Service Analytics: Putting the "Smart" in Smart Services

Niklas Kühl, Hansjörg Fromm, Jakob Schöffer, and Gerhard Satzger

Abstract Artificial intelligence in general and the techniques of machine learning in particular provide many possibilities for data analysis. When applied to services, they allow them to become smart by intelligently analyzing data of typical service transactions, e.g., encounters between customers and providers. We call this service analytics. In this chapter, we define the terminology associated with service analytics, artificial intelligence, and machine learning. We describe the concept of service analytics and illustrate it with typical examples from industry and research.

1 Introduction

As outlined in the previous chapters, modern economies are becoming more and more "servitized"—with over 75% of the gross value added being derived from the tertiary sector (Eichengreen and Gupta 2011) and with an increasing number of industrial companies proceeding to engage in service-type offerings (Neely 2008).

A prominent theory in the field of services—while still being discussed controversially—is the so-called *Service-Dominant Logic* proposed by Vargo and Lusch (2008) that advocates the perspective that value is not "embedded" in products or services but is rather created by the knowledge, skills, and resources employed by both provider(s) and customer(s). The particular challenge then is the so-called *co-creation of value*, i.e., partners aiming at incorporating potential contributions from both sides to come up with a solution that—from an overall system point of view—maximizes the generated value. This goes far beyond the typical customer integration in a traditional service context in that it elevates the viewpoint above the simple provider perspective. Moreover, it opens the view to analyzing and purposefully designing more complex (smart) service systems comprising a larger number of stakeholders (Maglio et al. 2018).

N. Kühl (✉) · H. Fromm · J. Schöffer · G. Satzger
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
e-mail: niklas.kuehl@kit.edu; hansjoerg.fromm@kit.edu; jakob.schoeffer@kit.edu; gerhard.satzger@kit.edu

M. Maleshkova et al. (eds.), *Smart Service Management*,
https://doi.org/10.1007/978-3-030-58182-4_13

However, this poses a systemic challenge. Looking at traditional service systems, we realize that one of the key challenges is the availability of the knowledge, skills, and resources available at one integration point to connect the partners. This becomes evident looking at a simple example with data being one of the important resources to be shared: in a supply chain with different providers and their customers, everyone would benefit from understanding and reacting to changes in the production processes within the system. So far, however, individual processes might be measured in real time, but their resulting data are typically not communicated to other stakeholders. Similar situations can be found in other industries. Automotive manufacturers, car dealers, and vendors of value-added services around the car do not know much about the usage of their products and services by the customers. Doctors and other medical service providers do not know about their patients' behaviors and health status once they have left their offices.

As an instrument to overcome this disconnect, modern information and communication technology may assist helping to create a system-wide view and, thus, exploit the inherent potential (Böhmann et al. 2014). For instance, rather than manually reading error logs of machines in a supply chain once a year, smart services could provide regular, more frequent or even real-time updates. At the same time, this data could centrally be made available, and service providers could manage capacity to the advantage of all parties: lower maintenance prices for consumers, faster reaction times, and higher profits for providers. As we can observe, this disconnect is actually overcome more and more with the emergence of new measuring sensor technologies in the field of the Internet of Things (IoT) (Martin et al. in press). An increasing volume of data is already collected by the users/customers themselves (e.g., through smart devices) or by smart services in different fields, like energy services, telematics in automotive and mobility services, RFID in logistics, condition sensors in engineering, data capture solutions in healthcare, etc. The further dissemination of electronic networks, led by the Internet "revolution," will increasingly enable sharing of the captured data across organizational boundaries and support their availability at the point of decision.

Where data are available already today, the potential is clearly visible and is already being leveraged in smart services: by design, these services require connectivity between providers and customers. For instance, customers visit the provider's web pages in order to obtain its service. Thus, the provider is able to analyze the customers' usage characteristics at any level of detail. Typical data of interest are the overall number of page visits, the number of page visits per customer, the time intervals between page visits, the path that an individual customer takes through the website, etc. With these data, the provider can perfectly analyze the behavior and preferences of individual customers, can make personalized recommendations, can assess the general acceptance and attractiveness of the web offering, and can discover possible usability problems related to navigating and finding information on its web pages.

