Literature Survey

Team ID : PNT2022TMID31896
Team Leader : Vishnu Kurup
Team Member : Vishnu Surya P
Team Member : Vishnu S
Team Member : Tamilselvan S
Team Member : Jagathesh S

Global Sales Data Analytics

Data Analytics

Authors : 1. Nada Elgendy , University of Oulu.

2.Ahmed Elragal, Luleå University of Technology .

Data Analytics :

The term "Data Analytics " has recently been applied to datasets that grow so large that they become awkward to work with using traditional database management systems. They are data sets whose size is beyond the ability of commonly used software tools and storage systems to capture, store, manage, as well as process the data within a tolerable elapsed time. Big data sizes are constantly increasing, currently ranging from a few dozen terabytes (TB) to many petabytes (PB) of data in a single data set. Consequently, some of the difficulties related to big data include capture, storage, search, sharing, analytics, and visualizing. Today, enterprises are exploring large volumes of highly detailed data so as to discover facts they didn't know before Hence, big data analytics is where advanced analytic techniques are applied on big data sets. Analytics based on large data samples reveals and leverages business change. However, the larger the set of data, the more difficult it becomes to manage

Data Storage and Management :

One of the first things organizations have to manage when dealing with big data, is where and how this data will be stored once it is acquired. The traditional methods of structured data storage and retrieval include relational databases, data marts, and data warehouses. The data is uploaded to the storage from operational data stores using Extract, Transform, Load (ETL), or Extract, Load, Transform (ELT), tools which extract the data from outside sources, transform the data to fit operational needs, and finally load the data into the database or data warehouse. Thus, the data is cleaned, transformed, and catalogued before being made available for data mining and online analytical functions . However, the big data environment calls for Magnetic, Agile, Deep (MAD) analysis skills, which differ from the aspects of a traditional Enterprise Data Warehouse (EDW) environment. First of all, traditional EDW approaches discourage the incorporation of new data sources until they are cleansed and integrated. Due to the ubiquity of data nowadays, big data environments need to be magnetic, thus attracting all the data sources, regardless of the data quality [5]. Furthermore, given the growing numbers of data sources, as well as the sophistication of the data analyses, big data storage should allow analysts to easily produce and adapt data rapidly. This requires an agile database, whose logical an physical contents can adapt in sync with rapid data evolution [. Finally, since current data analyses use complex statistical methods, and analysts need to be able to study enormous datasets by drilling up and down, a big data repository also needs to be deep, and serve as a sophisticated algorithmic runtime engine. lel Processing (MPP) databases for providing high query performance and platform scalability, to non-relational or in-memory databases, have been used for big data. Non-relational databases, such as Not Only SQL (NoSQL), were developed for storing and managing unstructured, or non-relational, data. NoSQL databases aim for massive scaling, data model flexibility, and simplified application development and deployment. Contrary to relational databases, NoSQL databases separate data management and data storage. Such databases rather focus on the high - performance scalable data storage , and allow data management tasks to be written in the application layer instead of having it written in databases specific languages. Alternatively, Hadoop is a framework for performing big data analytics which provides reliability, scalability, and manageability by providing an implementation for the MapReduce paradigm, which is discussed in the following section, as well as gluing the storage and analytics together. Hadoop consists of two main components: the HDFS for the big data storage, and MapReduce for big data analytics [9]. The HDFS storage function provides a redundant and reliable distributed file system, which is optimized for large files, where a single file is split into blocks and distributed across cluster nodes. Additionally, the data is protected among the nodes by a replication mechanism, which ensures availability and reliability despite any node failures [3]. There are two types of HDFS nodes: the Data Nodes and the Name Nodes. Data is stored in replicated file blocks across the multiple Data Nodes, and the Name Node acts as a regulator between the client and the Data Node, directing the client to the particular Data Node which contains the requested data.

