Industrial Marketing Management 69 (2018) 135–146

Contents lists available at ScienceDirect

Industrial Marketing Management

journal homepage: www.elsevier.com/locate/indmarman

Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice

Niladri Syam[a,*], Arun Sharma[b]

[a] University of Missouri-Columbia, Trulaske College of Business, Room No. 431, Cornell Hall, Columbia, MO 65211-0001, USA
[b] University of Miami, School of Business Administration, P.O. Box 248027, Coral Gables, FL 33124-6520, USA

ARTICLE INFO

Keywords:
Machine learning
Artificial intelligence
Sales
Sales process
Sales renaissance

ABSTRACT

Experts have suggested that the next few decades will herald the fourth industrial revolution. The fourth industrial revolution will be powered by digitization, information and communications technology, machine learning, robotics and artificial intelligence; and will shift more decision-making from humans to machines. The ensuing societal changes will have a profound impact on both personal selling and sales management research and practices. In this article, we focus on machine learning and artificial intelligence (AI) and their impact on personal selling and sales management. We examine that impact on a small area of sales practice and research based on the seven steps of the selling process. Implications for theory and practice are derived.

1. Introduction

We are undergoing a time of profound transformations powered by digitization, information and communications technology, machine learning, robotics and artificial intelligence (Gupta, Keen, Shah, & Verdier, 2017). Many commentators in the business and economic sphere suggest that this will usher in a new epoch—the Fourth Industrial Revolution (Marr, 2016). The fundamental shift in the fourth industrial revolution will be in the area of decision-making. Whereas traditional informational technology helped with processing of communications and data, the decision-making was human. The new shift will be evident in emerging technologies allowing computers to also make reliably appropriate decisions. This digitization shift has begun and will have profound implications for personal selling and sales management functions. The sales profession has always changed in response to changes in the larger macro-environment (e.g., technological, macro-economic, demographic, cultural) within which it operates. As an example, with the advent of advanced telephones and rapid transportation, the sales profession moved away from the stereotypical Willy Loman (from *The Death of a Salesman*) type of traveling salesman with defined routes to visits based on demand. Similarly, with the advent of the Internet and databases, information became more widely available and some of the ordering moved away from written orders to ordering on the internet.

We hypothesize that selling in future decades will be disruptive and discontinuous, owing primarily to shifts in technology. In other words, digitalization of sales functions with the addition of artificial intelligence and machine learning represent a discontinuous change compared to the non-digital era. For example, an emerging firm, Node, uses machine learning and artificial intelligence to harness large databases and match them with data available on the web to create prospect lists (Node, 2017). Their website promises to provide "strategic insight, tactical guidance and cutting edge technology to help anyone find the right person at the right business at the right time with the right message…" (Node, 2017). Analogous to changes as Europe transitioned from the Middle Ages to the Renaissance, we label this shift as the "Sales Renaissance" where the focus of sales management will transition from traditional sales functions to new functions that may involve bridging inter-organizational and intra-organizational boundaries.

In this paper, we explicate the anticipated technology and environmental changes, and discuss existing machine learning and artificial intelligence technologies and processes. We then discuss the implications of technologies on selling functions. Our focus is on the sales process as a critical element of research and practice, and we use the seven steps of selling to describe the basic selling process (Dubinsky, 1981). In order to provide a deeper focus, the paper does not address big data in-depth for two reasons. First, discussion of big data in sales would be a paper by itself. Second, the definition of big data will change with advances in computing such that data sets that are 'big' today will become normal in the future. We have therefore focused on the more fundamental shifts, which are digitalization driven by machine learning and AI, rather than the size and 'quality' of available data.

The next section presents emerging trends in technology and

* Corresponding author.
E-mail addresses: syamn@missouri.edu (N. Syam), asharma@bus.miami.edu (A. Sharma).

https://doi.org/10.1016/j.indmarman.2017.12.019
Received 9 November 2017; Received in revised form 29 December 2017; Accepted 29 December 2017
Available online 17 January 2018
0019-8501/

environment. In the section after that we will describe machine learning and artificial intelligence and emphasize certain tools. In subsequent sections, we will discuss the impact of these technologies and methodologies on the selling process and highlight areas where further attention is needed. The final section will summarize and discuss areas for further research.

2. The fourth industrial revolution

In this section, we discuss the emergence of the fourth industrial revolution with a historical perspective on the previous industrial revolutions. The first Industrial Revolution occurred when mechanization, water power and steam power multiplied the efficiency of productive technologies that had previously depended on human and animal labor. The second Industrial Revolution was ushered in by mass production and the assembly line style of production. This stage was facilitated by widespread availability of electricity that made factory production even more efficient. The third Industrial Revolution saw the advent of computers and automation. Automation that has increased with each industrial revolution, is seen as having both dramatic advantages and disadvantages. In the context of the third revolution, the processing power of computers and increased access to memory, which continues to increase, is ideally suited to automate repetitive and non-value added tasks (e.g., delivery information).

The fourth Industrial Revolution involves, among other things, cyber-physical systems of how humans interact with machines and also the Internet of Things (IoT) where machines interact with machines (Marr, 2016). In an article featured in April via Forbes online Tech/Big Data column Marr, an established expert on IoT, mentions that,

> "While some would argue IoT got off to a rocky start with a lower adoption rate than was predicted, most would agree the IoT is growing and will continue to grow in 2017 and beyond. Whether it reaches the lofty predictions of 50 billion connected devices by 2020 remains to be seen, but I strongly believe that businesses who learn to harness the data created by the Internet of Things are the ones who will survive and thrive in the future."
>
> (Marr, 2017)

