# Comparing the performance of basketball players with decision trees and TOPSIS

Erhan Cene<sup>1</sup>, Coskun Parim<sup>1</sup>, Batuhan Ozkan<sup>1</sup> <sup>1</sup>Department of Statistics, Yildiz Technical University, Davutpasa Campus, 34220, Istanbul, Turkey

Abstract— In this study, individual game statistics for basketball players from Euroleague 2017-2018 season are analysed with Decision Trees and Technique for Order-Preference by Similarity to Ideal Solution (TOPSIS) methods. The aim of this study is to create an alternative ranking system to find the best and the worst performing players in each position eg. guards, forwards and centers. Decision trees are a supervised learning method used for classification and regression. The aim of the decision trees is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. On the other side, TOPSIS is another method to construct a ranking system by using a multi-criteria decision-making system. All the individual statistics such as points, rebounds, assists, steals, blocks, turnovers, free throw percentage and fouls are used to construct the rankings of players. Both decision trees and TOPSIS results are compared with the Performace Index Rating (PIR) index of players which is a single number expressing the performance of the player. Comparing these 3 measures revealed the over and underperformers in the Euroleague for the 2017-2018 season. The results of individual players' performance are visualized with the proper methods such as Chernoff's faces.

*Keywords*—Decision Trees, Multi-Criteria Decision Making, Sports Statistics, Data mining, Multivariate Statistics

#### I. INTRODUCTION

Using statistics in sports is getting popular for the last two decades with the improvement in computer science. Team or player statistics had been used for a much longer time but implementing statistical methods to make inferences is a relatively new field. Statistics is widely used in sports in terms of measuring team success, predicting game outcomes, evaluating a player or team performance and efficiency, ranking players or teams.

Basketball is one of the most popular sports in the world. NBA and Euroleague is the biggest two organisations of basketball where NBA founded in USA and Euroleague facilitates in Europe. Total of 16 teams from 9 countries participated in Euroleague for the 2017 - 2018 season. This study focuses on the Euroleague players from the 2017 - 2018 season and aims to adopt an alternative ranking system for the players by employing the Technique for Order-Preference by Similarity to Ideal Solution (TOPSIS) method.

In order to achieve the new ranking system, various

player statistics from the 2017 - 2018 Euroleague season are gathered from the Euroleague website. Players are divided into three positions: Guards, forwards and centers. ANOVA analysis and decision trees are employed to find which variables are statistically significant among groups. Different weights are assigned to those variables. TOPSIS results provide a new ranking for players. The results of the new ranking system are visualised with the Chernoff's faces and the new ranking of the players is compared with the Euroleague's Performance Index Rating (PIR) index.

#### II. PREVIOUS WORKS

Bozbura, Beşkese and Kaya [1] used TOPSIS as Multiple Criteria Decision-Making (MCDM) method and ranked 6 NBA players by using rebounds, points, blocks, assists, age and salary.

Cooper, Ruiz and Sirvent [2] and Cooper, Ramon and Ruiz (2011) employed data envelopment analysis to rank basketball players from Spanish Basketball League and developed an alternative performance index.

Piette, Pham and Anand [3] used network analysis techniques to evaluate basketball players. Play-by-play data from multiple seasons of NBA is used and over and underperformers are determined on offence, defence and overall.

Reza Kiani Mavi et al [4] ranked football teams with AHP and TOPSIS for the Bundesliga data from the 1999 / 2000 season.

Radovanovic et al [5] used data envelopment analysis (DEA) and distance based analysis (DBA) to rank 26 NBA players with the data from 2011/2012 season.

Chen, Lee and Tsai [6] used AHP and TOPSIS methods to find which players should start in a baseball match by using data from Chinese Profeessional Baseball League (CPBL) from 2011 season.

Changwu [7] used TOPSIS and gray correlation method to the 12 basketball teams that participated 2012 London Olympic Games and competitiveness of each team is determined.

Moreno and Lozano [8] employed a network DEA approach to assess the efficiency of 30 NBA teams from 2009-2010 season.

Atefeh Masoumzadeh and Amirteimoori [9] used Spanish League data for 35 basketball players and rank them with DEA.

