Temporal Activity Outcome Prediction of Players in Ice Hockey

Chayan Shrang Raj

Linköping University, SE-58183, Linköping, Sweden

Abstract. Temporal activity prediction is a challenging task in sports especially in sports like ice hockey because of its fast dynamics and interactions among the players. In this paper, we are trying to predict whether the next action of the players will be successful or not given the various parameters describing its location and other factors. Basically, it is the process of predicting if the player's next action/move would be a success or not. Ice hockey has a large number of rules and techniques that this approach is not easy to understand neither to implement but latest advancements in machine learning has given us the opportunity to learn complex findings/patterns in the data. Choosing the kind of output is in itself a big task as it requires a good sports knowledge to grasp the nuances of the game and hence effectively produce meaningful results. Given a sequence of time stamps and player's action outcome at each time stamp, we are leveraging the power of deep fully connected neural networks with residual connections to understand the given historical data at hand [1]. The proposed model learns better context information from the residual/skip connections in the network along with using sophisticated hyper parametrization optimization process.

Keywords: Residual, Neural Network, Ice Hockey, fully connected, Optimization

1 Introduction

Ice hockey is one of the most popular sports in the world, particularly in US, Canada, Scandinavia and much of Europe. Thanks to the digital inclusion in this field that today we can collect a large amount of data of each match played and leverage it for better team planning and strategies. Sports activity prediction of each individual players is a complex task due to rapid movements of players in the field, large number of rules, occlusions, rapid change in play environment and more. Individual players add their value in winning as a team but better game plan for each player makes the game even more exciting and strategically robust. In this paper, we are focusing on analyzing and finding patterns of player's next action that leads up to certain events such as success or failure of pass, loss puck recovery (lpr) or shot and much more. The inputs to our Machine Learning model are timestamps of match played between 20 teams in one season. Time stamps include every activity like a shot, team possession, offside, score differential, etc. And the output is the probability of an action being successful, action can a shot, lpr, pass, etc. We can then use these insights learned from our model to predict and create better strategies for upcoming matches. This technique could be very useful for teams to understand their gameplay and make improvements in their strategy for becoming match winners [2]. Since, we are talking about temporal activities, we have full view of previous and present actions taken by players and the results that lead up to certain events that later decides the fate of the match in future. It could also allow players to change/enhance their tactics in field to maximize success probability during the time in each period. Machine learning is being used more and more in sports analytics area because it gives stakeholders the power to make better decisions and freedom of choosing of motely options. We are attempting to create a specialized profiling of teams by analyzing matches played with different teams in a tournament. The research in this area has been sparse where player activity prediction is in question, but other sports have conducted analysis where homographic projections of temporal and spatial snapshots were fed into convolutional neural networks that produced promising results [3].

2 Background

Many of the models for evaluating player activity in ice hockey define a particular statistics or evaluation metric that gives values based on types of actions in the game [4]. Some background knowledge is required in neural networks with its types of layers and non-linearity of inputs along with different types of loss functions. Sufficient familiarity with Ice hockey rules is required to assess the type of objective function to use and to select good features.

3 Algorithms

First step in applying any machine learning algorithm is to have representative data of real life and also must be clean (devoid of any noise that does not represent real life situation). For this paper, the data was provided by Sportlogiq with permission of SHL, the Swedish Hockey League, representing event data from the 2020-2021 SHL season. It consists of 76041 rows and 22 features describing each game with a unique game id and different time stamps. Firstly, an exhaustive exploratory data analysis was performed using one of the 'gameid' e.g. 66445. It consists of match time stamps between two teams encoded as 742 and 916, where every time stamp describes the event played by any one of the players in one of teams from a certain point on the field and whether it was successful or not. Keeping in mind the structure of the data, I have used a different idea of separating training, evaluation and test sets as different matches uniquely defined by their 'gameid'. It means, if one match is used as a validation data and the other is used as testing data, then all the other matches are used as training data. This allows our model to analyze whole data space and extract complex patterns in each match. Looking at the data at hand, I have come up with an idea of choosing the best describing train/val/test sets because it would be a biased analysis if we only took one match for all processes. We will be using 'outcome' feature as our class label to learn and predict, that demonstrates if a particular event was successful or not. To work on this, I have first trained an ensemble of four Residual neural networks that will loop over all the 'gameid' with a stride of 2. So, first ensemble works on 'gameid' 0 and 1, then next on 2 and 3 and so on for test and validation sets respectively. We then average the accuracy of the 4 NN ensemble models and choose the games that produced highest accuracy. This is done because neural networks are powerful algorithms and can learn complex patterns with enough data, so we are feeding our data first to this ensemble to understand the best explainable data blocks for our next ensemble step [5]. In figure 1, these are 'gameid' 6 and 7 produced highest accuracy and are chosen for test and validation sets respectively.

```
{66445: ((742, 916), 0),
 83522: ((907, 824), 1),
 73649: ((916, 915), 2),
 63799: ((916, 650), 3),
 88237: ((907, 564), 4),
 60432: ((915, 729), 5),
 86583: ((787, 579), 6),
 71102: ((824, 583), 7),
 81893: ((787, 771), 8),
 84953: ((807, 915), 9),
 80711: ((729, 650), 10),
 89409: ((896, 771), 11),
 77265: ((824, 771), 12),
 73282: ((807, 896), 13),
 78204: ((583, 579), 14),
 87892: ((787, 583), 15),
 75425: ((564, 729), 16),
 78500: ((579, 807), 17),
 87080: ((896, 742), 18),
 65884: ((742, 907), 19)}
```

