Prediction of tiers in the ranking of ice hockey players

Timmy Lehmus Persson   Haris Kozlica   Niklas Carlsson   Patrick Lambrix

Linköping University, Sweden

Abstract. Many teams in the NHL utilize data analysis and employ data analysts. An important question for these analysts is to identify attributes and skills that may help predict the success of individual players. This study uses detailed player statistics from four seasons, player rankings from EA's NHL video games, and six machine learning algorithms to find predictive models that can be used to identify and predict players' ranking tier (top 10%, 25% and 50%). We also compare and contrast which attributes and skills best predict a player's success, while accounting for differences in player positions (goalkeepers, defenders and forwards). When comparing the resulting models, the Bayesian classifiers performed best and had the best sensitivity. The tree-based models had the highest specificity, but had trouble classifying the top 10% tier players. In general, the models were best at classifying forwards, highlighting that many of the official metrics are focused on the offensive measures and that it is harder to use official performance metrics alone to differentiate between top tier players.

1 Introduction

The success of a sports team depends a lot on the individual players making up that team. However, not all positions on a team are the same. In ice hockey there are three main types of players: goalkeepers, defenders and forwards. While evaluating players it is therefore important to take into account these types.

In this paper, we compare and contrast which attributes and skills best predict the success of individual ice hockey players in different positions. First, using the method in [14] we investigate which performance features were important for the three main position types in the National Hockey League (NHL) for four different seasons. For the data processing, feature selection and analysis we used R 3.6.3 and packages dplyr 0.8.3, ggplot2 3.0.0, gridExtra 2.3 and caret 6.0 as well as Weka 3.8.4 [6]. Our work (including [14] for football) distinguishes itself from other work on player valuation or player performance, by working with tiers of players, i.e., the top 10%, 25% and 50% players in different positions (in contrast to individual ratings). An exact ranking of players may not always be available, and for several tasks, e.g., scouting, an exact ranking of players is not necessary. In these cases using tiers is a useful approximation. Further, we deal with many skills.

Second, we evaluate different techniques for generating prediction models for players belonging to the different top tiers of players. We used Weka 3.8.4 for estimation of the models. We found that the two Bayesian classifiers performed best and that, in general, the models were best at classifying forwards.

The remainder of the paper is organized as follows. Sect. 2 presents related work. Sect. 3 discusses the data sets and the data preparation. Sect. 4 and 5 present the feature selection and prediction methods, respectively, and show and discuss the corresponding results. Finally, the paper concludes in Sect. 6.

2   Related work

In many sports work has started on measuring player performance. For the sake of brevity, we address the related work in ice hockey.

Many of the models for evaluating player performance in ice hockey define a particular stat or evaluation measure that assigns values based on particular types of actions in the game. For instance, the well-known goal measure, assist measure, and the more recent Fenwick and Corsi measures[1] attribute a value to goal-scoring actions, to passes that lead to goals and to different types of shots, respectively. To deal with some of the weaknesses of traditional measures new approaches have been proposed, including regression models replacing the $+/-$ measure (e.g., [12, 13, 3]). One main recognized weakness is the lack of influence of the context in which the actions are performed. This is the basis of the work on added goal value [15] that attributes value to goals, but the value of the goal is dependent on the situation in which it is scored.

Recent works often take several kinds of actions into account for defining a measure. For instance, in [4] principal component analysis was performed based on 18 traditional measures and a performance measure based on the four most important components was proposed. Further, many of these approaches also take some context into account. For instance, event impacts for different kinds of actions in [19] are based on the probability that the event leads to a goal (for or against) in the next 20 seconds. Several works model the dynamics of an ice hockey game using Markov games (e.g., [22, 7]). In [16, 20, 21, 9] action-value Q-functions are learned with respect to different targets. The proposed measure in [9] showed the highest correlation to 12 out of 14 traditional measures compared to measures such as $+/-$, goal-above-replacement, win-above-replacement and expected goals. In [17] the action-value Q-functions are used to define variants of these player impact measures. In [11] action-value Q-functions are used to define measures for pairs of players. Player rankings used for the NHL draft are presented in [18, 10].

---

[1] See, e.g., `https://en.wikipedia.org/wiki/Analytics_(ice_hockey)`.

