Prediction of Basketball Free Throw Shooting by OpenPose

Conference Paper · October 2018

4 authors, including:

Masato Nakai
Advanced Institute of Industrial Technology
9 PUBLICATIONS **11** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Application skill science by OpenPose View project

Prediction of Basketball Free Throw Shooting by OpenPose

Masato Nakai,Yoshihiko Tsunoda, Hisashi Hayashi, Hideki Murakoshi

School of Industrial Technology, Advanced Institute of Industrial Technology
1-10-40, Higashiooi, Shinagawa-Ku,Tokyo, 140-0011, Japan

Abstract. OpenPose, which is developed by Carnegie Mellon University (CMU) presented in CVPR 2017, takes in real-time motion images via a simple web camera and is capable of recognizing skeletons of multiple persons in these images. It also generates recognized skeleton point coordinates to files. OpenPose is featured by CMU's original top-down method for real-time recognition and it is open online especially for research purposes. Thus we aimed to build a posture analysis model using OpenPose skeletal recognition data and verifying the practicality of OpenPose by verifying the accuracy of the model. As a posture analysis model, we adopted a logistic regression model that predicts the shooting probability of the basketball free throw with skeleton posture data as explanatory variables and the fact whether the ball enters the basket or not as a binary target variable. As the result, sufficiently significant prediction accuracy was obtained. Therefore, posture analysis using OpenPose has been verified to be practical with our model. We consider that with many skeleton data which are easily provided by a simple web camera, OpenPose makes statistical diagnostic approach possible. We also consider it could lower costs (in both financial and time-wise) of such an analysis which has previously required more equipments and more time for preparation regarding motion capture analysis systems.

1 Introduction

A large number of various human posture data with high precision are required to improve performance in statistical posture analysis. However, we found that there are not enough posture data available, because of complexity for acquiring posture data. In order to acquire whole posture data as a time-series, there are two major methods. One is video motion analysis(manual marking) and the other is motion capture analysis. Video motion analysis is very laborious, because we have to divide video into photographs and mark annotations on the pose images. Motion capture analysis is necessary to attach so many sensors on limbs that acquisition of data is complicated in operation. At CVPR 2017 conference, Carnegie Mellon University (CMU) presented OpenPose[1][1] which can recognize

skeletons of multiple players in real-time, using a simple web camera, as shown in Fig. 1. OpenPose adopts unique top-down position recognition using Deep Learning and also the unique algorithm as affiliation recognition of body parts by PAF (Part Affinity Fields)[2]. As a result, in the moving skeletal pictures generated by OpenPose, the skeleton marks are shown and overlapped well with the figure of people. And it seems that recognition accuracy is very high even for various people in various environments.

Although Results of OpenPose's paper[2] said that OpenPose had achieved State-of-the-Arts in the COCO2016 keypoints challenge[3], we decided to evaluate the performance of OpenPose on our own in two aspects. At first we evaluated the correspondence between the actual body positions and the output data generated by OpenPose. As a result of column (actual/openPose rate) in Fig. 2 , we found that the CV(Coefficient of Variation : std/mean) of rates of the distance from the neck of OpenPose to the actual body was 0.08. For example,since a neck is wide in the range of several cm,measurement errors can not be avoided to measure actual distance from the neck to each part. we estimate that CV shows acceptable accuracy. So we can evaluate that OpenPose recognizes body points in a static pose.

However the skeletal recognition of OpenPose in dynamic motion is not clear. In order to evaluate OpenPose's performance in a dynamic motion, next we decided to build a basketball shooting prediction model using real-time skeletal data generated by OpenPose. The "OpenPose's performance" we aim to clarify includes the accuracy of dynamic recognition and usefulness as a dynamic data generator with the accuracy of our shooting prediction model using dynamic data generated by OpenPose.

As a result, we found that the free throw prediction model indicated sufficiently significant accuracy. Thus, we found that OpenPose is a convenient and practical generator of posture data.

The rest of the paper is as follows. In Section 2, we briefly review previous dynamic posture analysis models. In Section 3, we show our experimental methods and selection of the prediction model. In Section 4, we show experimental result and estimation the accuracy of prediction. In Section 5,we conclude this paper. In Section 6, we present future works.