Data Analytic Processing :

After the big data storage, comes the analytic processing. According to [10], there are four critical requirements for big data processing. The first requirement is fast data loading. Since the disk and network traffic interferes with the query executions during data loading, it is necessary to reduce the data loading time. The second requirement is fast query processing. In order to satisfy the requirements of heavy workloads and real-time requests, many queries are response-time critical. Thus, the data placement structure must be capable of retaining high query processing speeds as the amounts of queries rapidly increase. Additionally, the third requirement for big data processing is the highly efficient utilization of storage space. Since the rapid growth in user activities can demand scalable storage capacity and computing power, limited disk space necessitates that data storage be well managed during processing, and issues on how to store the data so that space utilization is maximized be addressed. Finally, the fourth requirement is the strong adaptivity to highly dynamic workload patterns. As big data sets are analyzed by different applications and users, for different purposes, and in various ways, the underlying system should be highly adaptive to unexpected dynamics in data processing, and not specific to certain workload patterns

Map Reduce is a parallel programming model, inspired by the “Map” and “Reduce” of functional languages, which is suitable for big data processing. It is the core of Hadoop, and performs the data processing and analytics functions [6]. According to EMC, the MapReduce paradigm is based on adding more computers or resources, rather than increasing the power or storage capacity of a single computer; in other words, scaling out rather than scaling up [9]. The fundamental idea of MapReduce is breaking a task down into stages and executing the stages in parallel in order to reduce the time needed to complete the task

Data Analytics and Decision Making :

Data is becoming an increasingly important asset for decision makers. Large volumes of highly detailed data from various sources such as scanners, mobile phones, loyalty cards, the web, and social media platforms provide the opportunity to deliver significant benefits to organizations. This is possible only if the data is properly analyzed to reveal valuable insights, allowing for decision makers to capitalize upon the resulting opportunities from the wealth of historic and real-time data generated through supply chains, production processes, customer behaviors, etc. Moreover, organizations are currently accustomed to analyzing internal data, such as sales, shipments, and inventory. However, the need for analyzing external data, such as customer markets and supply chains, has arisen, and the use of big data can provide cumulative value and knowledge. With the increasing sizes and types of unstructured data on hand, it becomes necessary to make more informed decisions based on drawing meaningful inferences from the data Accordingly, [8] developed the B-DAD framework which maps big data tools and techniques, into the decision making process [8]. Such a framework is intended to enhance the quality of the decision making process in regards to dealing with big data. The first phase of the decision making process is the intelligence phase, where data which can be used to identify problems and opportunities is collected from internal and external data sources. In this phase, the sources of big data need to be identified and the data needs to be gathered from different sources, processed, stored, and migrated to the end user. Such big data needs to be treated accordingly, so after the data sources and types of data required for the analysis are defined, the chosen data is acquired and stored in any of the big data storage and management tools previously discussed After the big data is acquired and stored, it is then organized, prepared, and processed, This is achieved across a high-speed network using ETL/ELT or big data processing tools, which have been covered in the previous sections.

Consequently, the following phase in the decision making process is the choice phase, where methods are used to evaluate the impacts of the proposed solutions, or courses of action, from the design phase. Finally, the last phase in the decision making process is the implementation phase, where the proposed solution from the previous phase is implemented. As the amount of big data continues to exponentially grow, organizations throughout the different sectors are becoming more interested in how to manage and analyze such data. Thus, they are rushing to seize the opportunities offered by big data, and gain the most benefit and insight possible, consequently adopting big data analytics in order to unlock economic value and make better and faster decisions. Therefore, organizations are turning towards big data analytics in order to analyze huge amounts of data faster, and reveal previously unseen patterns, sentiments, and customer intelligence. This section focuses on some of the different applications, both proposed and implemented, of big data analytics, and how these applications can aid organizations across different sectors to gain valuable insights and enhance decision making. According to Manyika et al.'s research, big data can enable companies to create new products and services, enhance existing ones, as well as invent entirely new business models. Such benefits can be gained by applying big data analytics in different areas, such as customer intelligence, supply chain intelligence, performance, quality and risk management and fraud detection . Furthermore, Cebr's study highlighted the main industries that can benefit from big data analytics, such as the manufacturing, retail, central government, healthcare, telecom, and banking industries.

