

CLUSTER FORMATION APPROACH TO PREDICT SPORTS TALENT ADJUSTING ANTHROPOMETRIC AND MOTOR ABILITIES

Utsav Chaware^{1*}, Dr. Aninidita Das²

¹Ph.D. Scholar, Department of Health Sciences, LNIPE, Gwalior

²Associate Professor, Department of Physical Education Pedagogy, LNIPE, Gwalior

Abstract- Talent identification and location in India are still done using manual calculations, coaches' gut instincts, and selectors' experience. Most of these are either biased to a certain ability of athletes or subjective ways of talent selection. The objective of the study is to build a cluster-based model to identify sporting talent based on their age and adjusted values as per their anthropometric and motor abilities, total of 175 subjects were part of the study. Anthropometric and motor abilities measurements were taken Using standard procedures guided by the Khelo India scheme, a flagship program by Govt. Of India. Results revealed 6 clusters in which homogeneous selections 27, 53, 20, 28, etc. could be done with determining factors in each cluster. It also displays the elements most responsible for the selection process among the clusters.

Keywords- Talent Identification, Clusters, Athletic ability, Anthropometrics, motor abilities, Physical Abilities

Introduction

Talent identification is an intriguing subject to research and argument regarding the outcomes. But talent identification procedures are always subjected to criticism due to convoluted measurements and analysis. With time data analysis and decision-making had been more precise using modern techniques of data processing. Manual processing and uni dimensional dependency solely on either bio motor abilities could lead to faulty judgments or coaches' approaches to select players purely on basis of anthropometric measurements also do not justify the process of talent identification. It has to be a multidimensional approach to select athletes on basis of bio-motor abilities adjusted as per their anthropometrics. This approach could bring better results in sports-specific domains as well, as reported by various researchers that many good athletes are subjected to rejection in a certain phase of the talent identification process and cornered due to the presence of specific abilities and not all the bio-motor abilities as in the case of endurance athletes and sprinters or weight lifters as they perform excellently in some elements but fail to deliver in the others. Also, many of the average height players with superior athletic displays fail to make the selection process as their counterparts with better anthropometric values get placed even with poor bio-motor abilities testing results.

A multidimensional modeling approach to talent identification includes adjustments of both the abilities prime in talent identification. Many recent pieces of research in various sports reported valuable and efficient results using modern data analysis procedures. Jamil et al., (2021) report some breakthrough characteristics which were distinct in Elite and Sub-elite Goal Keepers, using various machine learning algorithms which included cluster analysis too. The accuracy showed was more than 90% which reflects the impetus of the approach. In another recent study by Rasid et al., (2019) a multidimensional analysis for youth performance assessment for talent identification was done resulting in better outcomes and judgments.

The study practices a cluster formation approach to fill the gap in talent identifications' manual and subjective judgments. Since talent identification is based on various abilities, where athletes differ in their athletic abilities and anthropometric measurements as well to having inform approach in talent selection could result in flawed implementation. Since talent identification is majority on sports or to the specific club of sports the talent identification process should follow the same. The study tried to adjust the anthropometric measurements of athletes to their biomotor abilities to predict valuable and precise results Cluster analysis.

Cluster analysis

Cluster analysis is an advanced statistical tool to classify data and rectifies the judgments on basis of the data. For better decision-making of a domain, the data should be homogeneous to compare and select the best. Cluster

analysis diminishes the within heterogeneity and makes a group homogeneous. Cluster analysis is used for the segmentation of respondents of the same nature and similar characteristics, Data simplification Relationship identification in data. Using cluster analysis, we will find clusters of athletes having similar nearly similar abilities and anthropometric variables Primarily, we may end up having a lot of clusters but later these clusters will be agglomerated into more broad clusters to get the talent with precision.

Cluster analysis is an imperative and explicit technique to aid coaches in better and proportionate selection procedures.

Methodology

Anthropometric and motor abilities were assessed for 175 male athletes in the age group of 15-17years. Anthropometric measurements include Height and weight and bio-motor abilities include speed, endurance, upper body strength, core strength, balance, and flexibility. All the measurements were taken using standard procedures guided by the khelo India talent identification scheme. Norms to evaluate the testing were also taken from the scheme. All the subjects were from the rural and tribal areas of Madhya Pradesh.