Table 1: Attributes for field players and goalkeepers. Attributes in italics were removed during data preparation.

Position	Attributes
Field players	Player, Age, Team, POS (position), GP (games played), G (goals), A (assists), *PTS* (points), +/-, PIM (penalty minutes), PS (point shares), EVG (even strength goals), PPG (powerplay goals), SHG (shorthanded goals), GWG (game-winning goals), EVA (even strength assists), PPA (powerplay assists) , SHA (shorthanded assistss), S (shots on goal), *S%* (shots on goal percentage), TOI (time on ice), *TOI/60*, BLK (blocks), HIT (hits), FWON (face-offs won), FOL (face-offs lost), *FO%* (face-off percentage), CF (Corsi For), CA (Corsi Against), *CF%* (Corsi For percentage), CF%Rel (Corsi For percentage relative), FF (Fenwick For), FA (Fenwick Against), *FF%* (Fenwick For percentage), FF%Rel (Fenwick For percentage relative), oiSH% (on ice shooting percentage), oiSV% (on-ice save percentage) *PDO*, oZS% (offensive zone start percentage), dZS% (defensive zone start percentage), TOI(EV) (time on ice even strength), TK (takeaways), GV (giveaways), E+/- (expected +/-), SAtt. (shot attempts), Thru% (through percentage), SHFT (shift length), EVTOI (even strength time on ice), GF/60 (even strength Goals For per 60 minutes), GA/60 (even strength Goals Against per 60 minutes), PPTOI (powerplay time on ice), *PPCF%Rel* (powerplay Corsi For percentage relative), *PPGF/60* (powerplay goals for per 60 minutes), *PPGA/60* (powerplay goals against per 60 minutes), SHTOI (shorthanded time on ice), *SHCF%Rel* (shorthanded Corsi For percentage relative), *SHGF/60* (shorthanded Goals For per 60 minutes), *SHGA/60* (shorthanded Goals Against per 60 minutes)
Goalkeepers	Player, Age, Team, GP (games played), GS (game starts), W (wins), L (losses), OTL (overtime losses), GA (goals against), SA (shots against), *SV* (saves), *SV%* (save percentage), *GAA* (goals against average), SO (shutouts), GPS (goalkeeper point shares), MIN (minutes), QS (quality starts), *QS%* (quality starts percentage), RBS (really bad starts), *GA%* (goals against percentage), *GSAA* (goals saved above average), G (goals), A (assists), PTS (points), PIM (penalty minutes)

3 Data collection and preparation

3.1 Data collection

The data regarding players was taken from Hockey Reference[2] for the seasons 2015/16 to 2018/19. Different attributes were gathered for goalkeepers and field players. The lists of attributes are given in Table 1. Descriptions of the attributes are given in the extended version of this paper [8].

The ranking used as a response variable was directly taken from Electronic Arts NHL games between 2016 and 2019 (NHL17, NHL18, NHL19 and NHL20).

---

[2] https://www.hockey-reference.com/

Table 2: Number of players per position with ratings. In parentheses we show the number of players without ratings that were removed from the data set.

Season	Forwards	Defenders	Goalkeepers
2015/16	582 (10)	297 (9)	91 (1)
2016/17	572 (17)	287 (12)	90 (5)
2017/18	555 (28)	297 (10)	93 (2)
2018/19	545 (35)	302 (24)	87 (8)

We use the player rating value that is supposed to be a summary of a player's individual attributes[3]. The range for this value is between 1 and 99.

3.2 Data preparation

The data was then split using player position: goalkeepers, defenders, and forwards[4], resulting in 12 data sets (3 player positions × 4 seasons). As some of the players did not have a rating in the NHL games, data about these players was removed. Table 2 shows the number of retained players per position and the number of removed players.

For each of the data sets, attributes that were combinations of other attributes were removed. For field players these are G, A, PTS, S%, TOI/60, FO%, CF%, FF%, and PDO. For goalkeepers these are SV, SV%, GAA and QS%. Further, G was removed for goalkeepers as no goalkeeper scored those seasons. For other attributes data was missing and it was decided to impute the value 0 (Thru%, oiSH%, oiSV%, oZS%, dZS%) or remove the attribute (PPCF%Rel, SHCF%Rel, PPGF/60, PPGA/60, SHGF/60, SHGA/60, GA%, GSAA). All temporal attributes were rewritten into seconds. The value for Team was set to the team for which the player played the most games or in case of a tie to the team in which the player ended the season. Numerical data was normalized using the min-max-method to values between 0 and 1.