Customer Intelligence :

Big data analytics holds much potential for customer intelligence, and can highly benefit industries such as retail, banking, and telecommunications. Big data can create transparency, and make relevant data more easily accessible to stakeholders in a timely manner. Big data analytics can provide organizations with the ability to profile and segment customers based on different socioeconomic characteristics, as well as increase levels of customer satisfaction and retention . This can allow them to make more informed marketing decisions, and market to different segments based on their preferences along with the recognition of sales and marketing opportunities . Moreover, social media can be used to inform companies what their customers like, as Additionally, using SNAs to monitor customer sentiments towards brands, and identify influential individuals, can help organizations react to trends and perform direct marketing. Big data analytics can also enable the construction of predictive models for customer behavior and purchase patterns, therefore raising overall profitability [4]. Even organizations which have used segmentation for many years are beginning to deploy more sophisticated big data techniques, such as real-time microsegmentation of customers, in order to target promotions and advertising . Consequently, big data analytics can benefit organizations by enabling better targeted social influencer marketing, defining and predicting trends from market sentiments, as well as analyzing and understanding churn and other customer behaviors.

Quality Management and Improvement :

Especially for the manufacturing, energy and utilities, and telecommunications industries, big data can be used for quality management, in order to increase profitability and reduce costs by improving the quality of goods and services provided. For example, in the manufacturing process, predictive analytics on big data can be used to minimize the performance variability, as well as prevent quality issues by providing early warning alerts. This can reduce scrap rates, and decrease the time to market, since identifying any disruptions to the production process before they occur can save significant expenditures . Additionally, big data analytics can result in manufacturing lead improvements . Furthermore, real-time data analyses and monitoring of machine logs can enable managers to make swifter decisions for quality management. Also, big data analytics can allow for the real-time monitoring of network demand, in addition to the forecasting of bandwidth in response to customer behavior. Additionally, the quality of citizens' lives can be improved through the utilization of big data. For healthcare, sensors can be used in hospitals and homes to provide the continuous monitoring of patients, and perform real-time analyses on the patient data streaming in. This can be used to alert individuals and their health care providers if any health anomalies are detected in the analysis, requiring the patient to seek medical help [22]. Patients can also be monitored remotely to analyze their adherence to their prescriptions, and improve drug and treatment options.

Risk Management and Fraud Detection :

Industries such as investment or retail banking, as well as insurance, can benefit from big data analytics in the area of risk management. Since the evaluation and bearing of risk is a critical aspect for the financial services sector, big data analytics can help in selecting investments by analyzing the likelihood of gains against the likelihood of losses. Additionally, internal and external big data can be analyzed for the full and dynamic appraisal of risk exposures . Accordingly, big data can benefit organizations by enabling the quantification of risks . High-performance analytics can also be used to integrate the risk profiles managed in isolation across separate departments, into enterprise wide risk profiles. This can aid in risk mitigation, since a comprehensive view of the different risk types and their interrelations is provided to decision makers.

Conclusion :

In this research, we have examined the innovative topic of big data, which has recently gained lots of interest due to its perceived unprecedented opportunities and benefits. In the information era we are currently living in, voluminous varieties of high velocity data are being produced daily, and within them lay intrinsic details and patterns of hidden knowledge which should be extracted and utilized. Hence, big data analytics can be applied to leverage business change and enhance decision making, by applying advanced analytic techniques on big data, and revealing hidden insights and valuable knowledge. Accordingly, the literature was reviewed in order to provide an analysis of the big data analytics concepts which are being researched, as well as their importance to decision making. Consequently, big data was discussed, as well as its characteristics and importance. Moreover, some of the big data analytics tools and methods in particular were examined. Thus, big data storage and management, as well as big data analytics processing were detailed. In addition, some of the different advanced data analytics techniques were further discussed.By applying such analytics to big data, valuable information can be extracted and exploited to enhance decision making and support informed decisions. Consequently, some of the different areas where big data analytics can support and aid in decision making were examined. It was found that big data analytics can provide vast horizons of opportunities in various applications and areas, such as customer intelligence, fraud detection, and supply chain management. Additionally, its benefits can serve different sectors and industries, such as healthcare, retail, telecom, manufacturing, etc. Accordingly, this research has provided the people and the organizations with examples of the various big data tools, methods, and technologies which can be applied. This gives users an idea of the necessary technologies required, as well as developers an idea of what they can do to provide more enhanced solutions for big data analytics in support of decision.

