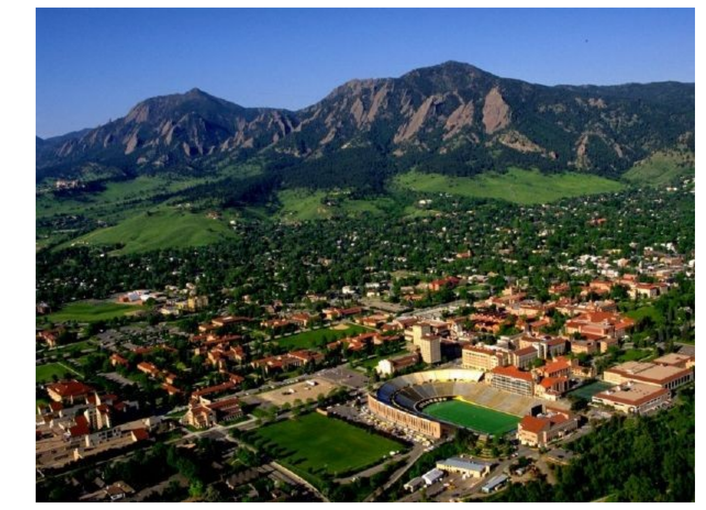
The Lord Stanley Seekers

University of Colorado Boulder

Kevin Crooks, Brian Lim, Arthur May

CSCI 5622 – Spring 2020, Professor Quigley



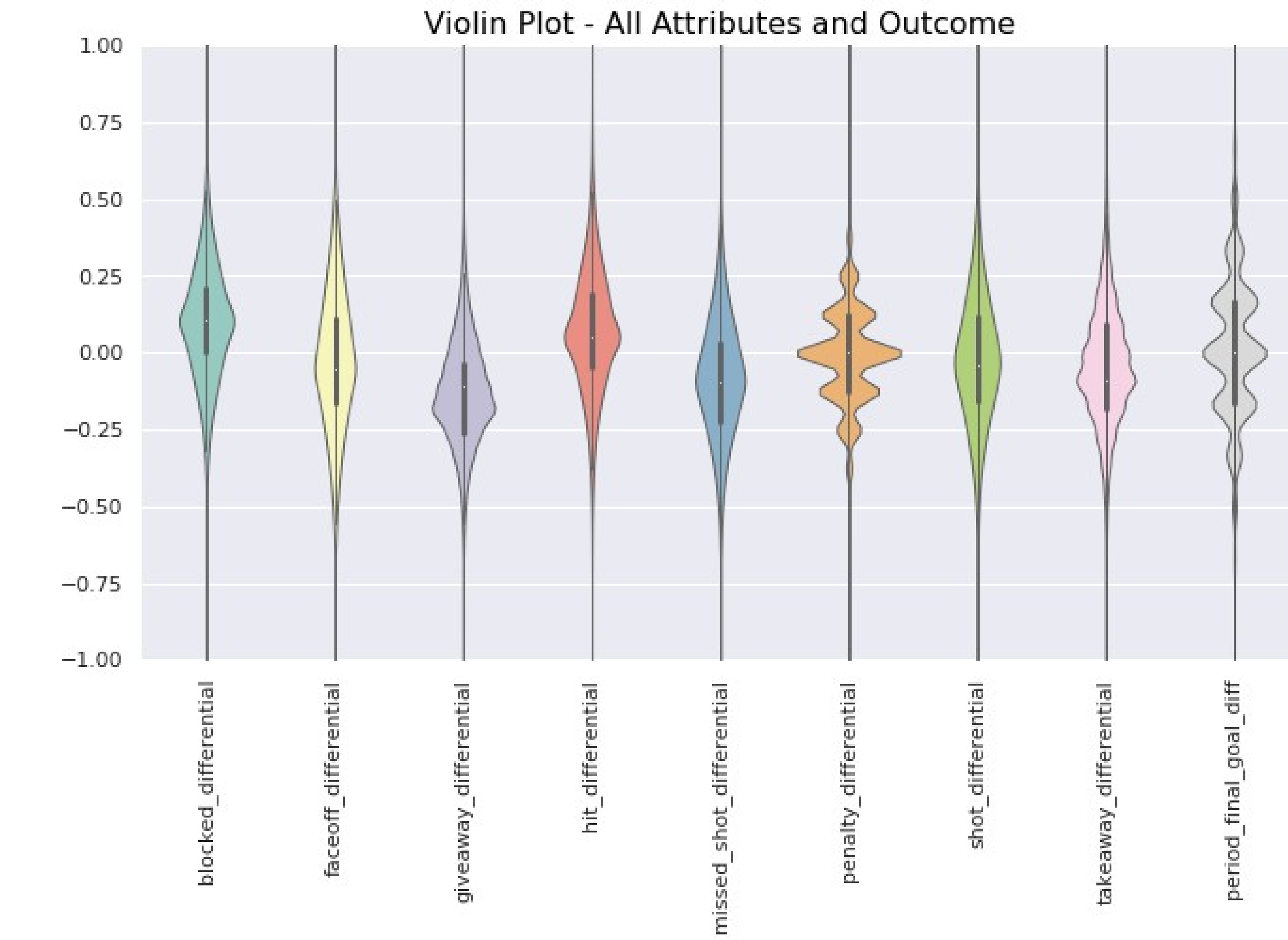
Introduction

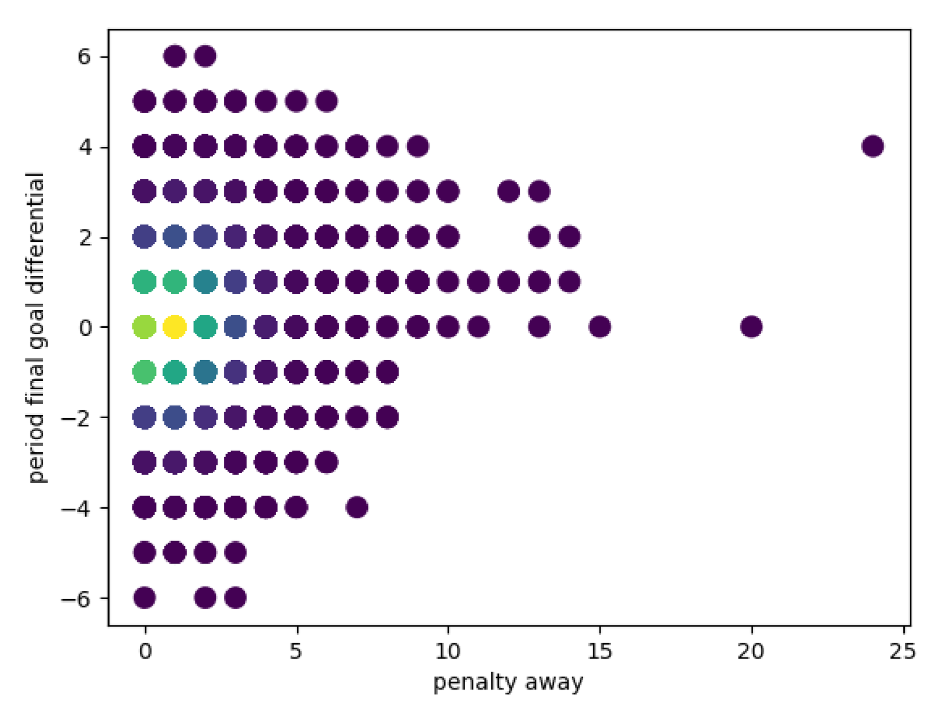
Most of the existing research in using Machine learning to predict outcomes in hockey games was focused on predicting the win or loss outcome for a given team. Rather than retreaded an old problem space, which would simply amount to wondering if the science has improved in the past few years, or training algorithms have, we decided to look at a new problem: period prediction. An NHL game is divided into 3 periods of twenty minutes each. If the score is still tied after these three regulation periods, a 4th overtime period is played with the current standard being a single 5-minute period. We included these in our analyses as well. Despite the fact that statistics will be different between a full 20-minute period and a shorter 5-minute period, we are looking at predictors for each period individually, so our overall models will not be affected by this bundling.