K-mean cluster analysis was applied using IBM SPSS 20, to the data adjusting their weight and height to their motor ability performance.

Results& Analysis

Table-1

Cases in each Cluster	
	1 27.000
	2 53.000
Cluster	3 20.000
	4 28.000
	5 13.000
	6 34.000
Valid	175.000
Missing	.000

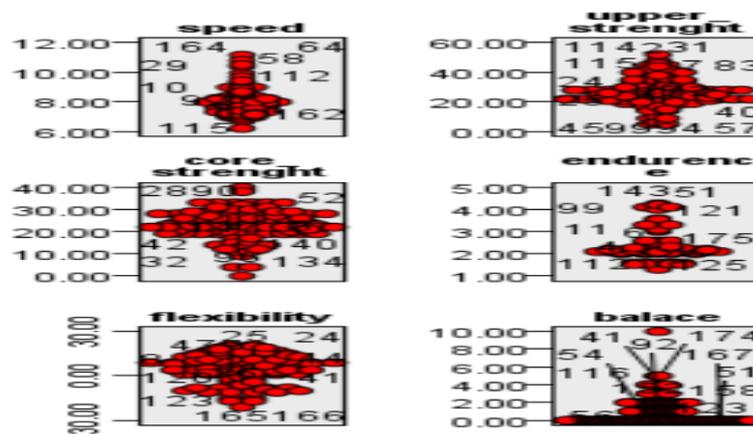
Table 1 displays athletes’ motor ability performances adjusted to their heights, or as per their height, in which only 27 out of 175 were classified in the first cluster revealing only 27 top height players performed all the motor abilities well when compared to other players of similar height. Cluster 2 highlights a majority in which the 2nd top hierarchy in terms of height lies consisting of 53 players who performed the motor abilities well to others. This could be a major cluster for the decision-making of athletes.

Table 2

Cluster Centers	Cluster					
	1	2	3	4	5	6
flexibility	11.59	12.55	2.10	3.18	1.15	12.41
balance	1.07	1.04	1.60	.71	.62	1.15

upper_strenght	30.52	25.11	13.70	24.82	22.08	40.24
core_strenght	25.07	21.15	23.80	25.79	9.77	28.18
endurance	2.26	2.26	2.62	2.31	2.59	1.98
speed	7.35	8.66	7.03	9.87	8.48	8.93

Table 2 displays significant differences in some of the motor components, There exist major differences in flexibility and upper arm strength in different clusters, but balance and endurances were found to be similar with no remarkable differences in the values which tells clusters could be neglected on these variables. Also, strength



fluctuations in some clusters could be seen which may create a significant difference in results.

Fig-1 shows the dominance of abilities in clusters

Fig-1. Shows the cluster prediction of variables of dominance amongst clusters on which selection procedure could be done. Speed was the most determining factor followed by upper strength and core strength, endurance and flexibility were also considerable factors while talent identification using cluster model but balance could be subsided as model show the concentration of all the data sets in one place which highlights the similar balancing abilities for all the subject and clusters.

Discussion of Findings and Conclusion

K-mean cluster applied to a data set discovered a lot of patterns and latent information which couldn't be judged manually by coaches and eventually ends up taking subjective decisions. Cluster analysis using the K-mean cluster classifies the data into homogeneous groups which makes data interpretation and decision-making easy and robust bagged by data and numbers. In the present study, data were divided into 6 clusters based on height and weight which emphasized that certain clusters differ in some abilities whereas others have similar traits approximately such as balance. Speed was reported to be the major decisive factor in selection among the clusters followed by upper body strength which was significant in cluster 1 from the other clusters, but speed was found more in later clusters which demonstrate higher speed abilities in players who are in the middle line of height. This justifies the reason as much height with poor core and body strength could result in altering the speed and agility of the player. Similar results could be obtained by the study by Mohamed et al., 2009; Siener et al., 2021 structuring abilities on the basis of their anthropometric abilities. Further studies could be done using more precise anthropometric variables in particular sports as they could bring exemplary decisions in the area of talent hunt and recognition.