The rating was used to create the top 10%, 25% and 50% tiers. However, as several players had the same rating it was not always possible to take a tier without having players with the same rating in the tier and outside the tier. Therefore, we decided to use a cutoff such that the actual percentages are less than or equal to the desired percentage for the tier. Using this strategy the actual percentages for the top 10%, 25% and 50% tiers for the different position and seasons were between 6.5% and 9.3%, 19.5% and 25%, and 39.6% and 49.3%, respectively. The exact numbers for each data set are given in [8].

For each of the data sets resulting from the steps above, we made an 80%-20% split where the 80% is used in the feature selection (Sect. 4) and as training set in the prediction (Sect. 5) while the 20% is used as test set in the prediction.

---

[3] https://www.ea.com/games/nhl/nhl-20/ratings

[4] In the original data the forwards were categorized as left wing, right wing, center and wing.

4 Feature selection

4.1 Filter method

Filter methods for feature selection examine data using statistical methods to determine which attributes are relevant. They often use relatively simple calculations and are often relatively fast. We used correlation-based feature selection (CFS) which aims to identify sets of attributes that are highly correlated to the classification, but not correlated with each other [5]. Essentially, CFS computes the Pearson correlation coefficient where all attributes have been standardized and uses this as a measure of merit for the attribute subsets. Further, we used 10-fold cross validation. This results in different subsets for the different runs. We retained the attributes that appeared in at least two of these subsets.

4.2 Wrapper method

Wrapper methods try to identify which subsets of attributes give the best results when used in a model by testing combinations of attributes. Wrapper methods employ a supervised learning method to compute the merit of each subset and are thus dependent on the chosen learning method.

We used the machine learning methods Logistic Regression (LR), Naïve Bayes (NB), Bayesian Network (BN) with $\alpha = 0.1$ and u = 1, Decision Tree (DT) with C = 0.25 and M = 2, k-Nearest Neighbor (KNN) with k = 3 and Random Forest (RF) with I = 100. For the Bayesian methods the attributes should be of nominal type and therefore the values of all numeric-type attributes were discretized by creating ten intervals with a width of 0.1 and ranging from 0 to 1 [2].

We used the Weka settings $\epsilon = 0.01$ and k = 5. This means that we started from the empty set and used best-first search with backtracking after five consecutive non-improving nodes in the search tree. As measure for merit we used AUC. Each algorithm was run over 10 folds and for each attribute and each algorithm the number of folds that contained the attribute was registered. Then for each attribute the mean over this number for the different algorithms was computed and if this mean was larger than 2 the attribute was retained.

4.3 Results and discussion

Table 3 shows the number of attributes that were retained per position, tier and season for both the filter and wrapper methods. Table 4 shows the most common attributes per position for the filter and wrapper methods. The full list of attributes for each data set is given in [8].

For goalkeepers W and QS were common for several tiers in the same season for both methods, while GPS was also common for the filter method. QS was important for all tiers over all seasons for the filter method. For the wrapper method SO was important over all seasons for the 25% and 50% tiers. For defenders PS and PPA were important for all tiers and all seasons for the filter method, while TOI(EV) and S appeared often. For the wrapper method PPTOI

Table 3: Number of retained attributes for the filter and wrapper methods, respectively. (filter/wrapper).

Season	Tier	Goalkeepers	Defenders	Forwards
	Top 10%	3/7	9/13	10/11
2015/16	Top 25%	2/5	17/8	12/13
	Top 50%	5/5	21/11	22/14
	Top 10%	5/5	11/11	16/14
2016/17	Top 25%	5/5	13/12	14/11
	Top 50%	9/6	17/13	23/9
	Top 10%	2/3	6/9	8/9
2017/18	Top 25%	7/7	11/8	13/8
	Top 50%	5/7	18/12	11/11
	Top 10%	4/6	15/10	11/11
2018/19	Top 25%	6/3	13/18	13/10
	Top 50%	9/6	18/11	20/11

Table 4: Most common attributes per position for filter and wrapper methods.