Fig. 1. Encoded dictionary.
Format: *Game Id: ((Team, Opposing Team), Game Id encoding).*

Residual networks have the capability to build deep neural networks and are used in this network to better understand the previous knowledge the model has already learned. It is created by skipping one connection from fully connected layer and then create an identity connection to the next layer before the non-linearity [6]. One residual block is shown in figure 2. We have used these after every second layer to carry forward the information.

Ensemble learning often proves to be performing superior to any one machine learning algorithm and hence final model/algorithms chosen for this paper is again an ensemble of 4 very powerful classification algorithms, namely, K-Nearest

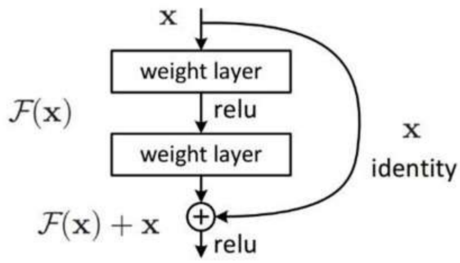

Fig. 2. Residual block

Neighbors, Logistic Regression, Random Forests, Support Vector Machines. The structure of our setup is shown in figure 3.

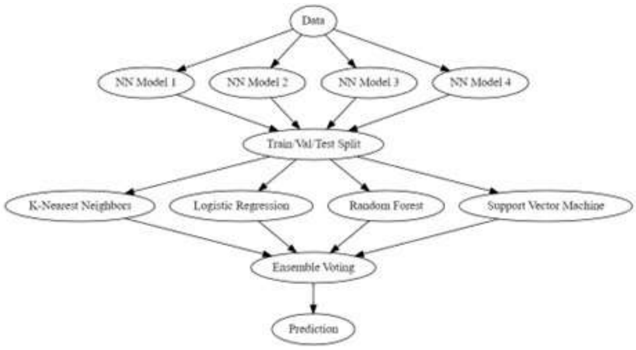

Fig. 3. Ensemble Graph

Feeding the data as block of matches gives a closed loop information to the network to learn from that specific game and environment. This new ensemble block creates a whole new space of hyper parameters for each individual algorithms and optimizing them is key to generalizing well on unseen data. For this purpose, we have used an open source hyperparameter optimization framework to automate hyperparameter search called 'Optuna' which has parallel and GPU support [7]. We have also leveraged acceleration for scikit-learn implementation using Intel® Extension for Scikit-learn called sklearnex. This reduced the training and optimization time by 30%. Post optimizing the hyper parameter space, we evaluated our model on test data and achieved an accuracy of 87% which is circa 3% higher than simply using residual neural networks. Also, it was our attempt to understand and make our analysis more interpretable, and so we trained SHAP (SHapley Additive exPlanations) which is a game theoretic approach to explain the output of any machine learning model [8]. The optimized hyper parameter space and feature explanation is shown in figure 4.

```
{'lr_penalty': 'elasticnet',
 'lr_l1_ratio': 0.2112744795805186,
 'lr_tol': 0.0007835048639900213,
 'lr_C': 0.623510727584215,
 'knn_neighbors': 26,
 'knn_weights': 'uniform',
 'knn_p': 2,
 'svm_C': 0.6301532188169567,
 'svm_kernel': 'rbf',
 'svm_tol': 0.009091792023585859,
 'rf_estimators': 299,
 'rf_criterion': 'gini',
 'rf_max_depth': 30,
 'rf_min_samples_split': 10,
 'rf_min_samples_leaf': 1,
 'lr_w': 0.10913043839614027,
 'knn_w': 0.4166692113067154,
 'svm_w': 0.025464590107005435,
 'rf_w': 0.6958783869393963}
```