The major difference from mere technological advances is the very close interaction between physical, digital and biological worlds. In this paper we concentrate on two aspects of the fourth industrial revolution—artificial intelligence and machine learning and provide a critical examination of how these will affect the selling function across all levels of firm sales and go-to-market strategy. By one estimate, as many as 47% of U.S. jobs are at risk from machine learning and AI fueled automation, and this will happen across a wide spectrum of industries, from salespeople, accountants, real estate agents, and insurance agents to drivers (Marr, 2016). Importantly, the jobs that will be eliminated are ones which currently minimally incorporate machine learning and AI. On the other hand, machine learning and AI also have the potential to vastly increase new jobs which will be driven by these emerging technologies. A recent report by the most prominent IT industry research firm Gartner estimates that enterprise technology powered by AI will create more jobs than it eliminates, and may well account for two million net new jobs by 2025. The new jobs will be both at "new positions of highly skilled, management and even the entry-level and low-skilled variety" (Loten, 2017). Moreover, for the sales function, there are clear limits to how much AI can eliminate sales jobs because of the critical role played by inter-personal buyer-seller interfaces in sales (Knight, 2017). Many, if not most, of the claims of job loss because of AI are due to exaggerated and wild extrapolations and generalizations of AI's capabilities (Brooks, 2017a). Brooks (2017b) mentions the 'performance versus competence' dichotomy that often lies at the heart of such exaggerations. We use cues about how a person performs some particular task to estimate how competent s/he may be at some different but related task. Say, that an experienced salesperson has reported to her manager that a major account that she has been working on has a 60% chance of closing. We naturally assume that she can reasonably answer questions like: *Where is the buyer in the buying process? How many decision makers are involved in the buying center? What are their personalities? What is the time frame for the buyer's decision?* The big mistake would be to assume that some AI powered software that has provided a 60% probability of close would also be able to answer these related, but very different, questions as competently as the experienced salesperson. In sum, the most reasonable expectation vis-à-vis the impact of machine learning and AI in the sales context is that salespeople will have to co-exist with AI and other technologies.

2.1. Machine learning and AI techniques used in sales

The term, machine learning, was coined by a pioneer in computer gaming and artificial intelligence, Arthur Lee Samuel (Lee, 1995), who used the term for the science of getting computers to act without being explicitly programmed. AI researchers use it as the best and most promising approach to the development of human-level AI. Merriam Webster.com (2017) defines artificial intelligence as, "a branch of computer science dealing with the simulation of intelligent behavior in computers." AI refers to the ability of machines to mimic intelligent human behavior, and specifically refers to "cognitive" functions that we associate with the human mind, including problem solving and learning. In the context of this study, machine learning is a prerequisite to the development artificial intelligence (Carbonell, Michalski, & Mitchell, 1983). In addition, machine learning requires substantial amounts of data (big data) and high processing power that is easily accessible. Therefore, our focus is on self-learning models and machine learning. A major advantage in using machine learning models for sales research is that the firm usually has customer exposure and purchase data. In contrast, in the advertising world, the cable TV providers and streaming services own the exposure data whereas the retailers own the purchase data. Thus, deploying these new approaches in the advertising context requires the researchers to first assemble data sets from various sources akin to assembling a 'jigsaw puzzle' (Malthouse & Li, 2017).

Traditionally, a model 'learns' from data when the parameters of a model are estimated based on data. Consider the simplest example of a linear regression problem that has only one explanatory (independent) variable. Given some data on the dependent (Y) variable and the single independent (X) variable, the Ordinary Least Squares estimates for the intercept and slope can be computed manually because there exists closed-form expressions for them. Thus, the linear regression model can learn from the data without the help of machines. However, in most realistic situations, parameters associated with multiple explanatory variables need to be related to their dependent variables in a highly nonlinear manner. Machine learning can be harnessed in these situations. The estimation of the parameters of such complex models require enormous computing power that only modern computers can provide. The complexity of estimating parameters is exacerbated when the data set is large and unstructured. Intuitively, large computers that continuously estimate parameters based on ever-increasing data that becomes available in real-time are utilizing machine learning. Machine learning and AI have clear advantages over traditional statistical methods when: (a) there are a multitude of variables available for analysis, (b) the associations between the variables is uncertain (and likely to be highly complex), (c) the values of each variable are evolving constantly (such as in the case of a GPS), and (d) when understanding correlations between variables is more important that causation. The great strength of machine learning models is in making predictions, especially where an atheoretical prediction will work well. This is the reason that machine learning models are evaluated on criteria such as scalability, real-time implementability, and cross-validated predictive accuracy rather than on internal and external validity and theoretical foundations which are more suited to the traditional models.

Machine learning methods can be broadly categorized into supervised learning and unsupervised learning. Gareth, Witten, Hastie, and Tibshirani (2013) define supervised learning by stating: "Supervised statistical learning involves building a statistical model for predicting or estimating an output based on one or more inputs … Furthermore, in unsupervised learning, "there are inputs but no supervising output." (Gareth et al., 2013, p. 1). In supervised learning the data set has the 'right answers,' in the sense that, for a set of explanatory variables in the data set, the right dependent variable value is provided to the computer. Said differently, in supervised learning we have an output variable corresponding to set of input variables. In unsupervised machine learning, programs are provided unstructured and sometimes unlabeled data and asked to determine the structure or patterns in the data. Among the more traditional techniques, the usual OLS regression which has an output variables (Y) corresponding to a set of input variables (Xs), is supervised whereas cluster analysis which has no well-specified output variable is unsupervised. A salient advantage of many machine learning tools (both supervised and unsupervised) is that, unlike traditional statistical techniques, they do not make a-priori assumptions about data that are often restrictive. Thus, at least in unsupervised machine learning there is *lower risk* that improper statistical techniques will be used when the underlying assumptions are violated by the data. Compared to traditional statistical techniques, machine learning is better at predictions because it can accommodate highly nonlinear and complex relations between inputs and outputs. However, the disadvantage is that machine learning lacks the ease of interpretability found in more traditional models (Hastie, Tibshirani, & Friedman, 2017, p. 351). The primary reason for lack of interpretability is that machine learning uncovers unmodeled complex relationships.

An additional dimension used to classify machine learning methods is based on generative versus discriminative models. Generative models try to model an underlying probability distribution and also the prior probabilities. They are called generative because the probability distributions can be used to generate synthetic data points. Hidden Markov Models (HMM) and Naïve Bayes classifiers are among common generative methods. In discriminative methods, no attempt is made to model underlying probability distributions, and the posterior probabilities are directly estimated. Logistic regressions, support vector machines (SVMs) and conditional random fields (CRFs) are common examples of discriminative methods. Both logistic regression and Naïve Bayes methods do classifications, but in generative models a probability distribution and the prior probabilities are updated, based on the Bayes theorem, to obtain the posterior. Thus, if the class variable is y and the feature variables are $X = (x_1, \ldots, x_n)$, then Naïve Bayes models $p(y, X)$ as well as the joint pdf (probability density function), and the classification is done by conditioning the pdf.