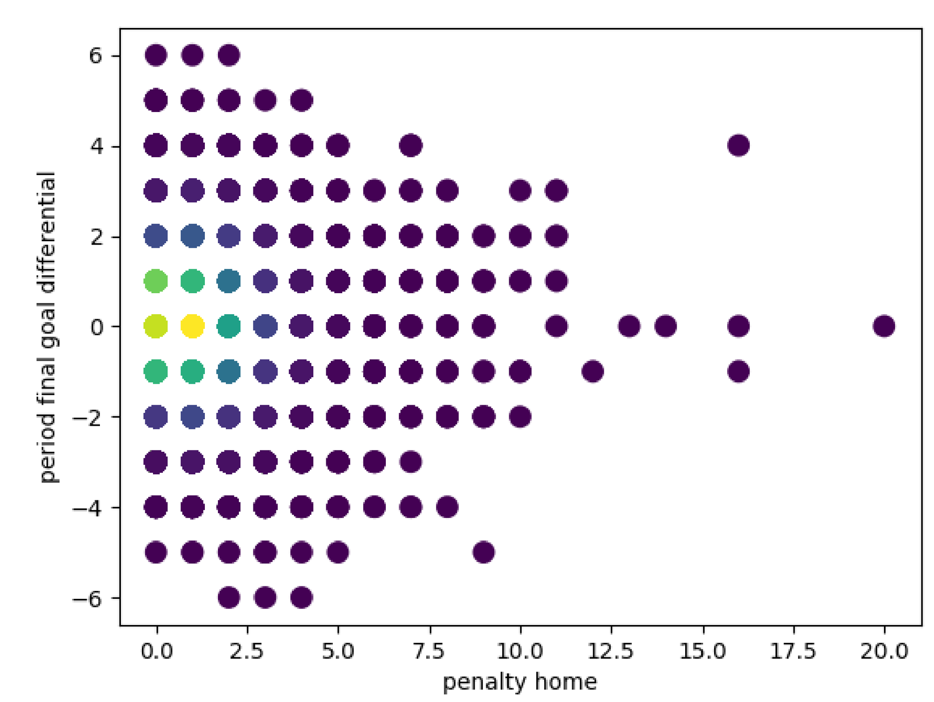
The advantage to such a prediction might be immediately obvious to a coach, although probably not as useful for Vegas. If your team enters the second period trailing by 2, how do you overcome that lead? Our goal is to see if we can find which statistics, taken from the standard metrics used in NHL statkeeping, can be beneficial in predicting the outcome of a period.

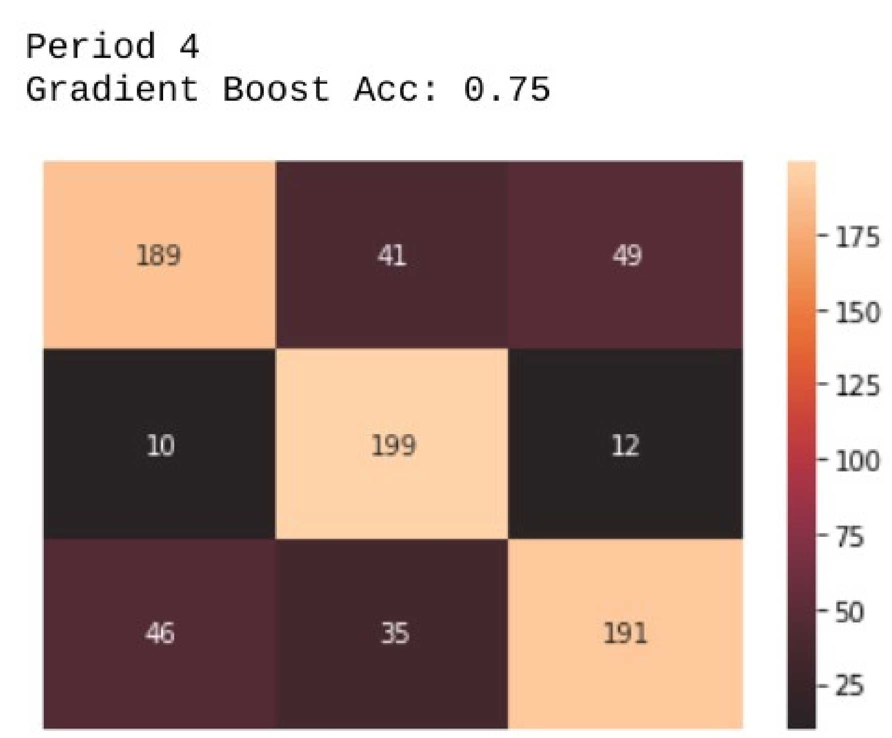
Methods

Our approach to address the problem iterated through three steps over multiple phases. The first step was to prepare the data. Once the data was prepared, we performed a number of visualizations around the data to get an understanding of relationships between features and the end of period goal differential. Once we had an understanding of the relationships between the features, we ran them through different machine learning classifiers based on the results we were seeking. We did not seek a binary prediction, but a goal differential which is a bit more challenging. Furthermore, as we settled on our models we experimented with different normalization approaches and ran a grid search on our classifier to help fine tune the results. Several classifiers will be tested including a gradient boost, random forest, and KNN. The gradient boosting classifier performed the best, and we used that to perform the search for which features were the most important. When our results were output we output confusion matrices to help us interpret how our model was performing relative to the other potential outcomes of the model. We performed forward stepwise selection to determine the best subset of features to be used to predict the outcomes of each period, and used the best subset per period to train and predict the models for the individual periods.