Predictive Sales Analytics

Authors : 1. Johannes Habel Associate Professor of Marketing University of Houston

2. Sascha Alavi Professor of Sales Management and Chair of the Sales Marketing Department University of Bochum
3. Nicolas Heinitz Research Associate University of Bochum

Introduction :

Due to the pervasive data ubiquity, sales practice is moving rapidly into an era of predictive analytics, using quantitative methods including machine learning algorithms to reveal unknown information such as customers' personality, value, or churn probabilities. However, many sales organizations face severe difficulties when implementing predictive analytics applications. This article elucidates these difficulties by developing the PSAA Model—a conceptual framework that explains how predictive sales analytics applications support sales employees' job peformance. Specifically, the PSAA Model posits that predictive sales analytics applications only improve job performance if (1) sales employees adopt these applications to revise their decision-making, and (2) these updates inherently improve the decision outcome. The mechanisms underlying these two preconditions fundamentally differ. While the former is explained by well-established technology adoption theories, the extent to which adoption improves decision-making is determined by the value potential in the PSA application and the decision-making environment. Thereby, this paper provides a theoretical frame for future studies on predictive sales analytics.

Predictive Analytics :

The sales function is a key focal area for firms' digital transformation (Alavi and Habel 2021). Owing to the increasing adoption of CRM systems and the high quantifiability of sales performance, sales managers are particularly interested in advancing their decision-making through analytics (Kelly 2017), and, more specifically, predictive analytics. When employing predictive sales analytics (PSA), managers aim to support sales employees' decision-making by providing them with information that they do not have, statistically predicted from information that they do have. As such, predictive analytics goes beyond the widely employed descriptive analysis of past developments and instead aims at estimating new observations (Shmueli and Koppius 2011; Wedel and Kannan 2016), for example, lead conversion likelihoods, cross- and upselling potentials, or customer churn (Antonio 2018; Bock et al. 2017; Hatami et al. 2015). By implementing PSA, organizations aim to improve sales managers' and salespeople's decisionmaking, for example, when building stronger customer relationships and increasing company revenue.

However, companies face continuous difficulties in the implementation of PSA applications (Kelly 2017; Sleep et al. 2019; Ascarza et al. 2021; HBR 2021). Many sales employees harbour deep fears when asked to work with predictive analytics (Ammanath et al. 2020) and lack the necessary abilities to effectively apply the new tools (HBR 2021). Consequently, all too often firms and sales employees alike tend not to leverage the full potential and benefits from adopting such tools. For instance, Luo et al. (2021) found that sales agents show aversion to receiving feedback from AI-based sales coaches which undermines productivity improvements from the coaching, while Kim et al. (2022) found that service employees refuse to rely on AI assistance.

Analytics :

Academic literature uses various analytics terminology, such as business intelligence, big data analytics, and business analytics (Davenport 2006; Holsapple et al. 2014; Chen et al. 2015). Although extant studies define and position each of these terms slightly differently, the terms have two recurring themes in common: first, the use of data as a source for statistical analysis and second, the objective of enhancing decision-making (Sleep et al. 2019). For example, Grover et al. (2018, p. 390) define big data analytics as "the application of statistical, processing, and analytics techniques to big data for advancing business." In a similar vein, Holsapple et al. (2014, p. 134) consider business analytics as "evidence-based problem recognition and solving that happens within the context of business situations." Analytics has developed considerably over the years mainly driven by increasing data quality (Wedel and Kannan 2016; Müller et al. 2018), enhanced technological opportunities (Wedel and Kannan 2016), and growing analytics capabilities within organizations (Kiron et al. 2011). In the past, managers' focus was mainly on descriptive analytics, that is, the summarizing of historic data to investigate what happened in the past (Morgan 2019a). For example, sales managers tracked sales performance per region and salesperson via monthly reports (Morgan 2019a). More recently, predictive models to develop forecasts and simulations of variables have grown in significance (Wedel and Kannan 2016; Shmueli and Koppius 2011; Sleep et al. 2019). An example of predictive analytics is the estimation of the likelihood of a successful customer win-back (Gerpott and Ahmadi 2015) or predicting a customers' cross-buying likelihood several months in advance. Moreover, pertinent research often views prescriptive analytics as the next, consecutive step in the development cycle of analytics applications, following predictive analytics (Huang and Rust 2018). Prescriptive analytics augments predictive analytics by adding precise behavioral guidance and recommendations on how to act on the predictions. For instance, a prescriptive analytics application like Showpad may not only analyze and uncover customers' psychological preferences but even recommend a communication strategy to effectively address the preferences. Naturally, such prescriptions may constitute a severe intervention in salespeople's work processes and autonomy, enhancing the risk of salespeople's aversion. Since numerous companies are still at a very early stage and struggle with the effective implementation of predictive analytics (Alavi and Habel 2021), in this paper we focus more on predictive and less on prescriptive sales analytics.