In logistic regression, the posterior probability, $p(y \mid X)$, is directly modeled (without modeling the joint pdf). Hidden Markov models (HMM) and CRF are generalizations of the Naïve Bayes classifier and logistic regression to *sequence models*. Sequence models have a sequence of observations $X = \{x_t\}_{t=1}^{N}$ and a sequence of states $Y = \{y_t\}_{t=1}^{N}$ (Srihari, 2010). They are also called graphical models since undirected or directed graphs are appropriate ways to models many random variables that are interdependent. In graphical models, the vertices are random variables and "the edges in a graph are parameterized by values or potentials that encode the strength of the conditional dependence between the random variables at the corresponding vertices" (Hastie et al., 2017, p. 626).

In the following section we provide a brief overview of the more common machine learning tools that have been, and could be, used in sales. The overview is not meant to be an exhaustive survey of all machine learning methods. Table 1 presents a summary of the advantages and disadvantages of these machine learning tools, their current applications, and their applicability in sales.

2.2. Machine learning models and tools

2.2.1. Neural network

A very common tool that has been used in many marketing and sales applications is a neural network (also called an artificial neural network or ANN). We provide a brief intuitive idea behind this methodology. Essentially a neural network is a 'black box' that searches many models (including nonlinear models with complex interactions) without being "told" the nature of the relationship. Neural networks solve problems differently compared to traditional computer algorithms. The traditional algorithmic approach is cognitive and depends on a stepwise and logical set of instructions. The problems solved by this method are those that we understand quite well and know how to solve. Neural networks are most effective in tackling situations with messy, noisy and complicated data, and can extract patterns and trends that humans and conventional computer programs cannot detect. Neural networks solve problems (process information) by examples, much like humans do, and are not programmed to perform a specific task. These networks simulate the human brain and are comprised of many highly interconnected nodes (like neurons in the brain). The key to the operation of a neural network are the firing rules that tell the neurons what output to generate based on the (many) inputs that act on them. The neuron has two modes—a training mode and a using mode. In the training mode, the neuron is trained to fire/not fire for different input patterns that are seen in the training data. In the using mode, when it analyzes new data, it sees whether an input training pattern is detected at the input (of the new, non-training, data). If an input training pattern is detected, then the associated output becomes the current output. However, if an input training pattern is not detected, the neuron turns to the firing rule to determine the current output.

The most commonly used neural networks in business applications are feedforward neural networks. If there are no feedback loops, and the signals travel only one way from input to output, the neural network is of the feedforward type. Most neural networks have three layers of variables (called units): The layer of input units is connected to the layer(s) of 'hidden' units, which in turn, is (are) connected to the layer of output units. The raw data that is fed into the system constitutes the input layers. There are weights on the connections between the input and hidden units, and these weights along with the activities of the input units determine the activities of the hidden units. Similarly, there are weights on the connections between the hidden units and output units, and these weights along with the activities of the hidden units determine the activities of the output units. The advantage of this type of architecture is that the hidden units can choose their own representations of the inputs by suitably modifying these weights.

2.2.2. Support vector machine

Another powerful machine learning technique that is used across many different applications is called the Support Vector Machine (SVM). SVMs can be used to perform classifications, and unlike traditional linear classifiers can efficiently perform nonlinear classifications using kernels. Essentially, the kernel projects the data into a higher dimension space, and a linear classification performed in the higher dimensional space is equivalent to a nonlinear classification in the lower dimension space. Moreover, unlike the traditional classification models which construct a separating hyperplane (in two-dimensional space, the separating hyperplane is a line that separates two classes), the SVM maximizes the 'margin' between the classes. The big advantage of SVM is that only a subset of the input data points, called the 'support vectors,' is required to determine the maximum margin (which does the classification). In the figure below, the two positive and two negative points on the dotted lines are the support vectors. The tradeoff in SVM is that, because it uses a margin (and not a line) for classification, widening the margin could cause more points to appear on the wrong side of the margin and become misclassified as shown in Fig. 1.

To capture this tradeoff, the SVM method maximizes the margin

Table 1
Comparing SVM, Artificial Neural Networks (ANN) and NLP.

	Advantages	Disadvantages	Current applications	Potential applications in sales
Support Vector Machines (SVM)	1. Can perform non-linear classifications. 2. Strong convexity property ensures global minima. 3. Strong theoretical foundations. 4. Allows simple geometric interpretation because of theoretical foundation. 5. Gives sparse solution because we can focus only on the 'support vectors.'	1. Easiest way to create an n-ary classifier is to create n support vector machines and train each of them one by one. 2. Not robust to outliers in input space. 3. Computational scalability is limited.	1. Predictive modeling. 2. Time series forecasting. 3. Handwriting recognition. 4. Face detection. 5. Text and Hypertext Categorization. 6. Classification of images. 7. Bioinformatics.	1. Targeting customers and sorting them by propensity to purchase. 2. Sales forecasting. 3. Advance warning of probability of switching to competitors. 4. Suggesting cross-selling and up-selling opportunities.
Artificial Neural Network (ANN)	1. May have any number of outputs, 2. Better at identifying very complex patterns, 3. Very good at predictions, 4. Very good at real-time analyses with evolving data,	1. Often converges to local minima, 2. Weak theoretical foundations. Heuristic origins and development, 3. Often overfits data if training goes on too long, 4. Interpretability of parameters not strong because of heuristic problem solving approach,	1. Handwriting recognition, 2. Face detection, 3. Predictions, 4. Classification of images, 5. Artificial intelligence in games,	1. Segmentation using unsupervised neural networks, 2. Segmentation using supervised neural networks to segment customers based on segment labels: frequent buyers; purchase amount etc. 3. Tracking customer behavior over time. 4. Sales forecasting 5. Keyword analysis for prospecting. 6. Text, audio and video analysis of salesperson-customer communications for lead generation and subsequent qualifying of prospects. 7. Designing optimal presentation strategies. 8. Designing pre-approach and approach strategies. 9. Optimizing sales operations management like operations planning, sales process activity scheduling, timetabling etc.
Natural Language Processing(NLP)	1. Can perform keyword search in text and emails. 2. Linguistic classification. 3. Highly flexible and highly representative of reality. 4. Easy to represent new and complex concepts. 5. Highly expressive.	1. Very difficult to make generic searches. 2. Problems with synonyms, unknown words, ill-formed sentences, syntactic ambiguities caused by structure, semantic ambiguity and contextual ambiguity. 3. Ambiguous, soft and fuzzy rules. 4. Needs very high computing power to encode all the rules that can understand human speech.	1. Machine translation. Making all the world's information across regions and languages. 2. Fighting spam. False positives and false negatives issues in span filters. 3. Information extraction. 4. Summarization. 5. Question answering.	1. Search engine insights. Finding keywords and phrases that potential customers search can help a company identify prospects. 2. Glean customer insights from their social media postings and reviews. 3. Text analysis of emails, audio and video records to generate and qualify leads. 4. Lead scoring for better efforts allocation based on quality of leads. 5. Trend analysis.

