Model Design Deep Learning And Its Applications In The NBA

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 19, 2021

1 Code Repository

https://github.com/jz5jx/deep_learning_NBA_MSDS

2 Model Tuning and Hyper-parameter Architecture

Given our baseline result of 64% Accuracy using both a logit model and a basic neural network architecture we wished to improve upon this for our project. It proved to be difficult to improve on this accuracy after a long list of attempts and trial and error through different neural network architecture's. Beginning with the preprocessing of the data, we conducted normalization and preprocessing using both pandas and the tensorflow dataset API. After the preprocessing, the team fit over twenty different models tweaking everything from the number of hidden layers (5,3, and 2) to the number of nodes in each layer to the activation functions used in each layer before settling on 5 layers with 1024, 512, 256, 128, and 1 nodes respectively using relu for the first 4 layers and sigmoid activation for the final layer. Relu activation was chosen after experimenting with both that and leakyRelu, however the loss function was decreasing more steadily and did not output negative values when using relu activation so we kept that throughout the model. The number of nodes chosen was because we had started with a shallow network and seen no improvement so we moved on to deeper layers till we reached our final structure. Additionally, different regularization methods to include dropout, batch-normalization, L1 and L2 weight decay regularization, and early stopping were implemented. However, our final model only included 2 batch-normalization layers and 2 dropout layers as well as early stopping. The main tuning of the model was based on trial and error through different long runs of each model structure. Lastly, we experimented with both adam and SGD before settling on adam since it is well established in the community and has an adaptive learning rate. After running our final model our accuracy had increased to 65.5%.

2.1 Model Interpretation / Issues

We faced several issues fine tuning our model process. One big challenge for our group was that at first we observed that our loss was not changing and our model was not learning from iterations of batches and epochs. This proved challenging as it seemed that there was nothing inherently wrong in the architecture of the model. However, we made a breakthrough when we realized that other FNN's had encountered and solved this issue by first standardizing the input data. Eventually we too solved this problem by standardizing the data and then also adding multiple batch normalization operations on top of our Dense Layers. This technique helped stabilize the gradient of our back propagation and allowed our model to learn.

Another issue we encountered was improving our model as most of the hyper-parameter tuning lead to lower accuracy scores. We were able to marginally improve by increasing our model width and depth and playing around with Early Stopping, Regularization, and Drop Out techniques as well.

Additionally, because we were working with positional data, the interpretation of this Neural Net is not practical and should not be attempted as it is not germane to the problem statement.

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.