For this process of capturing, processing, and analyzing data taken from a service system—in order to improve, extend, and personalize the service and create new value for both the provider and the customer—we use the term *service analytics* (Fromm et al. 2012).

2 Analytics, Data Mining, Machine Learning, and Artificial Intelligence

But, what is *analytics*? There is no single agreed-upon definition of the term *analytics*. Some authors use the terms *analytics* and *data mining* interchangeably (Kohavi et al. 2002). Others use *analytics* as a synonym for business intelligence (Davenport and Harris 2017). With the rise of artificial intelligence and machine learning, additional concepts are added to this nomenclature, calling for clear definitions of these terms and their interplay (Kühl et al. 2019).

Figure 1 and the terms defined within this section lay the foundation of our understanding of service analytics and the related concepts. However, the overall terminology and the concepts' relationships are discussed controversially (Emmert-Streib and Dehmer 2009).

Traditionally, some dissent is particularly related to the question if analytics should include or exclude data management and reporting technologies. Davenport and Harris (2017) distinguish between "access and reporting" and "analytics," both seen as subsets of business intelligence. Data management and reporting are often considered as basic analytics, which are a prerequisite for advanced analytics, built on methods from statistics and operations research. Recently, however, discussions have been focusing more on techniques labeled as *data mining* and *machine learning*—or *artificial intelligence* as an umbrella term. Not only are the terms *analytics*, *machine learning*, *artificial intelligence*, *data mining*, *deep learning*, and *statistical learning* related, but they also often appear in the same context and are sometimes used interchangeably. While the terms are common in different communities, their particular usage and meaning vary widely.

In the field of statistics, for instance, the focus lies on *statistical learning*, which is defined as a set of methods and algorithms to gain knowledge, predict

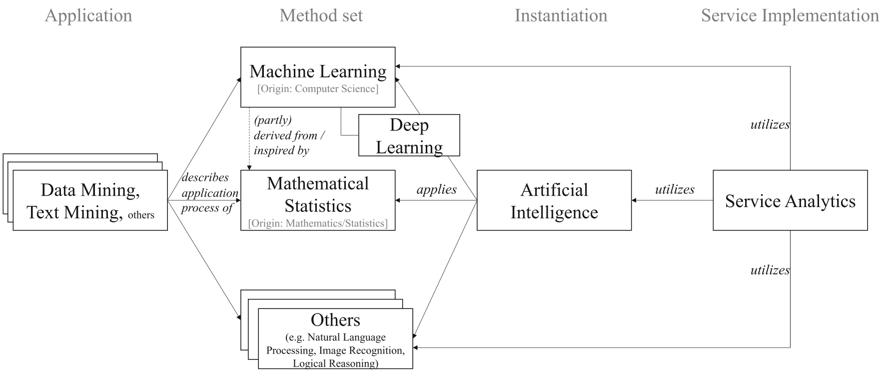

Fig. 1 Overview of terminology

outcomes, and make decisions by constructing models from a data set (Hastie et al. 2005). From a statistics point of view, machine learning can be regarded as an implementation of statistical learning (Bousquet et al. 2011).

Within the field of computer science, *machine learning* is focused on designing efficient algorithms to solve problems with computational resources (Mohri et al. 2012). While machine learning utilizes statistical approaches, it also includes methods not entirely based on statisticians' previous work, resulting in new and well-cited contributions to the field (Huang et al. 2004; Sebastiani 2002). Generally speaking, we can think of machine learning as a set of different tools used to derive meaning from data in an automated fashion. These tools are referred to as machine learning models—specific algorithms that usually take in large amounts of collected data and, through certain mathematical computations (training), accomplish learning general relationships or patterns in said data. There are several, fundamentally different types of machine learning, based on various scenarios that can occur. Three very important types are supervised learning, unsupervised learning, and reinforcement learning.