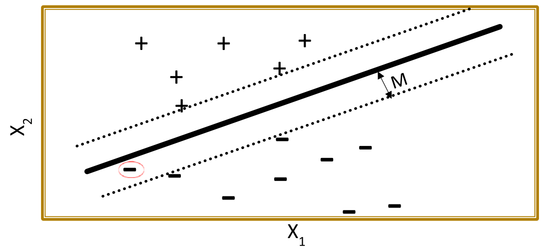

Fig. 1. Support Vector Machine. The negative point, circled in figure, is on the wrong side of the margin and is, thus, misclassified.

subject to the constraint of controlling 'losses' from misclassification. SVM uses the hinge loss function that only penalizes points on the wrong side of the margin (Sapankevych & Sankar, 2009). SVMs can also be used for predictions, often for time series predictions, using regression-like structures. Thus, they are a powerful method for forecasting. The variant of SVM used for regression is often referred to as support vector regression (SVR). Just as in the case of classification by SVM, the basic idea of regression using SVR is also to project the nonlinear input data to a higher dimension space via the use of kernels. Since the transformed data is linear in the higher dimensional space, the method performs a linear regression in that space. As opposed to the traditional OLS regression, which uses all the input data points to compute the error function that is subsequently minimized to estimate the regression function, SVR uses only a subset of the input data points for estimation, and the sparseness of the solution is one of the major advantages of this approach.

2.2.3. Natural language processing

Many practically relevant advances in business, are occurring due to rapid strides in natural language processing (NLP). NLP lies at the intersection of linguistics and machine learning, and falls within the area of computational linguistics. NLP started in the 1950s and some early attempts were made to come up with a context free grammar (CFG). Technically speaking, CFGs are appropriate for programming language syntax (the 'artificial' languages) and are inadequate for natural language, the language of humans. The initial NLP systems used handwritten rules to capture the rich nuances of how human beings speak, including parts of speech such as verbs, nouns, adjectives that capture the relationships between text units. In this way the early NLP systems tried to extract meaning (semantics) from text. However, due to the ambiguous meanings of words, which derive meaning from the context in which they are used, the handwritten rules had to accommodate subcategories and constraints. This made these rules extremely cumbersome, and this ultimately made way for *statistical NLP*, which could use fewer and broader rules but had to resort to statistical information to *disambiguate*. NLP has two steps—NLU or natural language understanding and NLG or natural language generation. NLU is the process by which the computer takes natural language and converts it to artificial language for use by the computer. NLU is much harder to accomplish and is also referred to as speech recognition, or speech-to-text, and the most commonly used methodology for this is the hidden Markov models (HMM). Many NLP systems also use conditional random fields (CRFs). At this point, it must be mentioned that NLP also has many problems in practical implementation, owing primarily to the richness of human communication which is very hard for computers to understand (Jacobs & Rau, 1988).

It is useful to distinguish between HMM and CRF models. In HMMs a variable can transition between states with certain probabilities, and these transitions generate output states which also occur probabilistically. The outputs are observed, but the state transition probabilities and output probabilities are 'hidden.' CRFs also predict state variables (*Ys*) based on observed variables (*Xs*), but these variables are the discriminative equivalent of HMMs. Fig. 2, taken from Srihari (2010), shows the relationship between Naïve Bayes, Logistic Regression, HMM and CRF.

Sequence models of machine learning, such as HMM and CRF, are the most commonly used methods for statistical NLP. Most NLP systems start with 'part-of-speech tagging' (POS), which assigns syntactic labels to each word—whether a word is a noun, verb, adjective, adverb past tense or present tense etc. The modern statistical approach to NLP does linguistic classification to find the most likely meaning for each word. Then there is a need for shallow tagging, which identifies *phrases* from constituent POS tagged items. The input in this stage are words in a sentence where each word has a POS tag. The task here is to label each word as (1) outside a chuck, (2) starting a chunk or (3) continuing a chunk (Srihari, 2010). Since we deal with sequences of words, therefore HMM and CRF type models are used heavily to accomplish these tasks.

While the above was an overview of machine learning techniques, below we provide the crucial link between the advancements in ML and AI techniques and their applicability to the sales function.

3. Impact of machine learning and artificial intelligence on the sales function

The sales function has already started observing some of the impacts of the Third Industrial Revolution through extensive use of computing technologies and automation. Personal selling and sales management has a number of routine tasks (e.g., order entry, new product announcement). Since these tasks take time and energy away from the primary task of salespeople in developing relationships, the advent of automation in routine tasks enhances the productivity of salespeople. Some examples of tasks that have been automated are inside sales, documentation, detailed call reports and provisioning of product/service data. Automation is continuing to ease the salesperson's burden for repetitive and non-productive processes, and the time that is freed up can be used by sales people for more productive customer facing tasks.

3.1. Categories of salesforce and machine learning and artificial intelligence

In examining the impact of machine learning and AI, we need to acknowledge that there are many categories of salesforce with distinct roles that are emerging. While these categories can most accurately be depicted on a continuum from simple to complex, for the sake of simplicity, we focus on three distinct and discrete categories that span the continuum. The first category is a salesforce that sells simple products. The simple sale features standard items and is increasingly being sold by an inside salesforce or online (since the customers can order them directly without much interaction with salespeople, or with minimal interaction with the sales staff that handles inbound calls), which is relatively less expensive and quick. Simple sales are used by firms whose products/services/solution are easy to understand and typically have low margins. Technology has had a dramatic impact on the simple sale whether inside or online. Once the firm has data on a customer, most tasks can be automated. The third industrial revolution has already had a deep impact on this type of sale and, with the advent of machine learning and AI-powered chatbots, we expect the fourth industrial revolution to have maximal impact in these areas.