Ergül [10] employed TOPSIS method on stock market

Manuscript received November 26, 2018; accepted December 21, 2018. \*Corresponding author: ecene@yildiz.edu.tr

data for Turkish football clubs and find that success in football provides positive impact on finance.

Geyik and Eren [11] employed AHP and TOPSIS methods to teams from Turkish Super Basketball League and Euroleague. They rank the teams according to TOPSIS results and compared the results with the real life results.

#### III. DATA AND VARIABLES

Euroleague uses the PIR index for ranking the players. In this approach, positive stats such as points, rebounds, assists, and steals are considered as a positive component of the index whereas, negative stats such as missed shots, turnovers, fouls committed are considered as a negative component of the index. PIR index is calculated as follows:

$$PIR = \begin{pmatrix} Points + Rebounds + Assists \\ + Steals + Blocks + Fouls Drawn \end{pmatrix} - \\ \begin{pmatrix} Missed Field Goals + Missed Free Throws \\ + Turnovers + Shots Rejected + Fouls Committed \end{pmatrix}$$
(1)

This approach has several weaknesses. First of all, it assigns the same weight to all of the statistics regardless of the importance of it. Also, it neglects the position of the player, some statistics have greater importance on certain positions such as blocks are crucial for centers and assists are important for guards. TOPSIS method aims to fill this gap as it can give different weights to each variable based on positions.

#### A. Data

Player statistics data is gathered from the website of the Euroleague. Data is separated into three subgroups based on the player position as: Guards, forwards and centers. Next, players who spent limited time on the court are excluded from the analysis. For this, players who played at least 10 games and 15 minutes per game are considered as eligible for the analysis. After excluding players, a total of 150 players consisted of 62 guards, 54 forwards and 34 centers are included in the analysis.

Two set of data is used in this study. First one is the raw data consists of average stats for players and the second one is the stats normalised for 40 minutes. Normalising the data for 40 minutes indicates which stats a player would record if he played full game. Such an approach eliminates the effect of MPG over other stats as MPG increases other stats would increase too. After normalising the data TOPSIS analysis is performed with the standardized Z scores within each position. This approach eliminates the inflation of some stats of players who played around the threshold minute of 15.

#### B. Variables

Selected variables reflects every aspect of the game such as shooting, rebounding, ball handling and defence and durability [2].

• Games Played (GP): Total number of games a player played through the season. This variable is related with the durability part of the game.

• Minutes Per Game (MPG): Average number of

minutes a player stay on court per game. This variable is related with the durability part of the game.

• Adjusted Field Goal (AFG): AFG is an advanced metric that shows the shooting ability of a player. AFG is calculated with the given formula,

$$AFG = (PPG - FTPG) \times AFG\%$$

(2)

where PPG is points per game and FTPG is free throws made per game and AFG% is the adjusted field goal percentage which is calculated with

$$\frac{PPG - FTPG}{2 \times FGA}$$
(3)

where FGA is the number of field goal attempts. This variable is related with the shooting part of the game.

• Adjusted Free Throw (AFT): AFT is related with the shooting part of the game and is defined with

$$AFT = FTM \times FT\%$$

(4)

where FT% is the free throw success percentage.

• **Rebounds Per Game (RPG):** Average number of rebounds a player made per game. This variable is related with the rebounding part of the game.

• Assists Per Game (APG): Average number of assists a player made per game. This variable is related with the ball handling and shooting part of the game.

• Steals Per Game (SPG): Average number of steals a player made per game. This variable is related with the defence part of the game.

• Blocks Per Game (BPG): Average number of blocks a player made per game. This variable is related with the defence part of the game.

• **Inverse of Turnovers (TOV\_INV):** Inverse of average number of turnovers a player made per game. Taking inverse of the turnovers provide consistency with other variables as normally higher turnovers indicates worse performance by a player. This variable is related with the ball handling part of the game.

• Non-Committed Fouls Own (NON\_PF): Average number of non-committed fouls of a player.

$$NON \_ PF = 5 - PF$$

(5)

where PF is the personal fouls. The logic behind this variable is same with the TOV\_INV. This variable is related with the defence and durability part of the game.

• Fouls Received (FOUL\_REC): Average number of fouls that opposition players made to the player. This variable is related with the shooting and ball handling part of the game.