References

- Ahamad, G., Naqvi, S. K., Beg, M. M. S., & Ahmed, T. (2015). A web based system for cricket talent identification, enhancement and selection (C-TIES). *Procedia Computer Science*, 62, 134–142. <https://doi.org/10.1016/j.procs.2015.08.426>
- Bailey, R., & Collins, D. (2016). The Standard Model of Talent Development and Its Discontents. *Kinesiology Review*, 2(4), 248–259. <https://doi.org/10.1123/krj.2.4.248>
- Bonal, J. R., Jiménez, S. L., & Lorenzo, A. (2020). The talent development pathway for elite basketball players in China. *International Journal of Environmental Research and Public Health*, 17(14), 1–21. <https://doi.org/10.3390/ijerph17145110>

- Byounggoo Ko. (2014). Sports Talent Identification and Selection in Korea. *IJASS(International Journal of Applied Sports Sciences)*, 26(2), 99–111. <https://doi.org/10.24985/ijass.2014.26.2.99>
- Guidelines on National Sports Talent Search Scheme (NSTSS). (n.d.). 1–38.
- Haghighat, M., Rastegari, H., & Nourafza, N. (2016). A Review of Data Mining Techniques for Result Prediction in Sports. *March*.
- Jamil, M., Phatak, A., Mehta, S., Beato, M., Memmert, D., & Connor, M. (2021). Using multiple machine learning algorithms to classify elite and sub-elite goalkeepers in professional men's football. *Scientific Reports* 2021 11:1, 11(1), 1–7. <https://doi.org/10.1038/s41598-021-01187-5>
- Jauhainen, S., Ayrämö, S., Forsman, H., & Kauppi, J. P. (2019). Talent identification in soccer using a one-class support vector machine. *International Journal of Computer Science in Sport*, 18(3), 125–136. <https://doi.org/10.2478/ijcss-2019-0021>
- Keogh, J. W. L., Weber, C. L., & Dalton, C. T. (2003). Evaluation of anthropometric, physiological, and skill-related tests for talent identification in female field hockey. *Canadian Journal of Applied Physiology*, 28(3), 397–409. <https://doi.org/10.1139/h03-029>
- Lidor, R., Côté, J., & Hackfort, D. (2009). ISSP position stand: To test or not to test? The use of physical skill tests in talent detection and in early phases of sport development. In *International Journal of Sport and Exercise Psychology* (Vol. 7, Issue 2, pp. 131–146). *Fitness Information Technology*. <https://doi.org/10.1080/1612197X.2009.9671896>
- Medicine, G. K.-B. J. of S., & 2010, undefined. (n.d.). Talent identification and development in children with reference to India. *Bjbm.Bmj.Com*. Retrieved 12 March 2021, from https://bjbm.bmj.com/content/44/Suppl_1/i65.3.short
- Mirkolae, E. G., Mohammad, S., Razavi, H., & Amirnejad, S. (2013). A Mini-Review of Track And Field's Talent-Identification Models in Iran and Some Designated Countries. *Annals of Applied Sport Science*, 1(3), 17–28.
- Mohamad, A. H., Ramli, R., & Ramli, A. F. (2020). A Software Engineering Approach in Netball Performance Analysis: Training and Activities Features for Automatic Players Position Selection. 2020 8th International Conference on Information Technology and Multimedia, ICIMU 2020, 371–377. <https://doi.org/10.1109/ICIMU49871.2020.9243476>
- Mohamed, H., Vaeyens, R., Matthys, S., Multael, M., Lefevre, J., Lenoir, M., & Philippaerts, R. (2009a). Anthropometric and performance measures for the development of a talent detection and identification model in youth handball. <https://doi.org/10.1080/02640410802482417>
- Mohamed, H., Vaeyens, R., Matthys, S., Multael, M., Lefevre, J., Lenoir, M., & Philippaerts, R. (2009b). Anthropometric and performance measures for the development of a talent detection and identification model in youth handball. *Journal of Sports Sciences*, 27(3), 257–266. <https://doi.org/10.1080/02640410802482417>
- Musa, R. M., Abdul Majeed, A. P. P., Taha, Z., Chang, S. W., Nasir, A. F. A., & Abdullah, M. R. (2019a). A machine learning approach of predicting high potential archers by means of physical fitness indicators. *PLoS ONE*, 14(1), 1–12. <https://doi.org/10.1371/journal.pone.0209638>