The second salesforce category is for products/services/solution where the profit margin is high, and the products/services/solution data is easily accessible to the customer (both data access and data understanding). The high profit margin allows firms to profitably use a salesforce. A distinct example of this category is a pharmaceutical salesforce where salespeople provide information to physicians that can be easily accessed on the Internet. Other examples are spirits and medical device industries. As customers also gravitate toward accessing information through devices (rather than human), information providers will provide more customized information (through machine learning and artificial intelligence); thus, as profit margins erode, we expect the fourth industrial revolution to have a dramatic effect on this

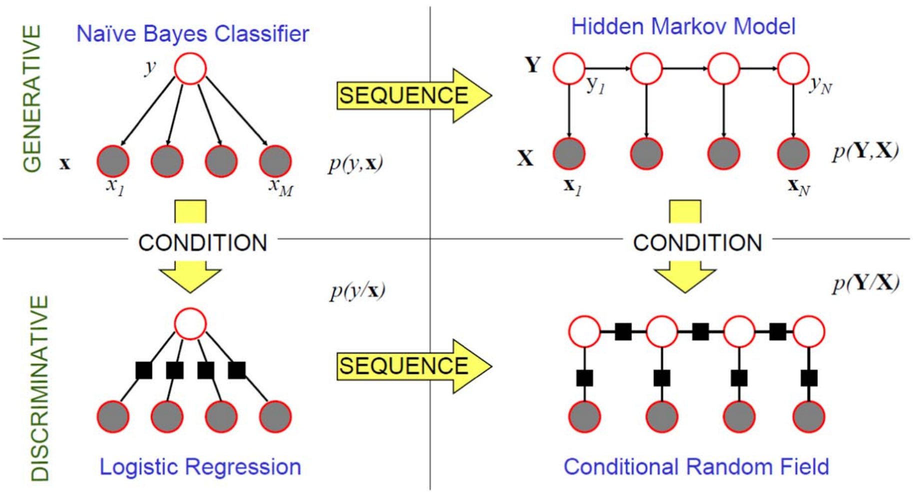

Fig. 2. Relationship between Naïve Bayes, Logistic Regression, HMM, and CRF.
Source: Srihari (2010).

type of salesforce. However, unlike the simple sale machine learning and AI may not replace the salesperson in this type of sales – technology will be an enabler, sometimes significantly so.

The third category is complex sales where multiple buying center members have different needs and information is not easily accessible as it needs to be customized for individual buying center participants. In this rapidly changing and complex environment, sales people have to become 'knowledge brokers' and are required to uncover customer needs that perhaps have not been explicitly stated even by the customers themselves (Verbeke & Dietz, 2011). This type of deep customer knowledge makes the role of customer development even more prominent than in the past. We expect a slower and more measured impact of machine learning and artificial intelligence in this case.

4. Classification scheme to study the impact of machine learning and artificial intelligence on the sales function

We needed a classification scheme to create a parsimonious framework that would allow us to focus on some key issues. An in-depth discussion on all possible areas of personal selling and sales management would require much more space than a single article. We examined two classification schemes—the selling process and the classification of sales research. Different industries and different companies within an industry have different sales processes depending on the complexity of the product/service, the sales cycle time, the sophistication of the buyer, and the length of the channel. Despite these differences, Dubinsky (1981) suggested the seven steps of selling apply to most sales situations—prospecting, pre-approach, approach, presentation, overcoming objections, close, and follow-up. This model of sales process is used extensively in research (Homburg, Müller, & Klarmann, 2011). The second classification scheme was developed by Plouffe, Williams, and Wachner (2008) who categorized the research conducted in personal selling and sales management. They have 20 categories which exceeds our research needs; therefore we choses to focus on the seven-step selling scheme developed by Homburg et al. (2011) as shown in Table 2.

In the next sections, we discuss the impact of machine learning and AI on the seven steps of selling and derive possible areas for future research (prospecting, pre-approach, approach, presentation, overcoming objections, close, and follow-up). Due to the similarity of two sets of stages in terms of sales and machine learning research, we combined the discussions on pre-approach and approach as well as objection handling and close. The research questions and some research avenues are summarized in Table 2.

4.1. Prospecting

An important stage in the sales process, and the one that overlaps most often with the customer development role of sales, is prospecting. In this stage, the firm needs to perform the task of finding customers and qualifying them (scoring, the potential customers based on some measure of their propensity to buy). In this regard, we discuss three inter-related topics which are segmentation, (1) targeting and positioning (STP), (2) demand estimation, (3) lead generation, and lead qualification. The reason to include STP in this review is that these typical marketing functions are intertwined with the initial stages of the sales process where marketing-sales integration is the most intense. Also, we would like to clarify that seen through the lens of the STP framework of marketing, segmentation would be associated with initial prospecting, and targeting would be associated with lead qualification. We discuss these issues next.

4.1.1. STP framework

In discussing the impact of machine learning on customer development, we use the STP (segmentation, targeting, and positioning) framework, specifically as it relates to the sales function. The usual starting point of most efforts at customer development start with segmenting customers into groups that share common characteristics, and segmentation methods have a very long history in marketing. This was well documented by Wedel and Kamakura (2000), who provided a very good summary of all the segmentation methods available up to the year 2000. The 21st century machine learning tools continue to place greater statistical power in the hands of marketers. Segmentation usually relies on 'traditional' clustering techniques like cluster analysis, and classification techniques like chi-squared automatic interaction detection (CHAID). Other current segmentation methods include support vector machines, classification and regression trees (CART), artificial neural networks (ANN), genetic algorithms (GA) and hidden Markov models (Ben-Hur, Horn, Siegelmann, & Vapnik, 2001; Gath & Geva, 1988; Hruschka & Natter, 1999; Huang, Tzeng, & Ong, 2007). Machine learning is increasing the efficiency of these segmenting algorithms and making it possible for massive and even unstructured data sets, i.e., big data, to be used for segmentation. Hruschka and Natter (1999) use the unsupervised learning variant of neural networks, and in their feed-forward network both the input and output variables (units) are segmentation criteria. Between the input and output, they have one hidden layer whose units are segment members. Using the segmentation criteria of a given person as inputs, the method uses the multinomial logit to determine segment membership for the hidden layer. For the output layer, segment memberships obtained in the intermediate hidden layer are weighted by segmentation criterion specific weights. The output

Table 2
Summary of Issues Worthy of Future Research.