Supervised learning deals with the so-called labeled data. This means that a supervised model is trained on data that include values for a given target variable (Kühl et al. 2020). Here is an example: Assume we have collected data from a sensor attached to a component of an industrial machine. At a given time, the sensor measures temperature and pressure of the component. We usually call these two variables input features or independent variables. For each collected value pair of temperature and pressure, we also know whether the component was intact or defective; this is our (binary) target variable—or label. The label is also often referred to as the dependent variable. Having collected these data over the course of some time, we are now able to train a machine learning model that can learn a relationship (if any) between temperature, pressure and whether the component is defective or intact. Based on its learning, the model is ultimately able to tell us whether a new, previously unseen combination of pressure and temperature values would rather correspond to a defective or an intact component. An exemplary popular supervised learning algorithm to accomplish this is logistic regression (Nelder and Wedderburn 1972).

Unsupervised learning does not have this target variable, or label, in the data. Instead, imagine a scenario where we have collected consumer data, including age, income, gender, education, etc. with the goal of performing a demographic market segmentation. These insights could ultimately be used to provide personalized online shopping recommendations. In this case, we are not interested in finding a relationship between input features and some target variable, but rather in finding new patterns in the data, i.e., disjoint clusters of customers, such that similar customers belong to the same cluster. In machine learning taxonomy, this approach is called cluster analysis, one of the key unsupervised learning techniques. An exemplary algorithmic implementation of cluster analysis is k-means clustering (Lloyd 1982).

The third of the three main pillars of machine learning is *reinforcement learning* (LeCun et al. 2015). This relatively young discipline started to get a lot of attention after DeepMind's AlphaGo implementation became capable of defeating the world champion in the game of Go (Silver et al. 2017). In brief, reinforcement learning follows a trial and error logic, trying to mimic human learning behavior. Through performing correct actions, i.e., actions that lead to some predefined success (e.g., not losing points in a video game), the algorithm receives a reward and thus learns to distinguish right from wrong. A problem with supervised learning, for example, is that correct relationships between independent and dependent variables are assumed to be known ex ante. However, in the game of Go, e.g., it is intractable to define the full set of correct actions given any specific game scenario. Reinforcement learning is trying to overcome this by learning "on the fly." It is expected that reinforcement learning will have a significant impact on a wide range of real-world applications, such as self-driving cars, robotics, education, etc. (Chollet 2017). A technique often associated with reinforcement learning is called *deep learning* (Goodfellow et al. 2016).

Deep learning has become increasingly popular in machine learning over the past years (LeCun et al. 2015). Generalizing the idea of the so-called *artificial (feed-forward) neural networks* (Basheer and Hajmeer 2000), deep learning models comprise multiple processing layers capable of learning complex data representations with multiple levels of abstraction. Deep learning has drastically improved machine learning's capabilities, for example, with regard to speech (Hinton et al. 2012) and image recognition (He et al. 2016). Despite their superior performance in certain areas, and several breakthroughs in the past, such as Krizhevsky et al. (2012), deep learning models remain challenging to interpret. This is why they are sometimes also referred to as "black box models" (Shwartz-Ziv and Tishby 2017).

Contrary to the above terms, *data mining* describes the process of applying quantitative analytical methods, which help solve real-world problems, for example, in business settings (Schommer 2008). From a machine learning perspective, data mining is the process of generating meaningful machine learning models. The goal is not to develop more knowledge about machine learning algorithms, but to apply them to data in order to gain insights and potentially derive certain actions. Machine learning can therefore be regarded as a basis for data mining (Witten et al. 2011).

In contrast, *artificial intelligence* applies techniques like mathematical statistics, machine learning, natural language processing or image recognition to mimic human intelligence, such as common sense reasoning (Nilsson 2014), in machines. More generally, it can be regarded as an umbrella term for any method with the ultimate goal of achieving machine intelligence.

Service analytics, eventually, applies techniques from all these fields, including machine learning, to improve, extend, and personalize a service, creating added value for both service providers and customers. These enhanced services can themselves be—or at least contain—analytics (e.g., "analytics-as-a-service" Delen and Demirkan 2013).

3 Practical Examples of Service Analytics

In this chapter, we will give some exemplary real-world examples of smart services, which are based on contributions to the minitrack "Service Analytics" from the Hawaii International Conference on System Sciences (HICCS). Other examples can be found in a variety of fields, ranging from industrial manufacturing and mobility to social sciences or health care.