- Musa, R. M., Abdul Majeed, A. P. P., Taha, Z., Chang, S. W., Nasir, A. F. A., & Abdullah, M. R. (2019b). A machine learning approach of predicting high potential archers by means of physical fitness indicators. *PLoS ONE*, 14(1). <https://doi.org/10.1371/journal.pone.0209638>
- Oytun, M., Tinazci, C., Sekeroglu, B., Acikada, C., & Yavuz, H. U. (2020). Performance Prediction and Evaluation in Female Handball Players Using Machine Learning Models. *IEEE Access*, 8, 116321–116335. <https://doi.org/10.1109/ACCESS.2020.3004182>
- Pearson, D. T., Naughton, G. A., & Torode, M. (2006). Predictability of physiological testing and the role of maturation in talent identification for adolescent team sports. *Journal of Science and Medicine in Sport*, 9(4), 277–287. <https://doi.org/10.1016/j.jsams.2006.05.020>
- Rasid, S. M. M., Abdullah, M. R., Juahir, H., Maliki, A. B. H. M., Kosni, N. A., Musa, R. M., Hashim, M. R., Alnamat, A. S. F., Alias, N., & Najmi, N. (2019). Applied multidimensional analysis for assessing youth performance in sports talent identification program. *International Journal of Recent Technology and Engineering*, 8(2 Special Issue 7), 207–211. <https://doi.org/10.35940/IJRTE.B1051.0782S719>
- Roberts, A. H., Greenwood, D., Stanley, M., Humberstone, C., Iredale, F., & Raynor, A. (2021). Understanding the “gut instinct” of expert coaches during talent identification. *Journal of Sports Sciences*, 39(4), 359–367. <https://doi.org/10.1080/02640414.2020.1823083>
- Rommers, N., Rössler, R., Verhagen, E., Vandecasteele, F., Verstockt, S., Vaeyens, R., Lenoir, M., D'Hondt, E., & Witvrouw, E. (2020). A Machine Learning Approach to Assess Injury Risk in Elite Youth Football Players. *Medicine and Science in Sports and Exercise*, 52(8), 1745–1751. <https://doi.org/10.1249/MSS.0000000000002305>
- Siener, M., Faber, I., & Hohmann, A. (2021). Prognostic Validity of Statistical Prediction Methods Used for Talent Identification in Youth Tennis Players Based on Motor Abilities. *Applied Sciences* 2021, Vol. 11, Page 7051, 11(15), 7051. <https://doi.org/10.3390/APP11157051>
- Sultana, D., & Pandi, J. M. C. (2013). Prediction of Athletic Ability from Talent Identification Model on School Students.
- Zheng, Y., & Liu, S. (2020). Bibliometric analysis for talent identification by the subject–author–citation three-dimensional evaluation model in the discipline of physical education. *Library Hi Tech*. <https://doi.org/10.1108/LHT-12-2019-0248>
- Jamil, M., Phatak, A., Mehta, S., Beato, M., Memmert, D., & Connor, M. (2021). Using multiple machine learning algorithms to classify elite and sub-elite goalkeepers in professional men's football. *Scientific Reports* 2021 11:1, 11(1), 1–7. <https://doi.org/10.1038/s41598-021-01187-5>
- Mohamed, H., Vaeyens, R., Matthys, S., Multael, M., Lefevre, J., Lenoir, M., & Philippaerts, R. (2009). Anthropometric and performance measures for the development of a talent detection and identification model in youth handball. 27(3), 257–266. <https://doi.org/10.1080/02640410802482417>
- Rasid, S. M. M., Abdullah, M. R., Juahir, H., Maliki, A. B. H. M., Kosni, N. A., Musa, R. M., Hashim, M. R., Alnamat, A. S. F., Alias, N., & Najmi, N. (2019). Applied multidimensional analysis for assessing youth performance in sports talent identification program. *International Journal of Recent Technology and Engineering*, 8(2 Special Issue 7), 207–211. <https://doi.org/10.35940/IJRTE.B1051.0782S719>

- Siener, M., Faber, I., & Hohmann, A. (2021). Prognostic Validity of Statistical Prediction Methods Used for Talent Identification in Youth Tennis Players Based on Motor Abilities. *Applied Sciences* 2021, Vol. 11, Page 7051, 11(15), 7051. <https://doi.org/10.3390/APP11157051>