Predictive sales analytics (PSA) :

As explained previously, the sales function is a key target of companies' predictive analytics endeavors—we label predictive analytics targeted at the sales function predictive sales analytics (PSA). That is, for variables relevant for sales employees, PSA quantitatively estimates values for observations not incorporated in the dataset on which the estimation is based. Put differently, when employing PSA, managers aim to support sales employees' decision-making by providing them with information they do not have, statistically predicted from information that they do have. To elaborate, the sales function is responsible for acquiring as well as retaining customers and is thus essential for a firm's success. As the firm's main decision-makers (Johnston and Marshall 2016), sales managers recruit potential candidates, train new entrants, and lead individual salespeople (Homburg et al. 2011; Albers and Krafft

2013). PSA can support these sales managers for example by forecasting sales to validate their budget plans (e.g., Pavlyshenko 2019) or by predicting employee turnover to improve their retention efforts (e.g., Bridges et al. 2007). Salespeople in turn are often key representatives of the firm, especially in the B2B context. Their decisions pertain to the various stages of customer acquisition and customer development (Albers and Krafft 2013; Söhnchen and Albers 2010). Similarly, PSA can support salespeople for example by predicting customers' conversion likelihoods enabling them to prioritize potential customers (e.g., Nygård and Mezei 2020) or by predicting customers' crosspurchase likelihoods to target customers with the right offerings (e.g., Kamakura et al. 2003). PSA applications can be integrated into CRM systems (e.g., Salesforce Einstein), provided in software independent of CRM systems (e.g., Allego, Chorus, Cogito, CrystalKnows, Showpad, bespoke tools as outlined by Burger and Habel 2020), or as individual reports for a decision maker in sales. Furthermore, note that some PSA applications are marketed as "artificial intelligence" (AI). We refrain from using the term AI in this paper because it is vague and employed in vastly different ways (Huang and Rust 2018; Singh et al. 2019; Ascarza et al. 2021). However, because the common core of AI and PSA is prediction (Agrawal et al. 2018), the model we advance in this paper might likewise apply for sales applications termed "AI."

From PSA adoption to job performance :

To explain the link between a sales employee's adoption of a PSA application and their job performance, we draw on the marketing capabilities model (Morgan 2019b). A firm's marketing capabilities comprise complex, coordinated patterns of skills, knowledge, and technology that become embedded as routines over time (Day 1994; Morgan et al. 2009; Jaworski and Lurie 2019). The concept originated from the resource-based view, which argues that firms acquiring inimitable resources and capabilities can gain a sustainable competitive advantage (Kozlenkova et al. 2014). This seminal theoretical lens suggests that different resources, such as the adoption of new technologies, can help companies develop certain capabilities which in turn enhance performance (Day 1994). In the marketing domain, Day (1994) particularly emphasizes the essential importance of the market sensing capability (see also Sett 2018). That is, firms may sense or identify business opportunities in the market through information gathering and analysis (Day 1994). Sensing capabilities are rooted in a firm's internal analytical activities, and typical tasks include for instance the identification of new potential customers, lead generation, and customer prioritization (Morgan et al. 2009). Applying this logic to our conceptual model, we argue that adopting PSA applications augments sales employees' market sensing capability and equips them with important information to manage sales organizations and sell to customers more effectively and efficiently (see also Day 2011; Feng et al. 2017; Guenzi and Habel 2020). Accordingly, we expect PSA adoption to improve job performance through leveraging sales effectiveness and efficiency. For example, PSA applications may increase effectiveness by allowing sales employees to generate higher quality leads, identify up-/cross-selling potentials, and prioritize customers by their predicted future value (Guo et al. 2020). PSA applications may increase efficiency by allowing sales employees to optimize effort allocation toward the most valuable customers and minimize time resources by automating analyses.

