Machine Learning for Individualized Training Support in Marathon Running

Ladislav Havaš[1], Vladimir Medved[2] and Zoran Skočir[3]
[1]*Department of Electrical Engineering, University North, Varaždin, Croatia*
[2]*Faculty of Kinesiology, University of Zagreb, Zagreb, Croatia*
[3]*Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia*

1 OBJECTIVES

Modern technology significantly influences the field of human physical exercise monitoring and control. Assuming a cybernetic-like approach to sports training, we witness a dynamic, synergistic interaction of the man-technology system realizing a training process. In this paper we focus to actual implementation of Information and Communications Technology (ICT) for intended improvement of marathon runners' training. While general principles of sports training are known (Milanović, 2009), ICT offers profoundly novel and original possibilities to influence and actively control training process of a particular individual. First author's sports experience (Havaš and Vlahek, 2006) is combined with an approach based on telecommunication platform which was gradually built, upgraded and validated over the years (Havaš, at al., 2013) to produce an original, comprehensive and intelligent system suited to individualized use (Havaš, 2014).

Here we focus in particular to machine learning features of the approach enabling a flexible, on line physiologically based-monitoring training support system for marathon running. Developed, tested and validated on marathon runners' data, the system however posesses capabilities for application in sports training in general.

2 METHODS

Data mining and On Line Analytical Processing (OLAP) analyses are used, autonomously or on demand. Realized and „imagined" trainings are validated and user progress is analyzed so that signs of overtraining, undertraining or possible disease are recognized. Using available telecommunication channels a message is automatically generated and sent to user and new (corrected) program is created for remaining number of days of the training process.

Among many available methods and techniques of data mining, concepts of neural networks are used (Neural_network, available at: http://en.wikipedia.org/Neural_network, accessed on 02 December 2012, Neural_network_theory, available at: http://fann.sourceforge.net/report, accessed on 05 December 2012). Program modules are developed for analysis of runners that have their active programs, and offer to them possible corrections, periodically, after their registration to the system, or based on special request by each user. Combining methods and concepts of data mining and OLAP analysis (Online_analytical_processing, available at: http://en.wikipedia.org/wiki, accessed on 14 November 2012), a new method of validation is developed, defined and realized for trainings of long distance runners (Havaš, 2014).

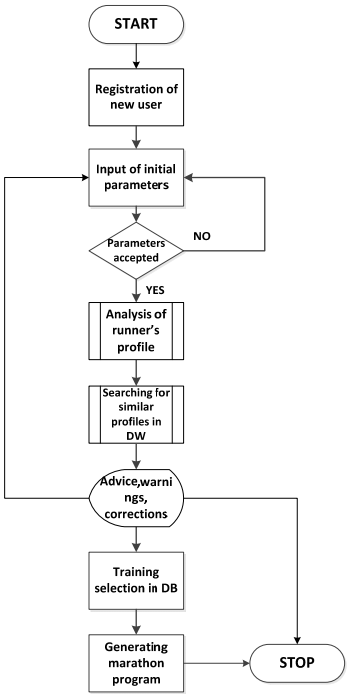

Figure 1: Searching for similar profiles.

Havaš L., Medved V. and Skočir Z..
Machine Learning for Individualized Training Support in Marathon Running.

Periodic control trainings are envisaged based on which one can monitor and calculate new values of maximum aerobic capacity $VO2_{max}$. Each positive (or negative) change is dynamically in calculated in all remaining training sessions, by means of which one attains improvement in remaining preparation period.

Developed prototype searches for similar runner profile patterns, and based on known history of their accomplishments forms and displays corresponding corrections, generating individualized program for preparation (Figure 1).

By filling data warehouse for even greater sample of runners in long periods, conditions are created for reliable prediction of future realizations and timely correction of preparation program.

3 RESULTS

The majority of tested marathon runners train many years and have achieved results in marathon, which were compared with the results that were achieved with the assistance of developed Online Running Trainer (ORT) system. The comparison between previous personal bests (achieved in the past three years) and those achieved by using the implemented system is shown in Figure 2.

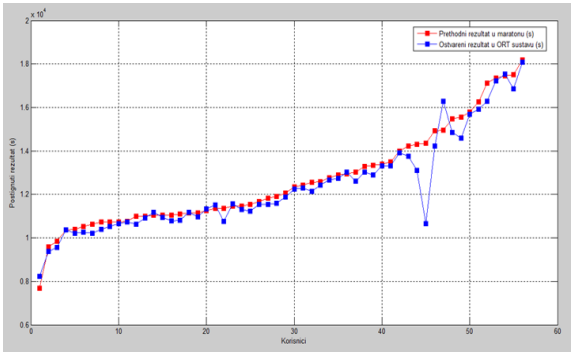

Figure 2: The comparison between previous and current results of tested users (in Croatian language).

X-axis of the Figure 2 represents tested users (58 marathon runners), while the y-axis represents the achieved results of those runners in seconds. The red curve in the Figure 2 shows previous results in marathon, while the blue curve shows achieved results with the ORT system. It can be seen that almost all tested users have improved their results and have successfully finished the training cycle without sport injuries. It is important to mention that they have not reached that result in several iterations (training cycles), but within one season in which they have been systematically guided.

4 DISCUSSION

Mentioned verification process has its flaws and was not carried out in a way which would (possibly) show more (or less) deviations from the anticipated results. Only one group of athletes (58 athletes) was tested, who used the developed system and ran on their key races. The results of those runners in the past three years, achieved without the help from the developed prototype were compared with the results with the assistance of ORT system, where there is only partially possible to compare the parameters of "different" groups of users who train in different ways. Because of the specificity of testing sport achievements, the comparison between the group of marathon runners and the group of people who do not actively engage in sport activities (running) would not have any practical value. Listed limitations of the conducted research should result in additional evaluations, for the purpose of obtaining quality and more certain judgments on a larger sample of athletes. Periodic loading of data warehouse (ETL process) will, in time, enable the evaluation of algorithms and methods on a larger data sample, while the developed ORT system will, by the method of self-learning, correct the training elements and generate better quality programs.

5 CONCLUSIONS

Machine learning means a computing method enabling future actions be improved based on information from past (Gamberger, 2011). Presented method, further, is based not only on the study of realization by one user, but the system learns as a whole based on registered data on remaining runners. This demonstrates that presented method successfully implements machine learning paradigm to the problem of improving sports training.

REFERENCES

Milanović, D., 2009., Theory and methodology of trainings *(in Croatian)*, Faculty of Kinesiology, University of Zagreb, Zagreb.

Havaš, L., Vlahek P., 2006. *Road running (in Croatian)*, TK Međimurje. Čakovec.

Havaš, L., Skočir, Z., Medved, V., 2013., *Modeling of the athlete's training decision support*, Vol. 20, pp. 315-322, March-April, Technical Gazette.

Havaš, L., 2014., Information system for athlete's training decision support *(in Croatian)*, Doctoral Thesis, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb

Havaš, L., Medved, V., Skočir, Z., 2013, *Application of mobile technologies in the preparations for long distance running*, icSPORTS 2013, Science and Technology Publications, pp.154-161., Lda, SCITEPRESS.

Neural_network, http://en.wikipedia.org/wiki

Neural_network_theory, http://fann.sourceforge.net/report

Online_analytical_processing,http://en.wikipedia.org/wiki

Gamberger, D., 2011, *Knowledge discovery by data mining (in Croatian)*, Institute Ruđer Bošković.