---

[1] OpenPose realizes three-dimensional acquisition by stereo (compound eye) camera in March 2017, but in this research, OpenPose of 2D position recognition version using monocular Web camera is used because of easy operation and sufficient use frequency.

Fig. 1. OpenPose

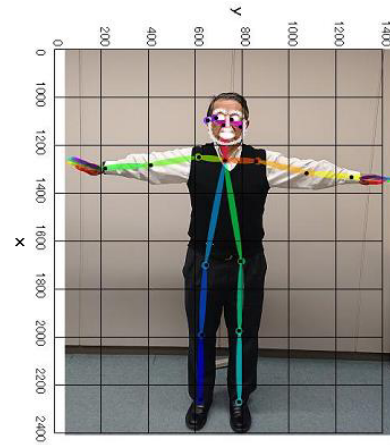| no	name	openPose x	openPose y	Distance from no.1	Actual distance (cm)	actual/openPose rate
0	nose	1114.4	722.1	147.67	20	0.13544
1	neck	1262.0	726.4	0.00		
2	R.shoulder	1266.3	861.2	134.86	18	0.13347
3	R.elbow	1316.9	1076.4	354.34	48	0.13546
4	R.wrist	1333.9	1266.3	544.67	64	0.11750
5	L.shoulder	1249.4	616.5	110.62	18	0.16272
6	L.elbow	1283.1	409.8	317.29	48	0.15128
7	L.wrist	1296.0	215.8	511.67	64	0.12508
8	R.hip	1684.0	793.7	427.30	60	0.14042
9	R.knee	1975.1	785.3	715.54	98	0.13696
10	R.ankle	2270.4	785.2	1010.12	136	0.13464
11	L.hip	1700.9	646.0	446.22	60	0.13446
12	L.knee	1992.0	629.1	736.48	98	0.13306
13	L.ankle	2262.1	624.9	1005.23	136	0.13529
					mean	0.13660
					std	0.01097
					std/mean	0.08034

Fig. 2. Comparison of distance from neck between OpenPose and actual body positions, Column Distance from no.1 is the distance from neck to each point of OpenPose. Column actual/openPose rate is actual distance divided by Distance from no.1

2   Previous research for posture analysis

As a previous analysis of sports motion, there is a method called video motion analysis which divides video into photographs and marks points manually on the pose image for annotation[4]. As a direct sampling posture data, motion capture is used to collect data from sensors on body and limbs[5]. However, these methods are so expensive to collect data that statistical models such as regression could not be applied. MicroSoft KINECT, which is not sold as of early

2018, can easily take 3D posture data, but the sensing range is very narrow and the recognition accuracy of skeletal points are somewhat lower[6].

On the other hand, as a statistical approach, it is realized that winning prediction of basketball game was modeled by logistic regression using records which include the winning / losing results and the frequency of shooting and robbing the ball in the game[7].

As a time-series analysis for motions, there is a research that tries to transfer abstracted motions from a human to a robot with hidden variables estimated by Hidden Markov and reversely predict the next action of the robot from estimated hidden variables [8]. However, our shooting prediction of basketball free throw is not a general time-series model that predicts the next action from the last time-series of motions because our model predicts a result whether to shoot in the basket or not rather than an action. The number of persons monitored in this experiment was limited to 51. We adopted a logistic regression[9] using the features which are composed with the positions at the start and the end or their difference, velocity and acceleration between start and end positions extracted from a time-series of the free throw motions.

3   Method

3.1   Subject of Experiment

In order to construct the shooting prediction model of basketball free throw, we took movies of basketball free throw motions with a full hi-vision video camera. For subject of experiment, we used 51 records which were obtained by two or three trials of various skill levels of 23 persons of a high-school basketball team and some members of an exchange student basketball circle. We generated their skeletal data by OpenPose from the movies. In this experiment, 20 out of 51 records succeeded in the free throw.

3.2   Output of OpenPose

The version of OpenPose[1] adopted in this paper is for 2 dimensional skeleton recognition, and the skeletal coordinates of 18 points (COCO keypoints[2]) as shown in Fig. 3 are outputted to files in about 10 to 20 frames per second depending on a computer performance[2] and connected as shown in Fig. 4 to make time-series data. A skeletal coordinate is composed of 3 values which are x as horizontal, y as vertical and p as confidence probability. We ignored low confident coordinates with less than 0.7 confidence probability.

