PROJECT PART 5: CHECKPOINT AND BREAK DOWN OF CONTRIBUTION

John Zhang (jz5jx)
School of Data Science
University of Virginia
Charlottesville, VA 22903
jz5jx@virginia.edu

Youssef Abubaker(yaa2vb)
School of Data Science
University of Virginia
Charlottesville, VA 22903
yaa2vb@virginia.edu

Joseph Cho (jc4jx)
School of Data Science
University of Virginia
Charlottesville, VA 22903
jc4jx@virginia.edu

April 12, 2021

1 Abstract

We attempt to use machine learning techniques to predict whether a shot attempt is successful or not based on a players release. SportVu provides us with spacial and temporal data of 20,772 3-point shot attempts. To set a baseline target for accuracy, we will compare our results to a logit model that predicts the shot outcome based on the final data point of the ball before it hits the rim (breaks the plane of the rim with respect to the z-axis). According to Sports Science, the average NBA player releases a shot in approximately 0.54 seconds. Following this statistic we utilized the first 8 frames of every shot as this would encapsulate the release points for every player. The first frame captures where the shot was released from and each frame after that captures the shot trajectory to the basket. The objective of this paper is to compare a Neural Net model with release data points to a logit model trained upon the final frame of the ball, these final frames are suspected to have a higher predictive performance as they are more indicative of shot outcome.

2 Motivation

The motivation for this project is exploring the applications of deep learning in the NBA. Currently the public domain of basketball research utilizes predominantly linear based models and regression techniques on play by play data. Therefore, to innovate from the existing architecture of this problem we attempt to utilize an ANN to explore shot outcomes of three pointers from shot release data.

3 Method

We developed two different models based on different perspectives of the movement of the ball. For our log-it model we utilized the coordinate positions from the last frame of each individual shot to predict shot outcome. The second model we evaluated was a custom neural net model. For the neural net model we started with a standard Keras Sequential Model. Then we added a Dense layer with 12 nodes and passed each input through "Relu" Activation Functions. From this first layer we then added a second layer with 8 nodes and another "Relu" Activation function. From the subsequent layer we then passed all the inputs through the output layer which mapped our matrix manipulations through a sigmoid activation function which predicted our inputs to either a "1" (Success) or "0" (Failure).

4 Preliminary Experiments

We implemented our logit model in python using the LogisticRegression() function with cross validation from the sklearn library. For cross validation we used a .67/.33 train-test split and observed 64.9% accuracy. For our Deep Learning model, we used the Keras framework to build a Neural Network of 3 dense layers. This model was trained on the "binary_crossentrophy" loss function and optimized using the Adam algorithm. Training consisted of 10 epochs with a batch side of 10 which yielded a accuracy of 64.1%.

5 Next Steps

Now that we have established a target from our logit model, we plan to implement other Deep Learning architectures in an effort to improve of our performance. We expect the implementation of a Convolutional Neural Network (CNN) to improve our accuracy as the temporal and spatial nature of this problem are well suited for this architecture.

6 Member Contribution

Each member helped with data wrangling. Additionally, Youssef was responsible for the explanatory data analysis. Joe helped with implementing our Logistic Regression Model. John was responsible for implementing the first iteration of the custom Keras Model. Every team member has been responsible for writing the papers and keeping in good communication and doing what's best for the team.

References

[1] Shah, Rajiv, and Rob Romijnders. "Applying deep learning to basketball trajectories." arXiv preprint arXiv:1608.03793 (2016).

[2] Wang, Kuan-Chieh, and Richard Zemel. "Classifying NBA offensive plays using neural networks." Proceedings of MIT Sloan Sports Analytics Conference. Vol. 4. 2016.

[3] Thabtah, F., Zhang, L. Abdelhamid, N. NBA Game Result Prediction Using Feature Analysis and Machine Learning. Ann. Data. Sci. 6, 103–116 (2019). https://doi.org/10.1007/s40745-018-00189-x

[4] Cheng G, Zhang Z, Kyebambe MN, Kimbugwe N. Predicting the Outcome of NBA Playoffs Based on the Maximum Entropy Principle. Entropy. 2016; 18(12):450. https://doi.org/10.3390/e18120450

[5] Fukushima, Takuya, Tomoharu Nakashima, and Hidehisa Akiyama. "Similarity analysis of action trajectories based on kick distributions." Robot World Cup. Springer, Cham, 2019.

[6] Hojo, Motokazu, et al. "Automatically recognizing strategic cooperative behaviors in various situations of a team sport." PloS one 13.12 (2018): e0209247.

[7] Miller, Andrew C., and Luke Bornn. "Possession sketches: Mapping nba strategies." Proceedings of the 2017 MIT Sloan Sports Analytics Conference. 2017.