Not all sales employees are expected to equally benefit from adopting PSA applications. That is, a "one-size-fits-all" perspective is less suited to fathom the intricate, multi-facetted effects of PSA applications on sales employees' job performance, indicating the necessity for a contingency perspective. Again, the marketing capabilities framework is instructive in identifying relevant contingency factors as it suggests that effects of marketing capabilities on employee performance should inherently depend on employees' work environment, particularly, the complexity of the environment (Day 2011). Following this logic, adoption of PSA should only increase employees' job performance if the adoption enables employees to effectively cope with complexities in their environment. We propose two sets of contingencies in this respect. First, we argue that effects of PSA adoption should be contingent on factors directly associated with how the PSA application is designed and deployed. As the underlying conceptual mechanism, we propose that these variables will determine the value potential inherent in the particular application for employees (Wedel and Kannan 2016). This value potential in turn will moderate effects of adopting the PSA application on job performance. We define value potential as the degree to which a PSA application helps sales employees achieve desired outcomes. Second, we propose that variables related to the decision-making environment such as competitive intensity or market turbulence, will moderate the effect of PSA adoption on job performance. This is because these variables will determine sales employees' demand for accurate information. The higher sales employees' information demand, the higher should be the value potential of adopting PSA in a certain environment (Wedel and Kannan 2016). That is, value potential of PSA in the decision-making environment again should constitute the key mechanism governing performance effects of adopting the PSA application.

Business problem :

PSA applications aim to help solve particular sales-related business problems (Schoenherr and Speier-Pero 2015; Habel 2019). In the sales context, the onus on solving these business problems is mainly on sales managers and salespeople (Johnston and Marshall 2016). Essentially, sales managers' business problems and decisions pertain to the management of their sales business unit (hereafter, strategic sales management), which comprises the sales strategy development and business planning with sub-activities such as prioritizing customer acquisition and retention, optimizing the salesforce size (Albers and Krafft 2013), and the leadership of salespeople (hereafter, operational sales management). Salespeople in turn are often the key, representative of the firm, especially in the B2B context. Their business problems and decisions pertain to the stages of the sales funnel (Habel 2019), from customer acquisition to customer development (Albers and Krafft 2013; Söhnchen and Albers 2010). Literature has developed PSA applications for a wide variety of sales managers' and salespeople's business problems (see Figure 2). For example, PSA applications support sales managers in decisions regarding allocating the efforts of salespeople, sales forecasting, sales force size and structure, incentive plans for salespeople, and salespeople retention (Figure 2, Panel A). They support salespeople regarding lead qualification, price quotation, time allocation on existing customers and prospects, customer feedback handling, cross-selling, customer churn prevention, customer lifetime valuation, and customer win-back.

Different use cases help sales employees in different work tasks and thus likely exhibit different value potentials, rendering the effect of a sales employee's PSA adoption on job performance more or less pronounced. To understand which use cases exhibit lower or higher value potential, consider that a predictive model aims to help sales employees improve their decisions (e.g., which product to offer to a customer). Accordingly, we propose that the value potential of a PSA use case should be determined by the value potential of the decision a use case aims to improve. The value potential of a decision can be quantified through the expected value theorem, and amounts to the multiplication of (1) the possible outcome (e.g., sales revenue generated in case the customer accepts the offer) with (2) the probability that the decision results in that outcome (e.g., probability that the customer accepts the offer) . As to the first, the possible outcome of decisions supported by PSA strongly depends on the specific use case. For example, for decisions that aim to retain salespeople, a firm's maximum possible outcome is given by their currently incurring costs due to salesperson turnover. For decisions that aim to improve cross-selling, a firm's maximum possible outcome is given by the sales potential in their customer base, given the firm's product portfolio. These examples illustrate that the outcomes of a PSA application are likely to be highly context dependent.