#### IV. METHODS

Four different methods are employed in this study. ANOVA analysis with the Bonferroni post-hoc test and decision trees are used for determining significant variables among player positions. Afterwards TOPSIS method is used to develop an alternative ranking system for the performance of basketball players. Finally results are visualized with the Chernoff faces.

R statistical programming language is used for the analysis. "rpart" and "rpart.plot" packages are used to construct decision trees, TOPSIS function from the "MCDA" package is used for the TOPSIS analysis, and faces function from "aplpack" package is used to construct Chernoff faces.

#### A. Decision Trees

Decision trees can be used for multiple purposes such as prediction, classifying or regression [12]. In this study decision trees are used to classify basketball players according to their positions to give an insight on which basketball related statistics separate them. This approach provides to assign different weights to different positions. Decision trees are widely preferred due to its simplicity and graphical representation. Decision trees consists of roots, branches and leafs, where dependent variables are separated into smaller fractions with the help of branches and leaves to the smaller fragments that show similar properties.

#### B. TOPSIS

The Technique for Order Preferences by Similarity to an Ideal Solution (TOPSIS) method was proposed by C. Hwang and K. Yoon [13]. The idea behind TOPSIS is to choose the best alternative nearest to the positive ideal solution (optimal solution) and farthest from the negative ideal solution (inferior solution). The distance measure used in this study is the Euclidean distance. With the Euclidean distance the procedures of TOPSIS can be described as follows [14]:

Let  $A = \{A_k \mid k = 1, ..., n\}$  denotes set of alternatives, denotes  $C = \{C_j \mid j = 1, ..., m\}$  set of criteria.  $X = \{X_{kj} \mid k = 1, ..., n; j = 1, ..., m\}$  indicates the set of performance ratings for each criteria and each alternatives where  $w = \{w_j \mid j = 1, ..., m\}$  is the set of weights for each criteria. Then the information table I = (A, C, X, W) can be given with the following form (Table – I).

TABLE I. THE INFORMATION TABLE OF TOPSIS

Alternatives	C <sub>1</sub>	<b>C</b> <sub>2</sub>	 Cm
$A_1$	X <sub>11</sub>	X <sub>12</sub>	 X <sub>1m</sub>
$A_2$	X21	X <sub>22</sub>	 X <sub>2m</sub>
An	X <sub>n1</sub>	X <sub>n2</sub>	 X <sub>nm</sub>
W	$W_1$	$W_2$	 Wm

Step 1: Calculating the normalized ratings.

$$r_{kj}(x) = \frac{x_{kj}}{\sqrt{\sum_{k=1}^{n} x_{kj}^{2}}}, k = 1, ..., n; j = 1, ..., m.$$
(6)

**Step 2:** For the benefit criteria calculate the weighted normalized ratings with

$$v_{kj}(x) = w_j r_{kj}(x), k = 1, ..., n; j = 1, ..., m.$$

(7)

**Step 3:** Positive Ideal Point (PIS) and Negative Ideal Point (NIS) are determined with the maximum and v

minimum values for 
$${}^{kj}$$
 in each criterion.  
 $PIS = A^{+} = \left\{ v_{1}^{+}(x), v_{2}^{+}(x), ..., v_{j}^{+}(x), ..., v_{m}^{+}(x) \right\}$   
 $NIS = A^{-} = \left\{ v_{1}^{-}(x), v_{2}^{-}(x), ..., v_{j}^{-}(x), ..., v_{m}^{-}(x) \right\}$ 
(8)

**Step 4:** Calculate the separation from the PIS and the NIS between alternatives.

$$D_{k}^{*} = \sqrt{\sum_{j=1}^{m} \left[ v_{kj}\left(x\right) - v_{j}^{*}\left(x\right) \right]^{2}}, k = 1, ..., n$$
(9)

$$D_{k}^{-} = \sqrt{\sum_{j=1}^{m} \left[ v_{kj}(x) - v_{j}^{-}(x) \right]^{2}}, k = 1, ..., n$$
(10)

Step 5: The similarities to the PIS can be derived with:

$$C_{k}^{*} = \frac{D_{k}^{-}}{\left(D_{k}^{*} + D_{k}^{-}\right)}, k = 1, ..., n$$

(11)

where 
$$C_k^* \in [0,1]$$

### C. Chernoff Faces

Chernoff faces is a graphical method proposed by H. Chernoff [15] which visualizes multidimensional data by using the properties of the faces. Each aspect of a face denotes a different variable.

