Chapter 20
Deep Analytics in Sport Community Forums

Dušan Fister
University of Maribor, Maribor, Slovenia

Iztok Fister
Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia

Iztok Fister Jr.
Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia

ABSTRACT

In recent years, some sport clubs have adopted web forums for online discussions about planning training sessions, races, club problems, sponsors and supporters, equipment and so on. Mostly, these forums are closed, because some discussions about critical information must be permitted only to registered club members. Indeed, various members are connected into a network representing a forum community that participates in forming the internal structure of a club and has taken an increasingly important role in solving the club affairs. This has influenced the classic way of doing business with meetings or calls and has allowed the community to participate in forming plans. This article deals with the deep analytics of data acquired from a web forum of a small cycling club located in Slovenia for a period of five years. The purpose of these analyses is to identify various members of the forum's community, track the dynamics of events as found in this forum, and search for hidden relationships.

1. INTRODUCTION

Sports training is a very complex process that usually involves many factors, e.g., coaches, cyclists, sports trainers, trainers' assistants and club officials (similar to team sports). In the past, sports training was performed slightly differently than nowadays, especially because of the rise of technological revolution in pervasive computing. The latter emerged due to the innovative way of thinking, which says that the information should be obtained from everywhere, anytime, for any objective. Monitoring and analyz-

DOI: 10.4018/978-1-7998-7707-3.ch020

ing various parameters during and after the sports training sessions therefore became a very important tool and the necessary step towards progression. However, cyclists must use various professional sport watches and a bunch of sensors that enable them to track and analyze the whole activity, as well as obtain vital parameters that play a big role in the cyclist's performance (Baca, 2017). A good example of this are power-meters, which are embedded into a bike's frame and are devoted to measure power during the cycling. After the training session, officials and cyclists can make deep analysis of the acquired data and determine how successfully the sports training session was performed and what went wrong. It is worth mentioning that not only professional cyclists use modern tech gadgets, but also amateur cyclists. However, the border line between professional and amateur cyclists is becoming thinner every day. Triathlon is an example of a sport where many amateurs behave like professionals.

In recent years, social scientists began to analyze the behavior and features of cyclists using deep analytics. The following three factors have given rise to the idea of a new paradigm in data processing (Henke et al., 2016):

- Data volumes have grown exponentially in the last century
- More sophisticated algorithms for data processing have been developed
- Computational power and storage have improved steadily

The first factor refers to data collection advances. Nowadays, new data sources allow the generation, calculation and storage of big data volume. These data sources produce data throughout internet searches, social media activities, online transactions, as well as sensors connected into the so-called Internet of Things (IoT) (Whitmore, Agarwal, & Da Xu, 2015). The second factor depends on the fact that the increased data volumes have demanded new self-learning algorithms, instead of the standard hard-coded algorithms. This class especially comprises machine learning algorithms (e.g., deep learning) that use an inductive approach to form a representation of the real world based on the seen data (Henke et al., 2016). The third factor, however, presents a basis for the first two. Computational power has been increased by introducing Graphical Processing Units (GPUs), enabling 10-30 times faster image processing in neural networks, than the traditional Central Processing Units (CPUs). The biggest step forward presents so-called cloud computing, where companies reserve as much computational power and storage volume as needed on the internet without buying any physical machine.

In summary, deep analytics is focused on the application of sophisticated data processing techniques to yield information from large distributed datasets that may contain the structured data, as well as unstructured and semi-structured data (Deep analytics, 2017). For instance, Roy et al. (2012) proposed a system for end-to-end processing of genomic data that are produced in large sequencing centers at the rate of 10 terabytes per day. Thus, massive amounts of raw data need to be transformed into biological information using complex processing that includes alignment of short read sequences, variation discovery and deep analytics. Sokol and Chan (2013) exposed the role of context-based analytics and argued that the application of the higher quality models to contextually- corrected data can lead to a better decision-making process, and, thus better outcomes. Fan and Bifet (2013) overviewed the topic of mining big data, its current status and a forecast. Chen and Zhang (2014) emphasized the valuable role of big data for predicting productivity in business and evolutionary breakthroughs in scientific disciplines. Big data explorations brought advances in business productivity and strengthened the competitive position of companies on the market. Najafabadi et al. (2016) explored how deep learning can be utilized for addressing some problems of data big analytics, including extracting complex patterns from massive amounts of data, semantic