Schoch et al. (2017) propose a paper, in which they propose a way to efficiently collect sensor data in electric vehicles, in order to analyze driving behavior and derive insights around battery degradation. The authors argue that the same sensor data can also be used to improve fleet allocation for car sharing providers or to implement predictive maintenance strategies, among others, both benefiting the end user (increased car availability) and the provider (cost savings through preventive maintenance).

Another example of smart services stems from a paper by Laubis et al. (2017). In this work, the authors describe a machine learning approach for estimating road roughness through smartphone-equipped passenger cars. This allows near real-time road condition monitoring and can benefit road users by warning them of hazardous situations, recommending appropriate driving behavior, or suggesting alternative routes altogether.

A smart service in the field of industrial maintenance is introduced by Wolff et al. (2018). Here, the authors propose the implementation of a technician marketplace that can be accessed by industrial maintenance customers to book technician capacity. They argue that both traditional pricing mechanisms and current dispatching of service workers are inefficient. The newly proposed simulation-based approach allows customers to book technician capacity for fixed time slots, while the price per slot is dynamic, depending on the remaining capacity. That way, the authors claim, customers are incentivized to buy slots in accordance with their objective task urgency, increasing the overall system efficiency.

In the automotive aftermarket domain, Steuer et al. (2018) propose a novel method for inventory planning of spare parts, based on clustering and classification techniques. The authors argue that this approach is particularly well suited for demand forecasting of new parts, where historical demand patterns might not be readily available. More accurate predictions of spare part demands are imperative for stock optimization in a market worth almost €1.0 trillion (Heid et al. 2018).

Steins et al. (2019) propose an approach to forecast the demand of emergency medical services in several Swedish counties. Being able to accurately forecast this demand can "help in providing quick and efficient medical treatment and transportation of out-of-hospital patients," as the authors of this paper phrase it. In addition to historical demands, they incorporate socioeconomic data as well as weather, time, traffic, events, and related information in their model. That way, the proposed method is able to outperform the traditional forecasting practice in place.

Kisore and Reddy (2015) conduct an empirical study to identify and evaluate relationships between demographic and other socioeconomic data on the one hand,

and people's preferences for ATM locations on the other. Based on their findings, the authors then propose a better informed decision-making process regarding ATM location planning for banks in India.

4 Conclusion

In today's connected world, large amounts of data are available and continue to grow every second. While the data can have many origins and purposes, a share is generated from (regular) service operations between providers and customers. Many of these interactions already capture meaningful information that can be utilized to generate useful knowledge as a basis for future decision-making.

The concept of "service analytics" provides researchers and practitioners with different techniques to uncover patterns and insights from data sets from the service space. On the basis of these findings, more sophisticated decisions can be made, and the results can be leveraged to understand multiple perspectives of service operations, which can then lead to further improvements, e.g., by delivering services more efficiently or increasing customer satisfaction.

References

Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods, 43*(1), 3–31.

Böhmann, T., Leimeister, J. M. & Möslein, K. (2014). Service-systems-engineering. *Wirtschaftsinf, 56*, 83–90 (2014).

Bousquet, O., von Luxburg, U., & Rätsch, G. (2011). *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, February 2–14, 2003, Tübingen, Germany, August 4–16, 2003, Revised Lectures* (Vol. 3176). Berlin: Springer.

Chollet, F. (2017). *Deep learning with python* (1st ed.). Greenwich, CT: Manning Publications.

Davenport, T., & Harris, J. (2017). *Competing on analytics: Updated, with a new introduction: The new science of winning*. Brighton, MA: Harvard Business Press.

Delen, D., & Demirkan, H. (2013). Data, information and analytics as services.

Eichengreen, B., & Gupta, P. (2011). The two waves of service-sector growth. *Oxford Economic Papers, 65*(1), 96–123.

Emmert-Streib, F., & Dehmer, M. (2009). *Information theory and statistical learning*. Berlin: Springer.