---

[2] Our experimental machine is CPU: AMD Ryzen 7 1800X, MEMORY: 16GB, GPU: NVidia GeForce GTX 1080ti,OS: Ubuntu 14.04 LTS,CUDA version: 8.0, cuDNN version: 5.1 for CUDA8.0
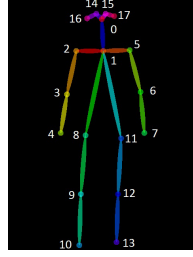
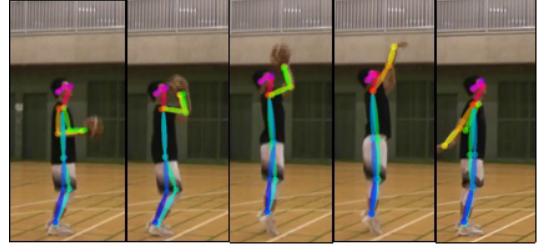**Fig. 3.** Skeleton points of OpenPose cited from [1]

Fig. 4. Connection of frames

Though we used only skeletal data for prediction of shooting model, Open-Pose can recognize also hands and faces as shown in Fig. 5 and outputs each recognized data to each file.
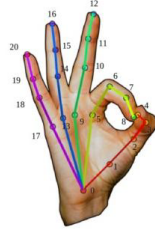

Fig. 5. Recognized finger and face points of OpenPose (cited from [1])

3.3 Statistical model

The shooting prediction model is a binary prediction as to whether to enter the basket or not. As major binary prediction models, there are logistic regression , SVM[10] and Xgboost[11]. The SVM using the kernel method is a nonlinear model which may make high accuracy but cannot calculate the shooting probability because SVM maps data space to higher dimensional space. Xgboost using the stochastic gradient method that has a high reputation for accuracy and robustness is not adequate for a diagnostic model because this model cannot indicate explicitly the degree of importance of explanatory variables. So we adopted a logistic regression model that is easy to interpret and commonly used.

Probability of logistic regression is as follows using α, β and features. The relationship between Z value and probability is shown in Fig. 6. Regression intercept α and coefficents β are calculated by multivariate logistic binary regression on Maximum Likelihood whose partial difference can be solved by Newton-Raphson method[9] because of no local minimum in this optimization.

$$\mathcal{Z} = \alpha + \sum_{i=1}^{n} \beta_i \cdot feature_i \qquad (1)$$

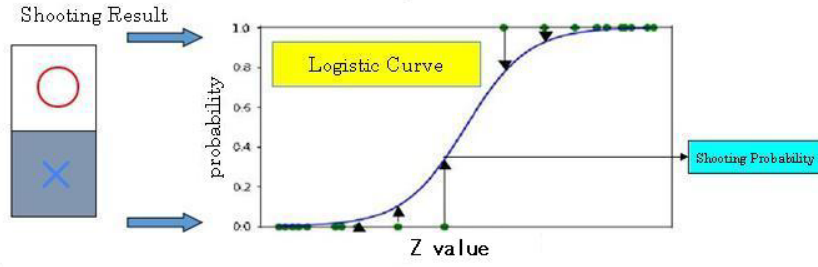$$probability = \frac{1}{1 + \exp(-\mathcal{Z})} \qquad (2)$$

Fig. 6. Logistic regression

3.4 Variables of Logistic regression

The explanatory variables of the logistic regression model are features calculated from the time-series data. Features include the positions of skeletal points, moving speed, acceleration, etc. The target variable is the fact whether the ball enters the basket or not In the time-series data, as shown in Fig. 7, everyone bent knees at the start of the throw and lifted hands to the highest level upon completion of the throw. So we decided to define the interval of a free throw between the time when the knees were bent the most at the start and the time when the hands were lifted to the highest physical point at the end. All the coordinate positions were relative from the neck point.
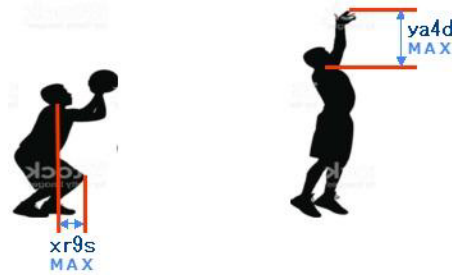

