best alternatives

Calculates the alternatives based on euclidean distance to the best case scenario

TOPSIS Results

Top Forwards:

1) Brady Tkachuk (2022)
2) Nathan MacKinnon (2019)
3) Connor McDavid (2022)
4) Nathan MacKinnon (2023)
5) Matthew Tkachuk (2022)

Top Defenders:

1) Cale Makar (2023)
2) Rasmus Dahlin (2022)
3) Roman Josi (2019)
4) Brandon Montour (2022)
5) Rasmus Dahlin (2023)

Top Goalies:

1) Juuse Saros (2022)
2) Andrei Vasilevskiy (2021)
3) Juuse Saros (2021)
4) Elvis Merzlikins (2021)
5) Alexandar Georgiev (2022)

Case Study - NYR VS CBJ (2023-24)

Forward Cluster	NYR Value	CBJ Value	NYR Salary	CBJ Salary
Middle Six Forward	68.75	114.71	7,270,000	7,175,000
Top Six Forward	461	123.61	36,842,857	11,600,000
Defensive Forward	57	90.11	4,462,500	6,275,000
Bottom Six Forward	0.09	24.57	800,000	11,135,000

Defensive Cluster	NYR Value	CBJ Value	NYR Salary	CBJ Salary
Bottom Pair Player	9.10	1.70	812,500	825,000
Top Pair Player	98.98	95.07	13,372.000	18,308,333
Defensive Player	3.22	18.57	800,000	3,518,333
Second Pair Player	109.08	37.25	12,750,000	8,583,333

Goalie Cluster	NYR Value	CBJ Value	NYR Salary	CBJ Salary
Elite	12.08	None	5,666,667	None
Poor	None	3.90	None	1,808,333
Average	9.84	-6.35	825,000	5,400,000

Case Study - DAL VS SJS (2023-24)

Forward Cluster	DAL Value	SJS Value	DAL Salary	SJS Salary
Bottom Six Forward	20.62	5.99	814,167	2,479,167
Top Six Forward	103.33	27.67	13,950,000	10,350,000
Defensive Forward	63.94	44	4,150,000	10,775,000
Middle Six Forward	489.55	90.37	37,594,167	7,313,333

Defensive Cluster	DAL Value	SJS Value	DAL Salary	SJS Salary
Bottom Pair Player	None	10.25	None	2,887,500
Top Pair Player	201.69	10.33	18,763,333	3,250,000
Defensive Player	26.37	3.45	925,000	7,912,500
Second Pair Player	42.47	8.37	26,250,000	46,750,000

Goalie Cluster	DAL Value	SJS Value	DAL Salary	SJS Salary
Elite	1.95	None	4,000,000	None
Poor	None	-7.62	None	867,500
Average	-9.37	2.45	1,000,000	2,350.000

Conclusions

- Forwards value is very scoring based (points, assists, goals) and puck awareness (rebounds), good teams rely on superstars
- Defensemen value based on puck possession (rebounds, giveaways), some scoring (assists)
- Goalies value based on saving less difficult shots (good goalies save most poor shots) - Could be the most valuable
- Clusters can determine tiers for evaluations for each position group
- Using the TOPSIS method, we can build models for evaluating the most valuable players at each position
- While we do not have an exact formula for game score, we can create fairly accurate estimates

Future Work

- Use more data and more sabermetrics
- Do a deeper dive into each team, where they get value and how they spend their money
- Further tune the TOPSIS model, as the weighting is very abstract
- Implement other machine learning techniques, such as XGBoosting
- Look into how to compare the three different positions - which position generally contributes most to team success? How much should be spent on each position?

Sources

- Moneypuck (statistics): https://moneypuck.com/index.html
- Capfriendly (salaries): https://www.capfriendly.com/
- What is TOPSIS?: https://robertsoczewica.medium.com/what-is-topsis-b05c50b3cd05
- TOPSIS method for Multiple-Criteria Decision Making (MCDM): https://www.geeksforgeeks.org/topsis-method-for-multiple-criteria-decision-making-mcdm/