G-filter	D-filter	F-filter	G-wrapper	D-wrapper	F-wrapper
QS(11)	PS(12)	PS(12)	SO(9)	PS(9)	PS(11)
W(10)	TOI(EV)(12)	PPA(12)	W(9)	TOI(EV) (8)	PPA(10)
GPS(8)	PPA(12)	TOI(EV)(12)	QS(8)	PPA(7)	TOI(EV)(9)
GP(6)	EVA(9)	SHFT(11)	GPS(6)	oiSH%(7)	EVTOI(9)
SO(6)	S(9)	EVTOI,	SA,	SHG, EVTOI,	PPTOI(9)
		PPTOI(9)	GS(5)	GA/60, PPTOI(7)	

and TOI(EV) appeared in all tears for several seasons. For the top 10% tier GA/60 was important for all seasons for the wrapper method, while PPA was important for the top 25% tier. For forwards PS and PPA were important for the filter and wrapper methods and TOI(EV) for the filter method. SHFT was an important attribute for the filter method for forwards, but not so much for defenders. In general, S is more common for top 50% tier players, while PPA is most common for top 25% tier players. Interestingly, PPA is selected more often than EVA. Further, in contrast to the wrapper method, for the filter method it is more common that attributes for a particular tier are selected in different seasons. Season 2017/18 was different in two senses. First, more attributes were selected for defenders and forwards than for the other seasons. Secondly, PPTOI and EVTOI were often selected in other seasons, but not in 2017/18.

We note that many of the selected attributes for field players are measures related to offense (e.g., related to assists, goals and shots) or neutral (e.g., related to time on ice), but the most often occurring measure (PS) relates to both offense and defense. For defenders, there are additionally measures related to goals against. This may reflect the kinds of stats that are collected for players.

In the data preparation step we removed attributes that are combinations of other attributes and these included much used metrics (e.g., goals and assists), which hockey professionals would want to use. Therefore, we investigated whether these metrics 'appeared' in the results, meaning that the attributes on

which they depend were selected. For goalkeepers QS% (combination of QS and GS) appeared often, while SV and SV% (combinations of SA and GA) appeared in filter data sets. We also note that whenever GA occurred, also GA% can be computed. Regarding field players, A (combination of EVA, PPA and SHA) appeared sometimes, but the interesting combination of EVA and PPA (which does not take into account boxplay) occurred often. To a lesser extent the same happened for G (combination of EVG, PPG and SHG). S% (combination of EVG, PPG, SHG and S) did not occur, but combinations of EVG, PPG and S, or EVG and S did. Further, also CF% (combination of CA and CF), FF% (combination of FA and FF), FO% (combination of FOW and FOL) and PDO (combination of oiSH% amd oiSV%) appeared in some data sets. For more information about the exact numbers of occurrences, see [8].

5 Prediction

5.1 Methods

For each data set that was used in the feature selection step, we then created two new data sets, one where we used the attributes selected by the filter method and one with the attributes selected by the wrapper method. For the top 10% and top 25% tier data sets we used SMOTE [1] to overcome the class imbalance. This oversampling technique synthetically determines copies of the instances of the minority class to be added to the data set to match the quantity of instances of the majority class.

5.2 Results and discussion

A detailed performance of all algorithms on all data sets is given in [8]. Fig. 1 shows specificity, AUC, F1, sensitivity and accuracy for different seasons, positions, tiers, filter/wrapper and machine learning algorithms. The largest variation among the measures was for F1. Fig. 2 shows F1 for different positions and tiers with respect to season, filter/wrapper and machine learning algorithm.

Overall, the choice between the filter and wrapper methods for different measures is not that important (Fig. 1), although for particular tiers and positions there may be a difference (e.g., goalkeepers top 10% and 25%, Fig. 2).