Fig. 7. Posture of start and end

4 Results

4.1 Accuracy of Logistic Regression

In general, the precision of the logistic regression model is expressed in the pareto diagram as shown in Fig. 8. In this diagram, the horizontal axis shows the composition rate of all the members in descending order of shooting probability predicted by the logistic regression, and the vertical axis shows the composition rate of the number of people who succeeded in free throw. About 40% people succeeded at this experiment. Red dots in Fig. 8 indicate composition rate of accumulated people who shot in basket. For example, if a player with a high shooting probability at 5% point of composition rate in descending order of the probability and this success shoot is at the fourth among the total 100 success shoot, the red dot is marked at (0.05, 0.04). If the model was perfect, it would be represented by the line of the perfect model with descending order of shooting probability, and if the shooting probability of the model was uncorrelated with the actual shooting in basket, it would be the line of the uncorrelated model. The accuracy of logistic regression model is indicated by the ratio of the area A of the cumulative curve shown in Fig. 9. This figure shows that sufficiently significant accuracy AR(AccuracyRatio) =41% was obtained. We also discuss this AR value in more detail in Appendix.

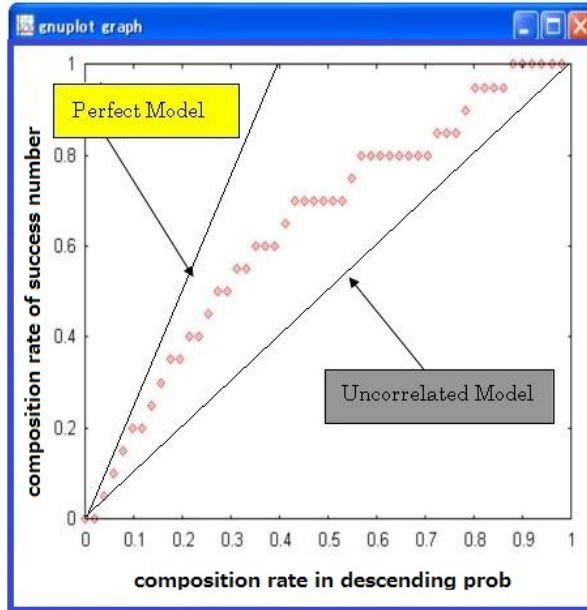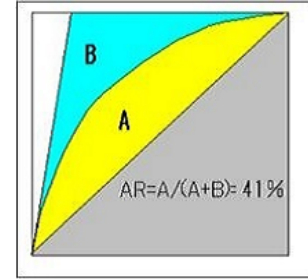**Fig. 9.** Area of AR value

Fig. 8. Parete figure

4.2 Interpretation of significant features in Logistic Regression

As a result, significant features that make the high shooting probability in the logistic regression were shown in Fig. 10. The shooting probability becomes higher when the blue color features are larger. The shooting probability also becomes higher when the red color features get smaller. From this result, the followings were found out. It shows that the shooting probability is higher if the bend of the knees is increased and knees are pulled quickly and at the same time the ball is pulled back and thrown over head. This motion uses the force of the knees' extension and the centrifugal force created when throwing the ball overhead.
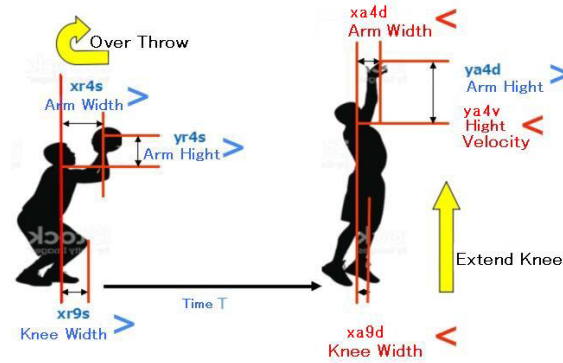

Fig. 10. Relation for shooting probability and features

4.3 Real time Diagnosis

The diagnostic system using OpenPose can display the shooting probability in real-time as shown in Fig. 11. and even if there are not any basket and ball, it becomes possible to judge the skill level directly just by gesture.
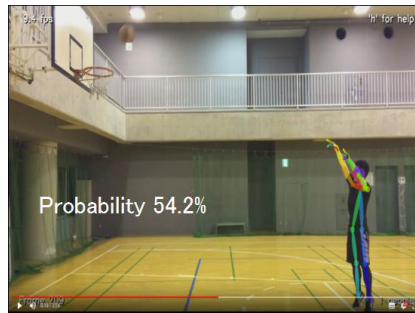