Second, the probability that a decision made using PSA achieves the desired outcome also depends on the specific use case. For different use cases, this probability should strongly differ. We expect a use case's probability to realize a specific value potential (1) to decrease with the the scope of action, that is, the number of different, potential actions or measures which a sales employee can take based on a prediction, and (2) to increase with the influenceability of outcomes, that is, the extent to which outcomes depend on sales employees' decisions rather than other factors. For example, when aiming to retain a salesperson with a high probability of departing, sales managers might face a high scope of action and moderate influenceability: A sales manager might offer a salesperson with a high turnover probability a different compensation plan, a reduction of workload, job enrichment, reassignment to a different team, or a new job title—to name just a few of many available options. Even when the sales manager identifies the right decision to take, retention will depend on other factors, such as a salesperson's disengagement from the current employer as well as alternative job offers. Conversely, when a predictive model suggests that a customer might cross-purchase a certain product, the scope of action requires offering this specific product. Similarly, influenceability might be moderate to high, increasing the probability of achieving a cross-sale.

Model development :

PSA applications are developed from predictive models based on appropriate data (Shmueli and Koppius 2011). We propose three model development choices determining the value potential of a PSA application: (1) The use of unstructured vs. structured data, (2) the use of non-parametric models, and (3) the model's predictive validity. We elaborate in the following. Unstructured data Most prior studies developed PSA models from structured data, that is, data in a tabular format in which observations are stored in rows and clearly defined variables in columns. Typical data sources in this respect include (1) transaction data, comprising information on the time, quantity, specific products, and terms at which a customer made a purchase (e.g., Glady et al. 2009), (2) CRM data, comprising master data of customers, information on company–customer interactions, and sales funnel-related variables such as leads, opportunities, and quotations (e.g., Eitle and Buxmann 2019), (3) web analytics data, such as website impressions and click-throughs (e.g., Nygård and Mezei 2020), (4) HR data, such as a salesperson's tenure, age, or compensation (e.g., Misra and Nair 2011), and (5) companyexternal data, such as macroeconomic information (e.g., Hadavandi et al. 2011). For example, papers that developed models to forecast sales primarily used historical transaction data complemented by macroeconomic data, such as consumer price index, unemployment rate, and gross domestic product (e.g., Fantazzini and Toktamysova 2015; Hadavandi et al. 2011).

Conclusion :

Academia and practice are increasing focus on PSA (Alavi and Habel 2021; Luo et al. 2021). However, despite its growing relevance, numerous firms face lingering difficulties when seeking to implement new PSA applications in their sales force. While these issues may arise from different sources such as insufficient data quality or frictions when integrating new PSA tools with existing data infrastructure and systems (Sleep et al. 2019), often sales employees harbour aversion against working with new algorithms. Such implementation hurdles at best, limit the productivity increases potentially enabled by PSA, but at worst, undermine established sales processes, even reducing sales productivity. Against this backdrop, our study integrates existing literature with pertinent theory into the PSAA Model, which predicts how PSA affects sales employees' job performance. To that end, we integrated key findings from our literature review with the marketing capabilities model (Day 1994, 2011; Morgan 2019b), which allowed us to conceptualize and propose important moderators thatreside in the PSA application and decision-making environment. Future research can readily employ this model as a framework for future research on PSA effectiveness. This is particularly important because prior literature conceived PSA as an umbrella term for a set of methods rather than a research construct. Our conceptual model provides such a research construct, which we perceive as essential to advance the PSA literature from building PSA applications to building PSA theory. A key prediction of the PSAA model is that adopting PSA applications constitutes a valuable resource to sales employees which provides them with a competitive advantage, increasing their job performance (Hunt 2015; Morgan 2015). However, we acknowledge as a limitation of our current conceptualization that this perspective presumes a rather traditional, static perspective on market development and individual decision making. That is, to date, our model does not comprise how new PSA applications may leverage diverse data sources to create innovations that may transform markets and hence take a dynamic, iterative, and longitudinal perspective on market development. More specifically, future research should extend the PSAA model by complementing it with a more dynamic perspective on how PSA adoption can transform and in fact form markets.