When comparing the resulting models, the two Bayesian classifiers were top performers for most data sets and evaluation measures and performed evenly across all combinations of comparisons. This is in line with the study in [14] regarding football. The tree-based models had the highest specificity, but had a lower sensitivity. They seemed to prioritize the majority class which resulted in lower performance when classifying the top 10% tier, and especially for the smaller data sets (e.g., goalkeepers). Overall, the models achieved high sensitivity, although for small data sets the tree-based models did not do well. In general, the models were best at classifying forwards, highlighting that many of the official metrics are focused on the offensive measures. This suggests that
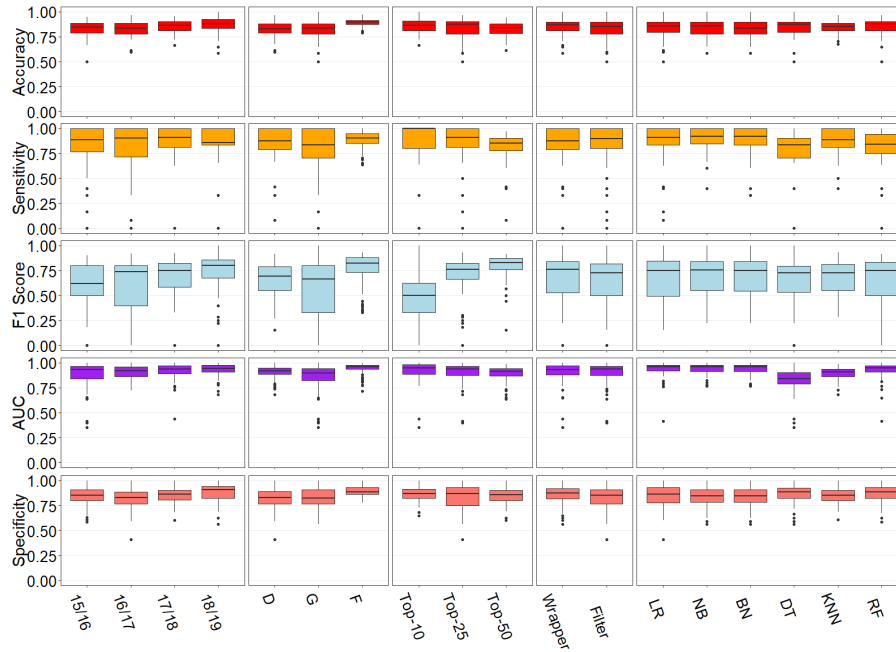
Fig. 1. Specificity, AUC, F1, sensitivity and accuracy for different seasons, positions, tiers, filter/wrapper and machine learning algorithms.

more work is needed to develop equally good defensive metrics. The models also achieved higher F1 for the top 50% highlighting that it is harder to differentiate between the highest rank top tier players using official performance metrics alone.

There is variation over the seasons, reflecting, among others, that different attributes were selected for different seasons.

A closer look at the misclassified players explains why the above problems are so hard. For example, of the top 10% forwards of the 2018/19 season, 19 players were misclassified by at least one out of 12 (2×6) combined models and the best model (BN) misclassified 8 players with the filter method and 10 with the wrapper method. However, some of these players either had weaker than normal years and therefore may have been classified lower than they normally would have by some models (e.g., Taylor Hall 4/12 wrong, Gabriel Landeskog 2/12 wrong, Joe Pavelski 1/12 wrong, and Auston Matthews 1/12 wrong, Patrice Bergeron 1/12 wrong), was a Rookie (Elias Pettersson 1/12 wrong) or were players outside the top 10% tier that were classified into this top tier at least once. For the first set we note that the most frequent player that should be in the set but sometimes is classified outside is Taylor Hall. He is a former Hart Trophy (league MVP) winner (2017/18 season) that had an injury plagued 2018/19 were he only played 33 out of 82 games. Similarly, the misclassification of the two most frequently
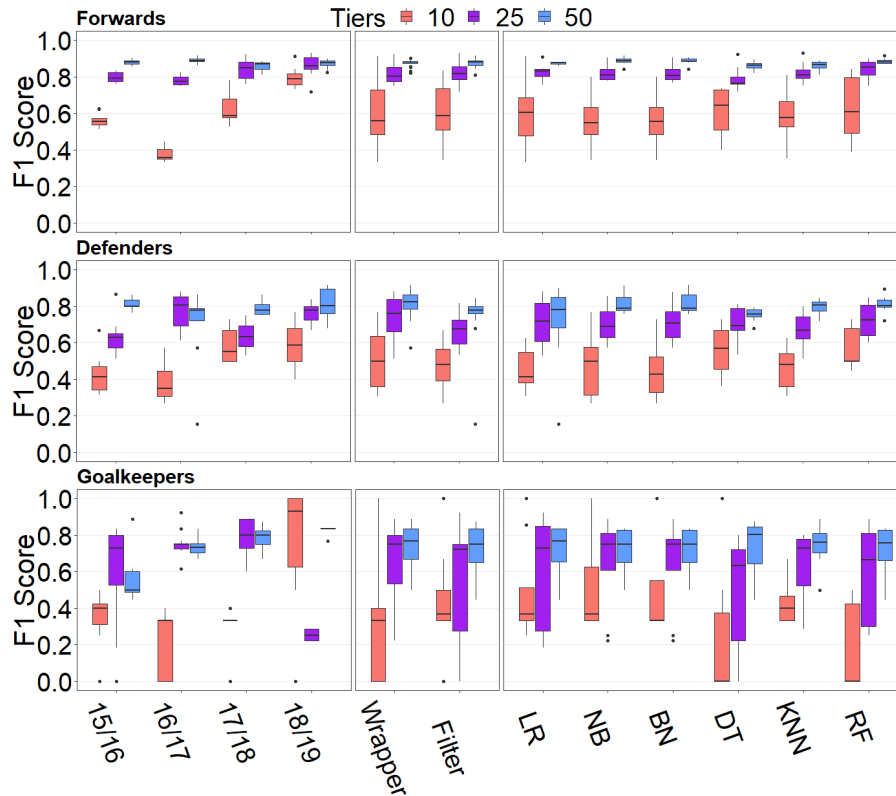
Fig. 2. F1 for different positions and tiers with respect to season, filter/wrapper and machine learning algorithm.