Fig. 11. Real time diagnosis by OpenPose

4.4 Posture diagnosis

By comparing feature quantities between a beginner (a person with low shooting probability) and an expert (a person with high shooting probability), it is possible to diagnose the amount of correction for beginner's postures. In the example of Fig. 12, one of the remarkable differences between the beginner and the expert is the position of the arm at the start in this experiment. The beginner pushed the ball from the chest, but the expert put the ball in front of the head and threw the ball over head. In this case, it is necessary to teach the beginner the form of overhead throwing.
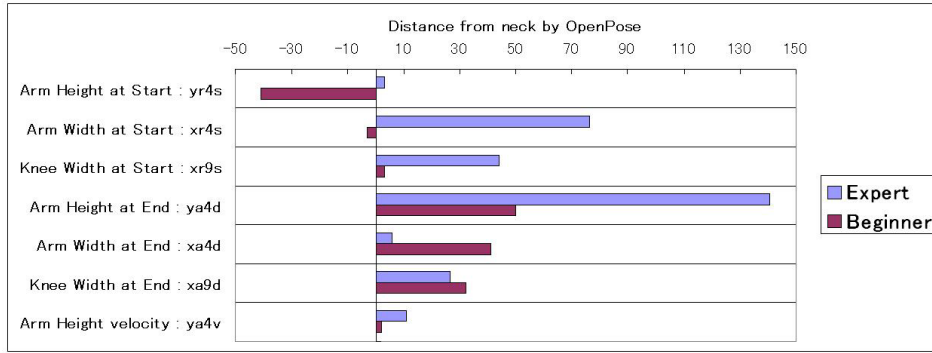

Fig. 12. Comparison of features between the beginner and the expert. Horizontal axis represents significant feature value in Fig. 10.

Next we tried to analyze the sensitivity. As the result in Table 1,improvement of shooting probability was obtained when the beginner's pose improves 2.0 (about 1cm) at each feature.

Table 1. Sensitivity Analysis, Column inc(%) shows incremental of shooting probability

Name	Val	yr4s	xr4s	xr9s	ya4d	xa4d	xa9d	xa4v	prob(%)	inc(%)
Arm Height at START	yr4s	2.0	0.0	0.0	0.0	0.0	0.0	0.0	6.44	0.20
Arm Width at START	xr4s	0.0	2.0	0.0	0.0	0.0	0.0	0.0	6.27	0.03
Knee Width at START	xr9s	0.0	0.0	2.0	0.0	0.0	0.0	0.0	6.41	0.16
Arm Height at END	ya4d	0.0	0.0	0.0	2.0	0.0	0.0	0.0	6.47	0.22
Arm Width at END	xa4d	0.0	0.0	0.0	0.0	-2.0	0.0	0.0	6.41	0.16
Knee Width at END	xa9d	0.0	0.0	0.0	0.0	0.0	-2.0	0.0	6.76	0.52
Arm Height Velocity	ya4v	0.0	0.0	0.0	0.0	0.0	0.0	-2.0	6.70	0.46
ALL		2.0	2.0	2.0	2.0	-2.0	-2.0	-2.0	8.55	2.31

5   Conclusion

We evaluated the performance of OpenPose. At first, we evaluated the correspondence between the actual body positions and the output data generated by OpenPose in the static conditions. Next, for evaluating OpenPose in the dynamic conditions, we built a basketball free throw prediction model by a logistic regression model. We found out the followings.

1. The skeletal data recognized by OpenPose are found to be highly applicable with sufficient accuracy.
2. In the previous posture diagnosis, data are generated by marking on a picture frame from a video stream or collected sensor signals by motion capture on the human parts. These methods are so expensive for data collection that statistical models could not be introduced. On the other hand, OpenPose can easily collect accurate data by using a simple web camera, it made it possible to obtain more accurate posture diagnosis by collecting more data.