#### V. RESULTS

Descriptive statistics for the selected variables for the raw data and for the normalised data for 40 minutes is given in Table I. Rebounds, assists, blocks and received fouls showed differences among positions for both per game stats and per 40 minutes stats according to one – way ANOVA and Bonferroni test results. In addition, non-made fouls are also effective among positions for per game stats. Rebounds and blocks have different means for each position and guards have higher assists than forwards and centers both in per game stats and 40 minutes stats. Centers make more fouls than other positions according to non-made fouls. In the received fouls forwards receive significantly less fouls than other positions in per game stats, and each position receive different number of fouls according to 40 minutes statistics.

Fig. 1. shows the decision tree results based on position. Similar to ANOVA results, assists, rebounds, blocks and non-made fouls are effective to make a separation among positions. Table III shows the results of classification tree.

Decision tree results suggest guards are specialised in assists and centers are specialised in rebounds and blocks. Forwards take more rebounds and blocks more shots than guards but less than centers. On the other side forwards make less assists than guards and make more assists than centers. Also centers take less minutes per game and forwards make less fouls.

Table III gives the classification table and accuracy.

Overall accuracy of the decision tree is 0.87. Decision tree, correctly predicted 91% of guards, 78% of forwards and 94% of centers.

Separate weights are given for each position and also for per game and per 40 minutes stats.

The weights for each position are given in Table IV.

## TABLE II. DESCRIPTIVE STATISTICS, ANOVA AND BONFERRONI TEST RESULTS FOR THE VARIABLES AMONG POSITIONS (a) PER GAME STATS (b) PER 40 MIN

	GUARDS (n=62)		FORWARDS (n=54)		CENTERS (n=34)		TOTAL (n=150)		ANOVA	
PER GAME STATS	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	p value	Source of Diff.
GP	28.10	6.42	27.80	6.53	28.12	6.02	27.99	6.33	0.960	
MPG	21.78	3.95	21.45	3.69	20.41	3.53	21.35	3.78	0.232	
AFG	4.08	1.80	3.79	1.17	4.14	0.98	3.99	1.43	0.438	
AFT	1.35	0.93	1.06	0.50	1.30	0.54	1.23	0.72	0.084	
RB	2.13	0.72	3.35	1.22	4.92	0.91	3.20	1.44	0.000***	G - FW - C
APG	2.85	1.67	1.43	0.63	1.03	0.55	1.93	1.41	0.000***	G-FW, C
SPG	0.75	0.33	0.64	0.27	0.60	0.30	0.67	0.31	0.040**	
TOV_INV	0.88	0.51	1.07	0.37	1.06	0.40	0.99	0.45	0.042**	
BPG	0.15	0.08	0.28	0.19	0.61	0.39	0.33	0.30	0.000***	G - FW - C
NON_PF	2.95	0.51	2.99	0.48	2.59	0.52	2.88	0.53	0.001***	C – G, FW
FOUL_REC	2.38	1.27	1.84	0.75	2.76	0.88	2.27	1.07	0.000***	FW – G, C

\*\*\* p < 0.001 \*\* p < 0.05 G: Guards, FW: Forwards, C: Centers

	GUARDS (n=62)		FORWARDS (n=54)		CENTERS (n=34)		TOTAL (n=150)		ANOVA	
PER 40 MIN STATS	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	p value	Source of Diff.
GP										
MPG										
AFG	7.33	2.50	7.07	1.89	8.19	1.74	7.43	2.16	0.055	
AFT	2.38	1.38	1.99	0.96	2.56	1.02	2.28	1.18	0.058	
RB	3.91	1.09	6.25	2.00	9.79	1.97	6.09	2.80	0.000***	G - FW - C
APG	5.06	2.45	2.64	0.98	1.99	0.93	3.49	2.19	0.000***	G – FW, C
SPG	1.35	0.50	1.20	0.51	1.16	0.53	1.25	0.52	0.147	
TOV_INV	1.75	1.30	2.12	1.04	2.14	0.82	1.97	1.12	0.132	
BPG	0.17	0.19	0.48	0.36	1.20	0.76	0.52	0.59	0.000***	G - FW - C
NON_PF	5.66	1.71	5.79	1.51	5.24	1.41	5.61	1.58	0.273	
FOUL_REC	4.22	1.74	3.44	1.36	5.47	1.68	4.22	1.76	0.000***	G - FW - C