misclassified players of the last set can also be explained. Teuvo Teravainen is an upcoming star who ranked 29th in the scoring race when the 2019/20 season shut down for a covid-19 break, and Evgenii Dadonov had a career year (scoring 72 points 2018/19) playing on a line with Aleksander Barkov and Jonathan Huberdeau (which both finished with over 90 points). The lists of misclassified players for all data sets are given in [8].

A limitation of the study is that for the algorithms with many parameters, we did not perform experiments to find the optimal parameter setting, but usually used the default values. An area for future work is, therefore, to experiment with optimal settings as well as other algorithms. Further, there are some choices in the experiments that may have an influence on the results. For instance, the choice of the number of occurrences in the feature selection step influences which attributes to retain and thus the data sets on which the machine learning algorithms are evaluated. It would be interesting to investigate these choices in a systematic way. Another track for future work is to use player performance methods for ranking instead of the EA player rating and to compare the results of the different methods.

6 Conclusion

In this paper we used 6 different machine learning methods (Logistic regression, k-Nearest neighbour, Decision tree, Random forest, Naïve Bayes and Bayesian network) and 2 different feature selection methods (filter and wrapper) to predict players' ranking tier (top 10%, 25% and 50%) for 3 player positions (forwards, defenders, and goalkeepers), looking at 4 seasons (2015/16 - 2018/19). The study highlights key performance metrics for the different player categories and provides insights into the difference in the complexity of identifying the key attributes and skills that may help predict the success of individual players.

When comparing the resulting models, the two Bayesian classifiers performed best and had the best sensitivity. The tree-based models had the highest specificity, but had trouble classifying the top 10% tier players. In general, the models were best at classifying forwards, highlighting that many of the official metrics are focused on the offensive measures. The development of equally good defensive metrics still remains an open problem.