Topics (seven steps of selling)	Important research questions
Prospecting	• How can firms study customer sentiment for products using tools such as machine learning and AI to estimate demand? How does demand estimation using customer sentiment compare with traditional methods? • How can machine learning and AI enhance forecasting and demand estimation when the external environment undergoes dynamic, rapid and unforeseen changes? • How can machine learning and AI use text analysis of emails and other communications to identify prospects? • How can machine learning and AI perform video and audio analyses of chat logs, archive of transcripts from online chat and instant messaging conversations with customer service personnel, to identify prospects? • How can machine learning and AI-based text mining and keyword analysis (e.g., identifying and validating qualifying words) assist sales organizations in lead qualifications? • How can machine learning and AI enhance smart and continuous customer targeting in real time? • How can machine learning and AI help develop recommendation systems to suggest cross-selling and up-selling opportunities?
Pre-approach and Approach	• How can text mining coupled with machine learning and AI enhance pre-approach and approach over traditional methods? • How can machine learning and AI help develop systems that recommend customer engagement strategies based on analysis of past engagements and their success/failure and depending on salesperson, industry and customer characteristics?
Presentation	• How can machine learning and AI help design dynamic sales presentations by incorporating customer feedback (verbal and non-verbal) and salesperson's goals while conducting a presentation? • How can machine learning and AI develop systems that automate the sales-finance link to provide data-based 'financial valuation' arguments for the company's products and design Total Cost of Ownership (TCO) presentations to counter price objections? • How can machine learning and AI develop systems that suggest multi-media presentation techniques based on 'attribution modeling' that can attribute the relative contributions of various communication alternatives in achieving successful sales? • How can machine learning and AI assist in leveraging 3D printing technology to construct customized prototypes for demonstrations to customers? • How can machine learning and AI develop systems to use Collaborative Filtering to improve recommendation systems which can be used to optimize presentations?
Overcoming Objections and Close	• How can machine learning and AI be used to conduct 'lost order audit' and help sales organizations better anticipate objections and handle them? • How can machine learning and AI develop systems to recommend negotiation strategies based on analyzing history of past negotiations and their success/failure rates and how they depend on customer and salesperson characteristics? • How can machine learning and AI develop systems to audit post-invoice sales recording process in accounting systems to identify sales anomalies such as 'channel-stuffing' because of salesperson's timing games and undue discounting of goods? • How can machine learning and AI develop systems to audit product returns by customers to identify problem customers and/or ineffective/problematic salespeople? • How can machine learning and AI provide advanced warning of lost sales based on analysis of history of objections and how they were handled?
Follow-up	• How can machine learning and AI better streamline sales-service links? • How can machine learning and AI create better advanced warning systems from the sales frontline to the supply chain and other related functions? • How can machine learning and AI better predict rebuy? • How can machine learning and AI better streamline the sales-manufacturing link? • How can machine learning and AI develop systems for seamless connections between sales invoicing and inventory systems to reduce delivery lag times? • How can machine learning and AI integrate 'upstream' supply chain (suppliers to firm) and 'downstream' supply chain (firm to customers) to enhance efficiency and effectiveness of a firm's entire supply chain operations?

units are obtained using a binomial logit function which uses the weighted sum of the memberships over all segments. The authors show that this segmentation neural network performs much better than the traditional K-means clustering approach, in that all square-error values for the neural network are smaller than those for the latter method for segments numbers between 2 and 11 (Hruschka & Natter, 1999).

Consistent with the customer-centric view of sales, an important type of segmentation uses some measure of customer lifetime value (CLV) or profitability as a segmentation basis. This type of segmentation groups the customer base into clusters wherein a given cluster has customers with similar lifetime values. Florez-Lopez and Ramon-Jeronimo (2009) implement such a segmentation study using machine learning techniques such as feature selection (Blum & Langley, 1997; John & Kohavi, 1995) and decision trees (Heath, Kasif, & Salzberg, 1993; Murthy, 1998). As is standard, feature selection allows the vast number of variables to be pared down to only those that are most useful for segmentation, thereby making the task of interpretation easy. Florez-Lopez and Ramon-Jeronimo (2009) then demonstrate a creative use of machine learning to accomplish segmentation by using decision trees.

Following segmentation, sales managers select a subset of segments to target. The selected segments are targeted with the appropriate products, which are offered at the appropriate prices, supported with appropriate communications and promotional strategy, after which the products are made available to customers through the appropriate distribution strategy. These personal selling activities fall under the rubric of the firm's positioning strategy which seeks to distinguish its offer from competitive offers.

4.1.2. Demand estimation

Once targeting is completed, there is a need for demand estimation and sales forecasting, within the target market(s). This demand can be facilitated by machine learning and AI. The key to success for firms, especially for technology-driven ones which are likely to narrowly focus on technical aspects and ignore broader market needs, is customer validation (Blank, 2013). Customer validation is the step that sales managers need to estimate customer demand and profitability. Therefore, demand estimation, or sales forecasting, is part of the overall customer development function of the firm. Demand estimation is also critical because most planning activities that a sales organization undertakes are critically dependent on generating accurate forecasts. Sales forecasts are also invaluable for financial planning, inventory planning and a host of other organizational planning functions. Moreover, designing appropriate sales territories depends on accurate estimates of sales potentials for all sales coverage units (SCUs). It is important to distinguish *sales forecasting* from *sales potential estimation*. The sales forecast is a prediction of how much a company *will* sell, whereas sales potential is an estimate of how much it *could* sell. Market potential of a territory is an aggregate measure how much all brands being sold in that territory could sell, whereas sales potential is a measure of how much a focal brand could sell.

Machine learning has been used extensively for sales (demand) forecasting. Carbonneau, Laframboise, and Vahidov (2008) use a combination of SVM and neural networks to forecast demand and show that these methods are superior to the traditional forecasting methods like trend, moving average and linear regression. More generally, machine learning has been found to be very successful in sales forecasting through time series prediction. Landt (1997) and Lawrence, Tsoi, and Gilles (1996) have shown how neural networks can efficiently handle predictions based on time series data, especially for chaotic time series. A vast majority of the traditional autoregressive methods for analyzing time series data assume linear stationary processes. However, chaotic time series data are very unlikely to obey simple linear relationships between the independent variables (called 'input cells' in the neural networks terminology) and dependent variables ('output cells'). Therefore, neural networks, which do not assume any such relationships, and are very effective at pattern recognition, are particularly suited for analyzing such time series data.