\*\*\* p < 0.001 \*\* p < 0.05 G: Guards, FW: Forwards, C: Centers

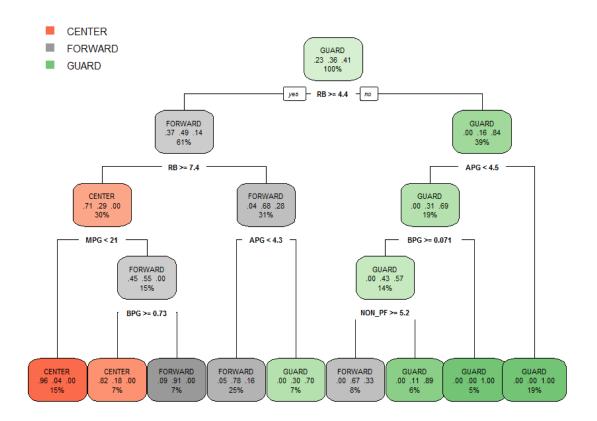


Fig. 1. Decision tree results for player positions.

TREE		PREDICTED							
IKEE	POSITION	GUARDS	FORWARDS	CENTERS					
	GUARDS	52	10	0					
A	FORWARDS	4	47	3					
OBSERVED	CENTERS	0	3	31					
ER	SENSITIVITY	0.9286	0.7833	0.9118					
BS	SPECIFITY	0.8936	0.9222	0.9741					
0	BAL. ACCURACY	0.9111	0.8528	0.9430					
	OVER. ACCURACY	0.8667	KAPPA	0.795					

TABLE III. DECISION TREE CLASSIFICATION RESULTS

: Guards, F: Forwards, C: Centers

Top 10 players according to TOPSIS results for the each position and both for per game and per 40 minutes statistics is given in Table V. Table V also gives PIR index results for the 2017-2018 season gathered from the Euroleague website, for comparison purposes.

Results for per game statistics did not differ widely from the PIR index results. The new ranking can be considered as a fine tune with the position specific weights. Looking at the Per 40 minutes stats, the results differ a lot from PIR index ranking especially for the forwards. This indicates the requirement of better weight selection for the forwards in the future studies. Other than this problem, TOPSIS results show great stability and stands as a solid alternative for the PIR index.

TABLE IV. WEIGHTS FOR EACH POSITION FOR TOPSIS	5
--	---

VARIABLES	WEIG	HTS (PE	R GAME)	WEIGHTS (40 MIN)				
	G	F	С	G	F	С		
GP	0.107	0.119	0.112	0.000	0.000	0.000		
MPG	0.214	0.238	0.224	0.000	0.000	0.000		
AFG	0.161	0.178	0.168	0.222	0.241	0.258		
AFT	0.107	0.119	0.112	0.111	0.138	0.129		
RB	0.054	0.079	0.093	0.111	0.172	0.215		
APG	0.125	0.079	0.075	0.222	0.103	0.086		
SPG	0.071	0.079	0.075	0.111	0.103	0.086		
TOV_INV	0.054	0.020	0.019	0.111	0.069	0.043		
BPG	0.036	0.059	0.075	0.037	0.103	0.151		
NON_PF	0.036	0.010	0.009	0.037	0.034	0.022		
FOUL_REC	0.036	0.020	0.037	0.037	0.034	0.011		
TOTAL	1	1	1	1	1	1		
G								

TABLE V. TOPSIS RESULTS FOR GUARDS, FORWARDS AND CENTERS PER GAME AND PER 40 MINUTES.