Fromm, H., Habryn, F., & Satzger, G. (2012). Service analytics: leveraging data across enterprise boundaries for competitive advantage. In: U. Bäumer, P. Kreutter, W. Messner (Eds.), *Globalization of professional services: Innovative strategies, successful processes, inspired talent management, and first-hand experiences* (pp. 139–149). Berlin: Springer.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.

Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: Data mining, inference and prediction. *The Mathematical Intelligencer, 27*(2), 83–85.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

Heid, B., Huth, C., Kempf, S., & Wu, G. (2018). Ready for inspection: The automotive aftermarket in 2030. Technical report, McKinsey & Company.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-R., Jaitly, N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine, 29*(6), 82–97.

Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2004). Extreme learning machine: a new learning scheme of feedforward neural networks. In *Proceedings. 2004 IEEE International Joint Conference on Neural Networks, 2004* (Vol. 2, pp. 985–990). Piscataway, NJ: IEEE.

Kisore, N. R., & Reddy, P. G. (2015). Empirical determination and evaluation of factors that impact ATM placement. In *2015 48th Hawaii International Conference on System Sciences* (pp. 1341–1348). Piscataway, NJ: IEEE.

Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). Emerging trends in business analytics. *Communications of the ACM, 45*(8), 45–48.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).

Kühl, N., Goutier, M., Hirt, R., & Satzger, G. (2019). Machine learning in artificial intelligence: Towards a common understanding. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Kühl, N., Hirt, R., Baier, L., Schmitz, B., Satzger, G. (2020). How to conduct rigorous supervised machine learning in information systems research: The supervised machine learning reportcard, communications of the association for information systems.

Laubis, K., Simko, V., Schuller, A., & Weinhardt, C. (2017). Road condition estimation based on heterogeneous extended floating car data. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.

LeCun, Y. A., Bengio, Y., & Hinton, G. E. (2015). Deep learning. *Nature, 521*(7553), 436–444.

Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Trans. Information Theory, 28*, 129–136.

Maglio, P. P., Kieliszewski, C. A., Spohrer, J. C., Lyons, K., Patrício, L., & Sawatani, Y. (2018). *Handbook of service science* (Vol. II). Berlin: Springer.

Martin, D., Kühl, N., von Bischhoffshausen, J. K., & Satzger, G. (in press). System-wide learning in cyber-physical service systems: A research agenda. In *Proceedings of the 15th International Conference on Design Science Research in Information Systems and Technology (DESRIST 2020)*, Kristiansand, Norwegen, November 2020.

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). *Foundations of machine learning*. Cambridge, MA: MIT Press.

Neely, A. (2008). Exploring the financial consequences of the servitization of manufacturing. *Operations Management Research, 1*(2), 103–118.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General), 135*(3), 370–384.

Nilsson, N. J. (2014). *Principles of artificial intelligence*. Burlington, MA: Morgan Kaufmann.

Schoch, J., Staudt, P., and Setzer, T. (2017). Smart data selection and reduction for electric vehicle service analytics.

Schommer, C. (2008). An unified definition of data mining. Preprint arXiv:0809.2696

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys, 34*(1), 1–47.

Shwartz-Ziv, R., & Tishby, N. (2017). Opening the black box of deep neural networks via information. arXiv preprint arXiv:1703.00810.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., et al. (2017). Mastering the game of go without human knowledge. *Nature, 550*, 354–359.

Steins, K., Matinrad, N., & Granberg, T. (2019). Forecasting the demand for emergency medical services. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.

Steuer, D., Hutterer, V., Korevaar, P., & Fromm, H. (2018). A similarity-based approach for the all-time demand prediction of new automotive spare parts. In: *Proceedings of the 51st Hawaii International Conference on System Sciences*

Vargo, S. L., & Lusch, R. F. (2008). Service-dominant logic: Continuing the evolution. *Journal of the Academy of Marketing Science, 36*(1), 1–10.
Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques* (3rd ed., Vol. 54). Burlington, MA: Morgan Kaufmann
Wolff, C., Vössing, M., Schmitz, B., & Fromm, H. (2018). Towards a technician marketplace using capacity-based pricing. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.