6   Future works

The data of basketball free throw in this experiment were taken from one side only by a web camera, so it was suitable to analyze with 2 dimensional data provided by OpenPose. However 3 dimensional motion data could bring better analysis in many occasions in general sports motion analysis field. so use of 3 dimensional OpenPose or expand 2D data generated by 2D OpenPose to 3D data[12] could be demanded. Also instead of the regression model, we would like to challenge a motion analysis by time series models reflecting correlation between the skeleton points using many data generated by OpenPose.

A   Appendix: Discussion on the AR value

In order to consider obtained our AR value in this experiment, we compared NBA(National Basketball Association) Free Throw data[13]. Fig.13 shows the histogram of free throw success rate of NBA's 238 players who threw more than 5 times last year. As shown in Table 2,we generated simulation data according to the number of NBA histogram. These data are consisted of the level and the

---

[3] Japan Student Services Organization

binary flag in each record. The level is set according to the success rate, but the success rate less than 0.5 was compiled to level 4 because of very few people. The binary flag is set randomly according to the success rate. But we generated 2380 records by multiplying the number by 10 to avoid bias of the random.

We made logistic regression using the binary flag as a target variable and the level[4] as a regression variable to obtain the AR value. As the result we obtain AR = 35%. In this simulation, even if the level as a regression variable has a strong correlation with the success rate explicitly, the AR was only 35%. We thought that the low AR value is due to the relatively small number of people at high and low levels. Because we obtained AR = 60% in the case of same number at each level in our simulation. Assuming expert or beginner players were somewhat few in our experiment, our experiment AR=41% can be considered as sufficiently significant accuracy.
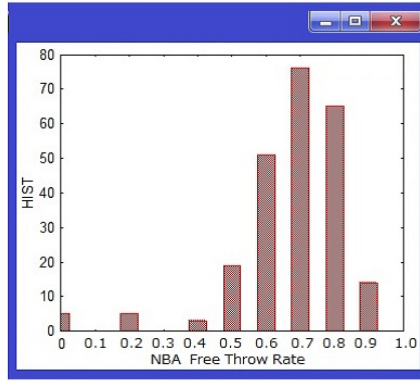

Fig. 13. Histogram of NBA Free Throw Success Rate

Table 2. Simulation data for AR

Level	Number	Success rate
9	140	0.9
8	650	0.8
7	760	0.7
6	510	0.6
5	190	0.5
4	130	0.25

---

[4] Since many same records are generated according to this table, we added a small perturbation of $\mathcal{N}(0, 0.01)$ to level value to avoid rank deficient by same records. For example 7.0026 at level 7.

References

1. CMU-Perceptual-Computing-Lab,
   https://github.com/CMU-Perceptual-Computing-Lab/openpose
2. Cao, Z., Simon, T., Wei, S., Sheikh, Y.: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. CVPR2017(2017).
3. MSCOCO keypoint evaluation metric.
   http://mscoco.org/dataset/#keypoints-eval
4. Hulka, K., Cuberek, R., Svoboda, Z.: Timemotion analysis of basketball players: a reliability assessment of Video Manual Motion Tracker 1.0 software. Journal of Sports Sciences(2013)
5. Wang, X.: An Optimization Method of Basketball Teaching and Training System Design based on Motion Capture Technology.(2017)
6. Dutta, T.: Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. Applied Ergonomics(2013)
7. Shanahan, J. K.: A model for predicting the probability of a win in basketball. University of Iowa(1984)
8. Inamura, T., Nakamura, Y., Toshima I.: Embodied symbol emergence based on mimesis theory. International Journal of Robotics Research,vol. 23, no. 4, pp. 363377(2004)
9. Bishop, C. M.: Pattern Recognition and Machine Learning. 4.3.2 Logistic regression, Springer(2006)
10. Platt, J. C.: Fast Training of Support Vector Machines using Sequential Minimal Optimization. (1999)
11. Chen, T., Guestrin, C.: XGBoost :A Scalable Tree Boosting System.
    arXiv2016.02754(2016)
12. Kudo, Y., Ogaki, K., Matusi, Y., Odagiri Y.: Unsupervised Adversarial Learning of 3D Human Pose from 2D Joint Locations. arXiv:1803.08244(2018)
13. ESPN Western Conference Player Free-Throw Shooting Statistics - 2018-19,
    http://www.espn.com/nba/statistics/player/_/stat/free-throws/sort/freeThrowPct/league/west/qualified/false/order/false