				GUA	RDS				
Р	ER GAN	ЛE			PER	R 40 MII	NUTES		,
NAME	C*	TOPSIS RANK	PIR	PIR RANK	NAME	C*	TOPSIS RANK	PIR	PIR RANK
SHVED, ALEXEY	0.73	1	20.38	2	CALATHES, NICK	0.61	1	25.53	4
DONCIC, LUKA	0.65	2	21.55	1	DE COLO, NANDO	0.60	2	29.38	2
DE COLO, NANDO	0.64	3	18.94	3	DONCIC, LUKA	0.60	3	33.21	1
CALATHES, NICK	0.63	4	18.52	4	JAMES, MIKE	0.58	4	24.56	6
NEDOVIC, NEMANJA	0.59	5	16.67	5	HEURTEL, THOMAS	0.57	5	23.57	7
RODRIGUEZ, SERGIO	0.57	6	13.83	11	NEDOVIC, NEMANJA	0.55	6	26.58	3
PANGOS, KEVIN	0.56	7	14.19	9	RODRIGUEZ, SERGIO	0.55	7	21.27	12
JACKSON, PIERRE	0.56	8	14.28	8	SHVED, ALEXEY	0.55	8	25.31	5
SPANOULIS, VASSILIS	0.56	9	11.88	17	ROCHESTIE, TAYLOR	0.53	9	22.09	10
JAMES, MIKE	0.55	10	15.67	6	PANGOS, KEVIN	0.52	10	20.61	14
	0.00	10	10107	FORW		0.02	10	20101	
Р	ER GAN	ЛE			PER	R 40 MI	NUTES		
		TOPSIS		PIR			TOPSIS		PIR
NAME	C*	RANK	PIR	RANK	NAME	C*	RANK	PIR	RANK
GILL, ANTHONY	0.63	1	13.09	4	HANGA, ADAM	0.59	1	14.74	33
CLYBURN, WILL	0.62	2	13.61	3	NUNNALLY, JAMES	0.57	2	17.01	21
WRIGHT, DORELL	0.62	3	12.81	6	PAPAPETROU, IOANNIS	0.55	3	11.17	50
JANKUNAS, PAULIUS	0.59	4	13.71	1	GARINO, PATRICIO	0.55	4	11.98	48
MELLI, NICOLO	0.57	5	12	8	SANDERS, RAKIM	0.53	5	14.85	32
BEAUBOIS, RODRIGUE	0.56	6	10.58	14	WHITE, AARON	0.51	6	18.2	17
PAPANIKOLAOU, KOSTAS	0.56	7	11.25	9	THOMPSON, HOLLIS	0.50	7	12.89	44
HONEYCUTT, TYLER	0.55	8	12.82	5	PRINTEZIS, GEORGIOS	0.49	8	21.38	5
PRINTEZIS, GEORGIOS	0.54	9	13.68	2	SASTRE, JOAN	0.48	9	12.01	47
RUBIT, AUGUSTINE	0.54	10	12.1	7	ABALDE, ALBERTO	0.48	10	14.41	34
			-	CENT	ERS				
Р	er gan	ΛE			PER	R 40 MII	NUTES		
		TOPSIS		PIR			TOPSIS		PIR
NAME	C*	RANK	PIR	RANK	NAME	C*	RANK	PIR	RANK
VESELY, JAN	0.65	1	16.03	2	MCLEAN, JAMEL	0.57	1	25.2	9
SHENGELIA, TORNIKE	0.65	2	18.21	1	VESELY, JAN	0.48	2	24.09	16
DUNSTON, BRYANT	0.65	3	15.53	3	MILUTINOV, NIKOLA	0.55	3	25.83	7
SINGLETON, CHRIS	0.59	4	12.47	11.5	SHENGELIA, TORNIKE	0.44	4	28.58	3
AYON, GUSTAVO	0.56	5	14.22	6	HINES, KYLE	0.51	5	22.24	21
GUDAITIS, ARTURAS	0.51	6	15.41	4	DUBLJEVIC, BOJAN	0.60	6	26.43	6
HINES, KYLE	0.50	7	12.1	14	GUDAITIS, ARTURAS	0.54	7	29.8	2
AUGUSTINE, JAMES	0.48	8	14.38	5	SERAPHIN, KEVIN	0.42	8	24.59	13
MILUTINOV, NIKOLA	0.45	9	14.03	7	STIMAC, VLADIMIR	0.41	9	30.44	1
DAVIES, BRANDON	0.45	10	9.78	26	DAVIES, BRANDON	0.54	10	22.45	19.5

Fig. 2. to 4. give Chernoff Faces results for the guards,

forwards and centers respectively. Each aspect of the face represents a different statistics of the player: Height of the face – GP, width of the face – MPG, structure of face – AFG, height of mouth – AFT, width of mouth – RB, smiling – APG, height of eyes – SPG, width of eyes – Inverse of Turnover, height of hair – BPG, width of hair – Non-made fouls, style of hair – received fouls, height of nose – GP, width of nose – MPG, width of ear – AFG, height of ear – AFT.

