11 Advanced Technology for Event Management

Hani Ragab and Mohamed Salama

Laerning outcomes

On completion of this chapter, the reader will be able to:

- ☐ Understand the basic concepts of data science and data analytics
- ☐ Compare and contrast different types of data.
- ☐ Understand the cross-industry process for data mining (CRISP-DM).
- ☐ Appreciate the application of advanced technologies in the context of event management.
- ☐ Understand the challenges associated with cyber security in events.

Introduction

Event management is a dynamic field that has always benefited from latest advances in technology. In this chapter we will review some of the newest and most promising fields in information technology and discuss how they could be used to support event managers.

Data is at the heart of information technology, in particular, *data science* aims to extract knowledge from data using machine learning techniques. The amount of data might make it not possible to process it on personal computers, leading to the field of *big data*. We will explore the fields of data science, big data as well as machine learning.

Stored and transiting data might hold high value that attracts cyber criminals, *information security* focuses on how to protect data from accidental release and tampering of data. Basic concepts of information security, particularly *cryptography*, had a major contribution in the creation of the new paradigm of *blockchains*.

Data science

Data science aims to extract knowledge from data. The data could be structured, semi-structured or unstructured. The techniques used to extract knowledge come from statistics, data visualization and machine learning. The fields related to the processing of data are quite interconnected and might be indistinguishable for the non-specialist. So, let us start by explaining the differences between a few keywords that are commonly used interchangeably, namely *data science*, *data mining*, *data analytics*, *business intelligence*, *machine learning* and *artificial intelligence*.

Data science and *data analytics* are quite close, in the sense that they apply similar techniques to extract knowledge from data. One simple difference is that data analytics focuses on answering concrete questions related to a business, whereas data science is the scientific field that empowers data analysts with tools that implement data science's innovations. One could see data science as the theoretical field, whereas data analytics is the application, in practice, of the results of data science.

Data analysts parse business data to draw charts, make predictions and produce actionable results. They look at the data with a pre-defined business-related goal, formulated as a question. Examples of business questions include "which individuals would be more susceptible to a mailing campaign about our new events?", "could past data provide us with a way to predict registrations for the next event?" and, for a telecommunication provider, "is it possible to provide customers who will churn?" Data analysts could also have roles where they maintain databases and data warehouses.

Data scientists can do the above and, using more advanced theoretical (mathematical, statistical and machine learning) and practical skills, come up with original solutions when the automated and common tools fail to produce desired results. Data scientists might work on solving abstract problems, without a direct link to a business problem, and they will be more tool-agnostic and tool-independent than a data analyst, whose job might be all around using one particular software to produce analyses.

Another field that is often confused with data science is *data mining*, aka *knowledge discovery from data* (KDD). There is no agreement on the difference (and even boundaries) of the two fields. It is quite common to find definitions that put data mining as part of data science. That is, data mining becomes that particular part of data science, that focuses on pulling the data and processing it to discover knowledge and apply that knowledge; while data science adds to these tasks many others, such as capturing of data and maintaining it. Other authors claim that data science is more of a business/media name of data mining.

While *statistics* and data science exhibit a few similarities, it is important to note that statistics focuses on data summarization and description, whereas data science focuses on higher level data manipulation, such as prediction and clustering. For example, statistics can tell us the average number of attendees of events and potentially identify basic features that lead to variations in attendance (e.g., venue). On the other hand, data science can be used to identify sentences in events description that led to higher attendance and can provide advanced models for the prediction of attendee numbers.

It is interesting to note that statistics has rigorous rules on how the sampling of data and its analysis are carried out, thus certain levels of confidence on the results can be established. This is not exactly the case in *machine learning* which uses available data to make insights from them. Businesses often must do their analyses with whatever data is available with no prerequisite of proper sampling techniques. Machine learning defines its own techniques for testing and validating results, e.g. it might validate outcomes using testing data or by the deployment of its outcomes in the field as a substitute for the theoretical guarantees provided by statistical techniques.

Machine learning aims to allow machines to improve their performance of a task, over time and by parsing more data (the learning). An older, but simpler, definition by Arthur Samuel describes machine learning as the field that enables computers to learn without being explicitly programmed (Samuel, 1959). The concepts of *artificial intelligence (AI)* and machine learning are quite intertwined; AI heavily uses machine learning techniques. For example, AI applications for image and video processing use deep learning techniques. AI applications on robotics also heavily rely on machine learning. AI does not limit itself to machine learning. For example,techniques like *minmax* and *alpha-beta pruning* are not machine learning techniques (at least for purists), but they are still widely accepted as AI ones. In fact, Deep Blue, the IBM machine that did beat Garry Kasparov in 1997 at chess, considered to be one of the biggest landmarks in AI's history, was a sophisticated implementation of the alpha-beta search algorithm.

It is interesting to note here that there is an intersection between data mining and machine learning since machine learning techniques are often used to learn from data. But this does not mean that either is a sub-discipline of the other. Indeed, machine learning applications are not limited to data mining, and data mining goes beyond machine learning by covering other disciplines such as data warehousing.

It is important to note that the above definitions and distinctions might not be widely used and respected in practice. A company might hire someone to act as a data analyst but gives them the title of data scientist or might