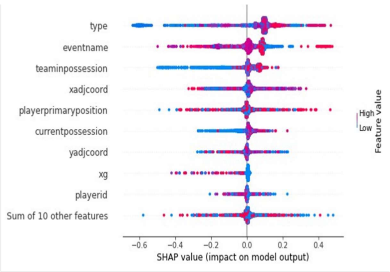

Fig. 4. Optimized Hyper Parameter Space / Impact of each feature on model output

4 Discussion

The evaluation of the models is done with the evaluation metrics accuracy, precision, recall and F1-score. In our case, we have binary classification where accuracy shows the total correct classification out of total values, precision and recall capture the limitations of accuracy and considers the curse of imbalanced class labels. Finally, F1-score encapsulates and keeps the the limitations of accuracy at bay and provides the best metric for our classifier which ranges from 0 to 1, higher the value, better the predictions. Figure 5 presents the result of our analyses.

	precision	recall	f1-score	support
class 0	0.68	0.59	0.63	779
class 1	0.89	0.92	0.91	2865
accuracy			0.85	3644
macro avg	0.78	0.76	0.77	3644
weighted avg	0.85	0.85	0.85	3644

Fig. 5. NN precision, recall and F1-Score

There have been several works in this area that try to quantify pre-match result prediction or live match prediction of the team. Temporal data is also leveraged to augment match outcome prediction based on historical data with the dawn of specialized implementations of Bidirectional LSTMs and GRU in highly efficient deep learning frameworks.

5 Summary and Future Work

This report aims to homogenize advanced machine learning algorithms and historical ice hockey data to express novel ideas in the field of sports analytics which is a burgeoning area for applied data analytics and artificial intelligence. The key point to perceive from this project is to better understand individual player strategy and game plan of a team against their opponent and to tap into opposing team gameplay. If a player's action is either successful or failure on the field, given a set of circumstances, this project tries to analyze/asses the cause of that event and would help stakeholders to pre plan or change their strategy driving attention to every minute detail leading to that event.

There is a scope of improvement in implementing this idea in future where more historical data could certainly help the algorithm to generalize even better. Data of 4 or 5 seasons would bolster the implementation even more. Expert Ice Hockey knowledge could be utilized in feature engineering to create impactful features for ravenous deep learning algorithms that are hungry for good features. There is also a possibility of applying sequence deep learning models and deep reinforcement learning where we could visualize a team's performance against an agent that learns using a feedback loop. Also, this idea could be used to evaluate individual player's weaknesses and strengths. It is also possible to augment the data with different hypothetical events and then capture the model's performance, this could allow the team to focus on successful outcomes.

Github Link

https://github.com/chayansraj/LINHAC-2022-Student-Competition

References

1. Marek, Patrice, Šedivá, Blanka and Ťoupal, Tomáš. "Modeling and prediction of ice hockey match results" Journal of Quantitative Analysis in Sports, vol. 10, no. 3, 2014, pp. 357-365. https://doi.org/10.1515/jqas-2013-0129

2. Tora, Moumita Chen, Jianhui Little, J.J.. (2017). Classification of Puck Possession Events in Ice Hockey. Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

3. 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). https://doi.org/10.24963/ijcai.2018/478

4. Lehmus Persson, T., Kozlica, H., Carlsson, N., Lambrix, P. (2020). Prediction of Tiers in the Ranking of Ice Hockey Players. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds) Machine Learning and Data Mining for Sports Analytics. MLSA 2020. Communications in Computer and Information Science, vol 1324. Springer, Cham. https://doi.org/10.1007/978-3030-64912-8_8

5. Pischedda, Gianni. (2014). Predicting NHL Match Outcomes with ML Models. International Journal of Computer Applications. 101. 15-22. https://doi.org/10.5120/17714-8249

6. Li, Fahong. Description, analysis and prediction of player actions in selected hockey game situations. Diss. University of British Columbia, 2004.

7. Akiba, T., Sano, S., Yanase, T., Ohta, T. and Koyama, M., 2019, July. Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 2623-2631).

8. Lundberg, S.M., Nair, B., Vavilala, M.S. et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. Nat Biomed Eng 2, 749–760 (2018). https://doi.org/10.1038/s41551-018-0304-0