The regression variant of SVM has been used by Carbonneau et al. (2008) for forecasting by exploiting the ability of regressions to make predictions. Similarly, Mukherjee, Osuna, and Girosi (1997) and Sapankevych and Sankar (2009) have demonstrated the effectiveness of support vector machines in handling time series predictions.

Another way in which machine learning based AI is improving sales forecasting is to use natural language processing (NLP) tools which enable computers to identify key words in speech and emails of potential consumers to predict the probability that these consumers will purchase. Natural language processing can be adapted to act as efficient approaches to linguistic classification in which linguistic categories can be predicted based on linguistic contexts (word choices in light of their perceived meaning based on the social setting, needs, and expectations of the target audience). Ratnaparkhi (1997) shows how various linguistic contexts can be combined to predict the probability of a certain linguistic category co-occuring within a certain linguistic context. The literature in sales and marketing has described many methods of determining the probability of purchase. Once the probability of purchase is known, then multiplying it by the potential (i.e., maximum) number of consumers will give the firm a good estimate of the expected number of purchasers. Syam (2017) provides a continuous-time Markov chain model that can predict the time-dependent win-rate of a project/opportunity that enters the top of the sales funnel and has to go through several stages before closing. This then gives a time-dependent estimate of the number of closed accounts which can be mapped to the sales forecast.

4.1.3. Lead generation and lead qualification

In business-to-business situations where the selling firm deals with fewer customers able to make large purchases, it is often feasible to do targeting at the individual customer level rather than at the segment level. Regardless of whether it is done at the segment level (for a buyer with a large customer base) or at an individual customer level, this stage of the sales process has seen enormous advances based on machine learning and AI. Almost an overwhelming array of cutting-edge tools have been developed in marketing, operations research, statistics, computer science, mathematics, economics and econometrics, which allow firms to do lead qualification much better than even a decade ago.

We now examine one aspect of prospecting – generating qualified leads. This may be a single stage (number of qualified leads) or a two-stage process (number of prospective customers + number of qualified leads). Lead qualification can benefit immensely from the efficiencies of automation brought on by machine learning and AI. One recent estimate states that, "On average, sales reps spend 80 percent of their time qualifying leads and only 20 percent in closing," (Porter, 2017). Lead qualifications seeks the answer two questions – can the firms buy our product/services/solutions (do they have the resources) and will they buy our product/services/solutions (do our offerings match the need). While the traditional method of lead qualification was a meeting with a customer, emerging technologies allow the sales organization to build mathematical and statistical models that can be then estimated using historical data. Researchers have demonstrated that machine learning and AI can be used for lead generation that requires less salesperson energy and has better outcomes than lead generation by salespeople.

It is important to note that we conceive lead qualification broadly as not only recording, vetting and certifying all the information about the customer, but also any attempt at determining the objective 'quality' of the lead. The most common metric for lead quality is some measure of propensity to buy, and many machine learning tools that perform predictive analytics can help managers in this regard. Some of the more common tools are SVMs, ANNs, Discriminant Analysis, Naïve Bayes and K-nearest neighbor. These methods can determine what actual online and offline behaviors have the highest probability of resulting in conversions, and then identifying customers, both online and offline, who exhibit these behaviors. Especially in the case of online commerce, the use of machine learning and AI for lead generation and conversion is essential, because of the enormous volume, complexity and 'real-timeness' of the data. Even attributing various search keywords to final conversion requires highly scalable algorithms that run iteratively and continuously considering the enormous number of keywords that are possible for an online search. In marketing, Potharst, Kaymak, and Pijl (2001) have used neural networks to target consumers who are likely to respond to direct mailings by a Dutch charity organization. They find a 70% response rate when the mailing is guided by neural networks compared to an overall 30% response rate based on traditional mailing. Kim, Street, Russell, and Menczer (2005) use a particular type of genetic algorithm, called the evolutionary local selection algorithm (ELSA) to do 'feature selection'—searching through the myriad features (demographic variables) to identify the feature subset that maximizes classification accuracy. The features selected by ELSA are then used to learn an artificial neural network that predicts 'buy' or 'not buy'. The authors show that this machine learning approach dominates the traditional methods of feature selection (done by principal components analysis) and classification (done by logistic regression).

A recent article featured an interesting application of machine learning and AI in the lead generation efforts of the Harley-Davidson company (Power, 2017). It presents the case of a Harley-Davidson dealership in New York, which, by using AI algorithms, went from getting one qualified lead per day to 40. Many of these new leads were "lookalikes" in the sense that these potential customers resembled previous high-value customers and, therefore, were more likely to make a purchase. In a spectacularly successful application of AI, the machine learning algorithm used the data generated through many text and visual campaign variables and customer variables to predict which online campaigns, implemented through different digital channels (SMS text, email, search, display, social media, etc.), were most likely to convert different customer segments. Within three months of implementing this machine learning and AI based lead generation and qualification program, the dealership's qualified leads had increased 2930%. In terms of lead generation, AI's main contribution comes from its ability to target customers with highly personalized, individually tailored advertising and marketing. Since machines do not need to sleep or rest, they can run continuously to compute personalized campaigns even for companies with a very large number of customers and prospects. Once the leads are generated and qualified and the optimized contact strategy for each lead is determined, AI can accurately and at scale figure out when and how prospects should be contacted, thereby putting more leads in the funnel and driving greater sales productivity.

4.1.4. Research questions

The previous discussion raises interesting research questions. For the sake of brevity, for all stages of the sales process, we only address a few research questions in the text and relegate to Table 2 additional research questions and research avenues. Marketers have used many different dimensions to position their products vis-a-vis competition, but traditionally most of the dimensions have been product attributes. With the ability of machine learning methods to mine textual data and do sentiment analyses, positioning no longer needs to be static and abstract but can be based on existing and real-time customer evaluations. In this context, researchers can address the following research question: Can firms study customer sentiment for products using tools such as machine learning and AI to enhance the positioning of products and services when compared to traditional methods?

One of the advantages of using machine learning and AI is the dynamic nature of the techniques in that they can always be running in the background and provide real-time data. Similar to the research question in positioning, an interesting research question then arises: How can machine learning and AI enhance forecasting and demand estimation, specifically when the external environment undergoes dynamic, rapid, and unforeseen changes?