Exceptional performances can be easily distinguished from Chernoff faces. In guards Luka Doncic, Alexey Shved, Nando de Colo and Nick Calathes are the first noticed players from Fig. 2. In forwards (Fig 3.), Will Clyburn, Georgioz Printezis, Antony Gill, Nicola Melli, Edgaras Ulanovas are first noticed players. In centers (Fig. 4), the most notable players are Jan Vesely, Tornika Shengalia, Chris Singleton and Bryant Dunston.

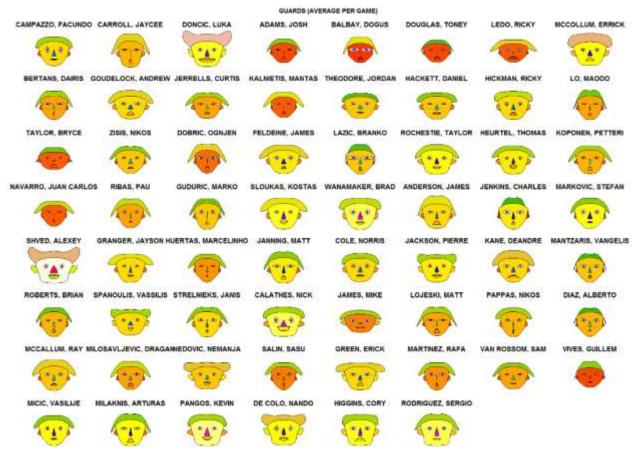
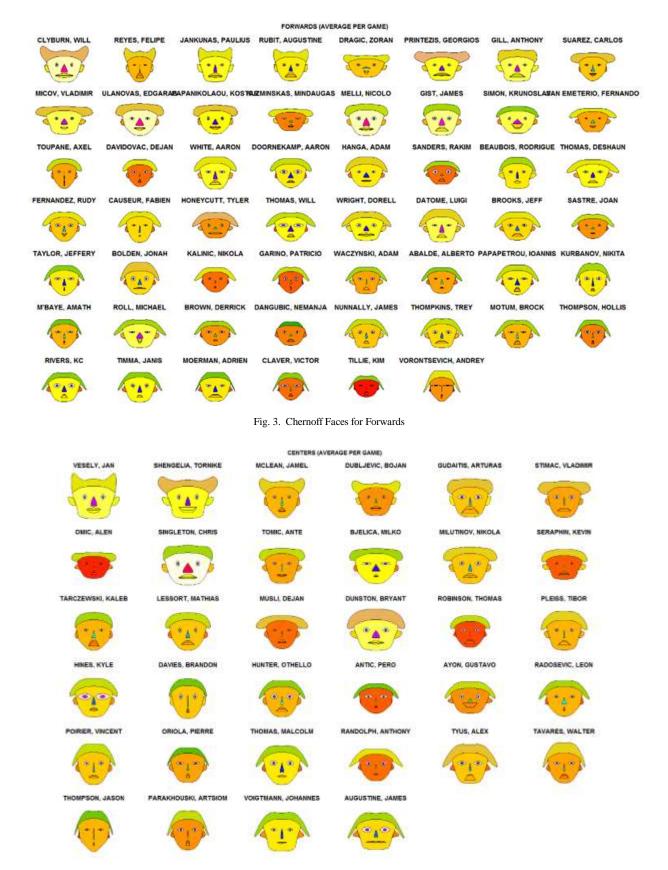


Fig. 2. Chernoff Faces for Guards

#### VI. CONCLUSIONS

In this study, individual game statistics for basketball players from Euroleague 2017-2018 season are analysed with Decision Trees and Technique for Order-Preference by Similarity to Ideal Solution (TOPSIS) methods. The aim of this study is to create an alternative ranking system to find the best and the worst performing players in each position eg. guards, forwards and centers. All the individual statistics such as points, rebounds, assists, steals, blocks, turnovers, free throw percentage and fouls are used to construct the rankings of players. Decision trees and one way ANOVA are used to determine the crucial variables for each position and TOPSIS results are compared with the Performace Index Rating (PIR) index of players which is a single number expressing the performance of the player. Comparing these measures revealed the over and underperformers in the Euroleague for the 2017-2018 season and provide an alternative way to determine player performances.