References

1. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research **16**, 321–357 (2002), doi: 10.1613/jair.953
2. Dougherty, J., Kohavi, R., Sahami, M.: Supervised and unsupervised discretization of continuous features. In: Prieditis, A., Russell, S.J. (eds.) Machine Learning, Proceedings of the Twelfth International Conference on Machine Learning. pp. 194–202 (1995), doi: 10.1016/b978-1-55860-377-6.50032-3
3. Gramacy, R.B., Jensen, S.T., Taddy, M.: Estimating player contribution in hockey with regularized logistic regression. Journal of Quantitative Analysis in Sports **9**, 97–111 (2013), doi: 10.1515/jqas-2012-0001
4. Gu, W., Foster, K., Shang, J., Wei, L.: A game-predicting expert system using big data and machine learning. Expert Systems with Applications **130**, 293–305 (2019), doi: 10.1016/j.eswa.2019.04.025
5. Hall, M.: Correlation-based Feature Selection for Machine Learning. Ph.D. thesis, The University of Waikato, New Zealand (1999)
6. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA Data Mining Software: An Update. SIGKDD Explorations **11**(1), 10–18 (2009), doi: 10.1145/1656274.1656278
7. Kaplan, E.H., Mongeon, K., Ryan, J.T.: A Markov Model for Hockey: Manpower Differential and Win Probability Added. INFOR Information Systems and Operational Research **52**(2), 39–50 (2014), doi: 10.3138/infor.52.2.39
8. Lehmus Persson, T., Kozlica, H., Carlsson, N., Lambrix, P.: Prediction of tiers in the ranking of ice hockey players - extended version (2020), `https://www.ida.liu.se/~patla00/publications/mlsa2020-hockey-extended.pdf`
9. Liu, G., Schulte, O.: Deep reinforcement learning in ice hockey for context-aware player evaluation. In: Lang, J. (ed.) Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. pp. 3442–3448 (2018), doi: 10.24963/ijcai.2018/478

10. Liu, Y., Schulte, O., Li, C.: Model Trees for Identifying Exceptional Players in the NHL and NBA Drafts. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) Machine Learning and Data Mining for Sports Analytics. MLSA 2018. Lecture Notes in Computer Science, vol. 11330, pp. 93–105 (2019), doi: 10.1007/978-3-030-17274-9_8

11. Ljung, D., Carlsson, N., Lambrix, P.: Player pairs valuation in ice hockey. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) Machine Learning and Data Mining for Sports Analytics. MLSA 2018. Lecture Notes in Computer Science, vol. 11330, pp. 82–92 (2019), doi: 10.1007/978-3-030-17274-9_7

12. Macdonald, B.: A Regression-Based Adjusted Plus-Minus Statistic for NHL Players. Journal of Quantitative Analysis in Sports **7**(3) (2011), doi: 10.2202/1559-0410.1284

13. Macdonald, B.: An Improved Adjusted Plus-Minus Statistic for NHL Players. In: MIT Sloan Sports Analytics Conference (2011)

14. Nsolo, E., Carlsson, N., Lambrix, P.: Player valuation in European football. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) Machine Learning and Data Mining for Sports Analytics. MLSA 2018. Lecture Notes in Computer Science, vol. 11330, pp. 42–54 (2019), doi: 10.1007/978-3-030-17274-9_4

15. Pettigrew, S.: Assessing the offensive productivity of NHL players using in-game win probabilities. In: MIT Sloan Sports Analytics Conference (2015)

16. Routley, K., Schulte, O.: A Markov Game Model for Valuing Player Actions in Ice Hockey. In: Meila, M., Heskes, T. (eds.) Uncertainty in Artificial Intelligence. pp. 782–791 (2015)

17. Sans Fuentes, C., Carlsson, N., Lambrix, P.: Player impact measures for scoring in ice hockey. In: Karlis, D., Ntzoufras, I., Drikos, S. (eds.) MathSport International 2019 Conference. pp. 307–317 (2019)

18. Schuckers, M.: Draft by Numbers: Using Data and Analytics to Improve National Hockey League (NHL) Player Selection. In: MIT Sloan Sports Analytics Conference (2016)

19. Schuckers, M., Curro, J.: Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events. In: MIT Sloan Sports Analytics Conference (2013)

20. Schulte, O., Khademi, M., Gholami, S., Zhao, Z., Javan, M., Desaulniers, P.: A Markov Game model for valuing actions, locations, and team performance in ice hockey. Data Mining and Knowledge Discovery **31**(6), 1735–1757 (2017), doi: 10.1007/s10618-017-0496-z

21. Schulte, O., Zhao, Z., Javan, M., Desaulniers, P.: Apples-to-apples: Clustering and Ranking NHL Players Using Location Information and Scoring Impact. In: MIT Sloan Sports Analytics Conference (2017)

22. Thomas, A., Ventura, S.L., Jensen, S., Ma, S.: Competing Process Hazard Function Models for Player Ratings in Ice Hockey. The Annals of Applied Statistics **7**(3), 1497–1524 (2013), doi: 10.1214/13-AOAS646