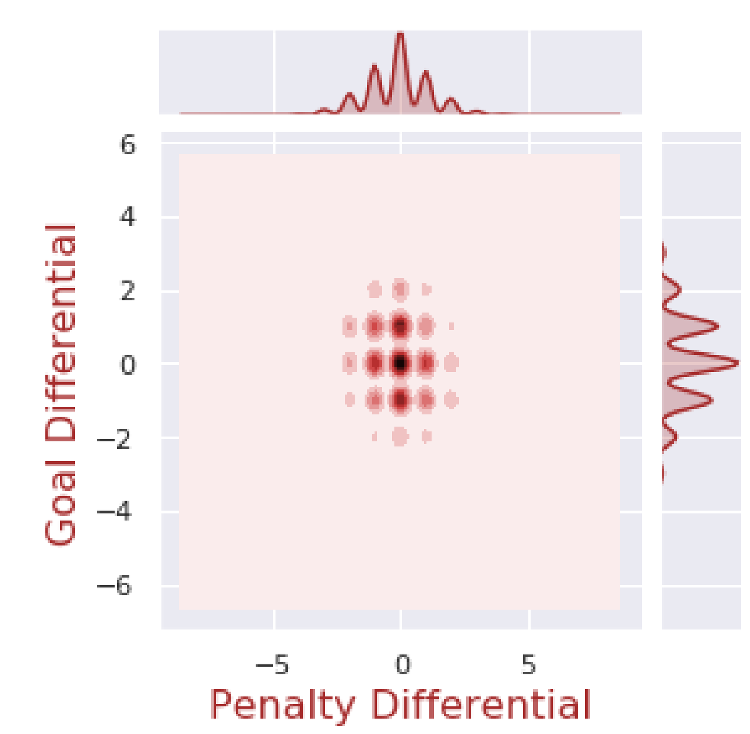








Overtime Period Confusion Matrix



Results

Our Gradient Boost classifier performed the best of the models that we explored, however, it still did not fully reach our desired prediction accuracy of 60% per period. We were able to achieve accuracies between 35-45% for periods 1 – 3, with the models getting more accurate as the games progressed. Interestingly, the model was capable of predicting the goal differentials of overtime periods with a ~75% accuracy. To achieve these results, the outcomes were binned to fit in {-1,0,1} goal differential, and we used a grid search to optimize the parameters of the model. A forward step-wise feature selection was performed for each period to determine which features were most important in predicting the outcomes of the period.

Discussion

The ability to predict the outcome for overtime is considered a significant contribution. While predicting the goal differential for each period was perhaps to difficult for the scope of this research effort, there are a number of reasons why the overtime period was as accurate as it was. First, there less classes of goal differential with a sudden death overtime format. When analyzing the features that were used in the models it was interesting to see how some features changed, while others were important across all the periods. For example, penalty differential was important in all the periods, while the number of hits was most important in the first and overtime periods, or the blocked shot differential only proved important in the third period. In terms of predicting goal differential, we may have had better results if we attempted to calculate a goal differential and allowed non-integer values to be a result, and while goals are only counted in integers, a fractional goal differential extrapolated over an entire season could have a significant impact on a teams placement in the standings.

References

Weissbock, J., Viktor, H., Inkpen, D.: Use of performance metrics to forecast success in the national hockey league. European Conference on Machine Learning: Sports Analytics and Machine Learning Workshop (2013) http://ceur-ws.org/Vol-1969/paper-06.pdf

Pischedda, Gianni: Predicting NHL Match Outcomes with ML Models. In: International Journal of Computer Applications (0975 – 8887), Volume 101– No.9, September 2014. (2014) http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.735.795&rep=rep1&type=pdf

Weissbock, Josh; Inkpen, Diana: Combining Textual Pre-game Reports and Statistical Data for Predicting Success in the National Hockey League. In: Advances in Artificial Intelligence 27th Canadian Conference on Artificial Intelligence, Canadian AI 2014, Montréal, QC, Canada, May 6-9, 2014. (2014)