#### DATA SCIENCE AND APPLICATIONS, VOL. 1, NO. 1, 2018



#### Fig. 4. Chernoff Faces for Centers

#### REFERENCES

- B. F. Tunç, A. Beşkese, and T. S. Kaya, "TOPSIS method on player selection in NBA," in 12th International Research/Expert Conference "Trends in the Development of Machinery and Associated Technology" TMT 2008, 2008, pp. 401–404.
- [2] W. W. Cooper, J. L. Ruiz, and I. Sirvent, "Selecting non-zero weights to evaluate effectiveness of basketball players with DEA," *Eur. J. Oper. Res.*, vol. 195, no. 2, pp. 563–574, 2009.
- [3] J. Piette, L. Pham, and S. Anand, "Evaluating basketball player performance via statistical network modeling," in *MIT Sloan Sports Analytics Conference*, 2011, no. June, pp. 1–11.

- [4] R. K. Mavi, N. K. Mavi, and L. Kiani, "Ranking football teams with AHP and TOPSIS methods," *Int. J. Decis. Sci. Risk Manag.*, vol. 4, no. 1–2, pp. 108–126, 2012.
- [5] S. Radovanovic, M. Radojicic, V. Jeremic, and G. Savic, "A novel approach in evaluating efficiency of basketball players," *Manag. - J. theory Pract. Manag.*, vol. 18, no. 67, pp. 37–46, 2013.
- [6] C. C. Chen, Y. T. Lee, and C. M. Tsai, "Professional baseball team starting pitcher selection using AHP and topsis methods," *Int. J. Perform. Anal. Sport*, vol. 14, no. 2, pp. 545–563, 2014.
- [7] H. Changwu, "Application of the TOPSIS method and gray correlation model in the competitiveness evaluation of basketball teams," *Comput. Model. New Technol.*, vol. 18, no. 12C, pp. 833–837, 2014.
- [8] P. Moreno and S. Lozano, "A network DEA assessment of team efficiency in the NBA," Ann. Oper. Res., vol. 214, no. 1, pp. 99–124, 2014.
- [9] Toloo, M. A. Mehdi, and A. Amirteimoori, "Performance assessment in production systems without explicit inputs: An application to basketball players," *IMA J. Manag. Math.*, vol. 27, no. 2, pp. 143–156, 2016.
- [10] N. Ergül, "Spor kulüplerinin futboldaki başarıları ile spor şirketlerinin finansal başarıları arasındaki ilişkinin test edilmesi," *Hacettepe*

Üniversitesi İktisadi ve İdari Bilim. Fakültesi Derg., vol. 35, no. 3, pp. 43–71, 2017.

- [11] O. Geyik and T. Eren, "Spor Toto Basketbol Süper Ligi ve Turkish Airline Euroleague Basketbol Takımlarının AHS-TOPSIS Yöntemleriyle Değerlendirilmesi," *Spor Bilim. Araştırmaları Derg.*, vol. 3, no. 1, pp. 32–53, 2018.
- [12] L. Rokach and O. Maimon, *Data mining with decision trees*, Second Edi. Singapore: World Scientific Publishing, 2015.
- [13] C. Hwang and K. Yoon, Multiple Attribute Decision Making: Methods and Applications, A State of the Art Survey, vol. 1. 1981.
- [14] G.-H. Tzeng and J.-J. Huang, *Multiple attribute decision making*. Florida: CRC Press, 2011.
- [15] H. Chernoff, "The use of faces to represent points in k-dimensional space graphically," J. Am. Stat. Assoc., vol. 68, no. 342, pp. 361–368, 1973.