# Prediction Olympic Athletes’ performance using Anthropometric Measurements, Age, Gender and Nationality Data

Timeline

1. Project Planning (2 weeks):
   * Define project objectives and scope.
   * Conduct initial research and gather resources.
   * Create a project plan outlining tasks, milestones, and timelines.
2. Data Preparation (3 weeks):
   * Clean and preprocess the data to ensure consistency and quality.
   * Perform exploratory data analysis (EDA) to understand the dataset's characteristics.
   * Select appropriate methods.
3. Model Development (4 weeks):
   * Choose suitable machine learning models for analysis.
   * Develop and train the models using the prepared data.
4. Documentation and Reporting (4 weeks):
   * Document the project process, including methodologies, findings, and challenges faced.
   * Prepare a final report summarising the project's objectives, methods, results, and conclusions.
   * Create visualisations, diagrams, and presentations to demonstrate key insights effectively.
5. Finalisation and Submission (1 week):
   * Make final adjustments based on feedback received.
   * Ensure all documentation and deliverables are complete and polished.
   * Submit the final project report, code, ethic form, title page as required.

Introduction

The vast majority of athletes consider the Olympic Games as the top of their sporting and physical achievements. A large number of advancements and discoveries can be associated with the drive to succeed in Olympic games.

Olympic athletes represent the highest level of human physical abilities, as highlighted by Borms and Hebbelinck (1984) in their study. They assert that athletes competing in the Olympics or world class athletes show the ideal composition of genetic predispositions and environmental factors, as a result they are able to reach a peak of their performance. Theoretically, those athletes who stand out the most in their specific events, possessing the optimal physical structure for those types of physical activity, as discussed by Carter (1985).

Within last century Olympic Games have been experiencing an enormous upward of competitiveness. In this evolving stage, coaches and supporting team should collaborate closely with data scientists to create and develop predictive models for athletes' performance, identify outliers, patterns and trends. This collaboration is essential not only for improving current athletes' metrics but also for the selection of young talents for specific events. Numerous studies, including those by Cuk and Karácsony (2002), Arkaev and Suchilin (2004), pointed out the significance of anthropometric characteristics in influencing the success of athletes in achieving their goals.

The significance of elite athletes' age in their performance also must be mentioned. Sports researchers acknowledge that even the month of birth compering to other athletes may determine an athlete's likelihood of reaching high performance, a phenomenon known as the relative age effect, as highlighted by J. Musch (2001).

Another vital factor which has significant impact on athletes’ successes is nationality or the country they represent. It is very important to consider because each country which shows up on Olympic Games has it own methodologies for many disciplines. Identifying the correlations between country which provides training and its success may a key for mutually beneficial exchanges of methods, practices, approaches and databases between counties. Additionally, there is a strong correlation between a country and an athlete's performance directly related to the athlete selection process. As was mentioned by De Bosscher V. (2007), the larger population provides the bigger talent sample for recruitment opportunities for arranging trainings and competitions.

This study will specifically concentrate on predicting athletic performance by analysing anthropometric characteristics, including weight, height and body mass index. Also, relevant data such as age, gender and nationality for each athlete will be included into the analysis.

References

1. Anik, A. I., Yeaser, S., Hossain, A. I., & Chakrabarty, A. (2018). Player’s Performance Prediction in ODI Cricket Using Machine Learning Algorithms. Proc. 4th Int. Conf. Electr. Eng. Inf. Commun. Technol. (iCEEiCT), pp. 500-505.
2. Arkaev, L. I., & Suchilin, N. G. (2004). How to Create Champions. Oxford, Meyer & Meyer Sport.
3. Arneson, R. (2015). Equality of opportunity. In E. Zalta (Ed.), The Stanford encyclopedia of philosophy (Summer 2015 ed.). Retrieved from https://plato.stanford.edu/archives/sum2015/entries/equal-opportunity/
4. Bahr, R., Clarsen, B., & Ekstrand, J. (2018). Why we should focus on the burden of injuries and illnesses, not just their incidence. Br J Sports Med, 52, 1018–1021. https://doi.org/10.1136/bjsports-2017-098160
5. Bahr, R., & Krosshaug, T. (2005). Understanding injury mechanisms: a key component of preventing injuries in sport. Br J Sports Med, 39, 324–329. https://doi.org/10.1136/bjsm.2005.018341
6. Borms, J., & Hebbelinck, M. (1984). Review of Studies on Olympic Athletes. In J. E. L. Carter (Ed.), Physical Structure of Olympic Athletes Part 11: Kinanthropometry of Olympic Athletes (pp. 7-27). New York: Basel-Karger.
7. Carter, J. E. L. (1985). Morphological Factors Limiting Human Performance. In D. H. Clarke & H. M. Eckert (Eds.), Limits of Human Performance. American Academy of Physical Education Papers, No. 18. Champaign: Human Kinetics.
8. Cuk, I., & Karácsony, I. (2002). Rings. Methods, Ideas, Curiosities, History. Norman (OK), Paul Ziert & Associates.
9. De Bosscher, V. (2007). Sports Policy Factors Leading to International Sporting success. Published Doctoral thesis, Vrije Universiteit Brussel, Brussel, VUBPRESS.
10. Johnson, E. (Ed.). (2000). Essentials of Exercise Physiology (pp. 500-527). Lippincott Williams and Wilkins, USA.
11. Flegl, M., & Andrade, L. A. (2018). Measuring countries’ performance at the Summer Olympic Games in Rio 2016. OPSEARCH, 55, 823–846. https://doi.org/10.1007/s12597-018-0347-8
12. Heinilä, K. (1982). The totalization process in international sport. Sportwissenschaft, 12(3), 235–254.
13. Musch, J., & Grondin, S. (2001). Unequal competition as an impediment to personal development: a review of the relative age effect in sport.
14. Michalsik, L. B. (2004). Analysis of working demands of Danish handball players. In What’s Going Gym. Copenhagen, Denmark: Forlaget Underskoven, pp. 321-330.
15. Frantisek, T. (2011). Competitive Loading in Top Team Handball. Vienna, Austria: European Handball Federation WEB PERIODICAL.
16. Donoghue, P. O. (2009). Research Methods for Sports Performance Analysis. Evanston, IL, USA: Routledge.
17. Sekeroglu, B., Dimililer, K., & Tuncal, K. (2019). Artificial Intelligence in Education: Application in student performance evaluation. Dilemas Contemporaneos Educacion Politica Valores, 7(1), 1-21.
18. Meayk, E., & Unold, O. (2011). Machine learning approach to model sport training. Comput. Hum. Behav., 27(5), 1499-1506.
19. Ofoghi, B., Zeleznikow, J., Macmahon, C., & Dwyer, D. (2013). Supporting athlete selection and strategic planning in track cycling omnium: A statistical and machine learning approach. Inf. Sci., 233, 200-213.
20. Hore, S., & Bhattacharya, T. (2018). A machine learning based approach towards building a Sustainability Model for NBA players. Proc. 2nd Int. Conf. Inventive Commun. Comput. Technol. (ICICCT), pp. 1690-1694.
21. Musa, R. M., Majeed, A. P. P. A., Taha, Z., Chang, S. W., Nasir, A. F. A., & Abdullah, M. R. (2019). A machine learning approach of predicting high potential archers by means of physical fitness indicators. PLoS ONE, 14(1).
22. Musa, R. M., Majeed, A. P. A., Taha, Z., Abdullah, M. R., Maliki, A. B. H. M., & Kosni, N. A. (2019). The application of artificial neural network and k-nearest neighbour classification models in the scouting of high-performance archers from a selected fitness and motor skill performance parameters. Sci. Sports, 34, 241-249.
23. de Jesus, K., Ayala, H. V. H., de Jesus, K., Coelho, L. D. S., Medeiros, A. I. A., & Abraldes, J. A. (2018). Modelling and predicting backstroke start performance using non-linear and linear models. J. Hum. Kinetics, 61(1), 29-38.
24. Maanijou, R., & Mirroshandel, S. A. (2019). Introducing an expert system for prediction of soccer player ranking using ensemble learning. Neural Comput. Appl., 31(12), 9157-9174.
25. Zhou, Z., Shakya, S., & Sha, Z. (2017). Predicting countermovement jump heights by time domain frequency domain and machine learning algorithms. Proc. 10th Int. Symp. Comput. Intell. Des. (ISCID), pp. 167-170.