Finally, an area of attention is enhanced lead qualification, an area that can take a disproportional amount of any sales team member's time. An interesting research area would be the application of text mining and keyword analysis (e.g., identifying and validating qualifying words) to assist sales organizations in defining and implementing lead qualifications. Some firms are attempting to use machine learning and AI in enhancing lead qualification but a more rigorous research examination is needed.

4.2. Pre-approach and approach

The pre-approach and approach has typically been studied together in sales research and some researchers have indicated that the stages are being merged (Sheth & Sharma, 2008). In the pre-approach and approach stages of the sales process, the impact of digitalization has been dramatic. Internet telephony, which has been around for a few decades, is typically a hardware and software combination that enables telephone calls that use the internet as a transmission medium.

Over the last decade, we have also seen a rapid increase in the use of mobile and web-based means through which the selling organization can contact the customers. A company called *6sense* has developed a service that does predictive analytics by uncovering buying signals that helps sales people decide on the optimal time to approach a potential buyer. The company uses data on the visits customers make to the client's site, along with third-party data and social media feeds to predict when the customer may be ready to buy. The most exciting development in AI-powered conversational software has been the emergence of chat bots. A chat bot (interactive agent or artificial conversational entity) is a computer program that can simulate human behavior when it carries out a conversation with a person. Because they pass the Turing test, which requires a computer program to impersonate a human in real-time, chat bots can be used by the selling organization to uncover the needs of the customer (through striking up a dialog) and then to process the order and payment, especially for routine purchases. Chat bots are powered by AI, which themselves depend on machine learning, and are flexible enough to allow the salesperson to take over the conversation, as needed, for more complex sales. The simpler chat bots, called *rule-based* chat bots, use specific keywords. Thus, if a promotional offer says, “Text VACATION to 77777 to get $10% off your ticket”, then customers have to text VACATION as written (all caps), since the software recognizes only the word in the form expressed by the text message. A misspelling or a text request saying “Give me 10% off on my ticket” will not be recognized. In contrast, AI-based chat bots perform sophisticated Natural Language Processing (NLP) and can be trained to understand the intent behind the request. In terms of machine learning tools, most AI-based chatbots use Markov chains to “…build responses that are more applicable probabilistically and, consequently, are more correct” (Abdul-Kader & Woods, 2015, p. 73). Recently genetic algorithms have been proposed that can build a new sentence depending on sentences retrieved from the database. Moreover, an important part of a chat bot is its classifier, which is the module that does filtering and segmentation of the input entered by the user, and this step benefits from all the advances that machine learning tools make in improving segmentation (Abdul-Kader & Woods, 2015).

What does this mean for sales staff? Clearly, members of any sales team will have to develop ‘machine intelligence’ (Baumgartner, Hatami, & Valdivieso, 2016), since they will have to deal with machines at various stages of the sales process. Even a cursory appreciation of what the AI-based machine can do, and how it does so, will allow a sales representative with machine intelligence to interpret each stage of the sales journey (project/prospect/opportunity) when handling the account. At some point, the machine may hand the baton to the sales representative, and give the reason the machine transferred the account. Conversely, at some point the sales representative may want to initiate the transfer of the account to the machine because s/he knows that the machine will be able to execute the remaining stages of the sales process more efficiently. This understanding of the capability boundaries of AI in sales will enable the salesperson to interact with the AI systems much more efficiently, to the point of treating the machine as part of the ‘team’ in closing the sale.

4.2.1. Research questions

The previous discussion raises an interesting research question. With virtually unlimited data on firms and individuals available, can machine intelligence enhance the pre-approach and approach stages of the selling process. Specifically, can text mining coupled with machine learning and AI enhanced pre-approach and approach methods override or replace traditional methods?

4.3. Presentation

In the presentation stage, major breakthroughs are being brought about by immersive technologies. Evidence has shown that immersive technologies drive higher user engagement than ‘plain’ videoconferencing by eliminating, or at least reducing, a major drawback of videoconferencing—the sense of the presenter not being present in the room. This element is missing in the major immersive technologies, which include mobile virtual reality (VR), 360-degree video and augmented reality (AR)—all of which enable the sales representative to make presentations remotely that very closely resemble the quality of on-site demos. Research has established that virtual reality also enhances consumer learning, especially in situations where consumers find it hard to imagine the product without an exact physical prototype (Suh & Lee, 2005). Since the cost of physical sales calls continues to increase, along with the hard costs of samples and shipping, remote presentations are increasingly attractive to companies. AI and machine learning drive virtual reality displays, in which often “…data is overlaid on a view of the outside world” (Dooley, 2017). In particular, machine learning fosters real-time image and audio processing as well as data visualization. These AI-backed systems can now be deployed from the cloud so that even smaller companies with limited budgets can avail themselves of these cutting-edge presentation technologies. As an example, Intel's RealSense Vision Processor can facilitate the presentation of prototypes in three dimensions and use advanced algorithms to process raw image streams in real-time, thereby computing high-resolution depth maps. Simultaneously, interactive product presentation tools are also being developed so that the machine learning embedded in them can train the system to recognize complex and nuanced communication patterns of the presenter and the audience, thus allowing digital-enabled remote collaboration. In retail, virtual showrooms use machine learning algorithms that can customize the buying experience, and virtual assistants provide interactive dialog between the customer and the sales organization. In the real estate industry, companies like

Kiawah Island Real Estate, SC, allows prospective customers to take virtual reality tours of its new-construction homes and to imagine themselves surrounded by personalized colors, carpets and furniture that aren't physically present and have yet to be added to the home plans. Given the advent made by these AI-based technologies in this space, the role of the realtor has to be rethought and redefined.

As a concrete example, selling in the financial services industry is being critically disrupted by Fintech firms which use robo-advisor apps (Friedberg, 2017). These are AI-driven bots that can explain complex financial products to potential customers, and can even answer frequently asked questions since they can draw from a database with much more information than is humanly possible for a salesperson to have. Clearly, in order to compete with these robo-advisors, salespeople will have to learn to better deal with ambiguity, offer the human-touch, and even anticipate and address the as-yet-unarticulated needs of the customer.

At the presentation stage, the selling organization also designs the offer (makes the 'product presentation') and often has to provide prototypes of the product or solution. Rapid prototyping has been around since the late 1980s and has been used to produce samples and prototypes without suffering from the diseconomies-of-scale of short production runs. Rapid prototyping is a name given collectively to a group of techniques comprised of three-dimensional computer aided design (CAD), 3D printing or "additive layer manufacturing" technology. The last two techniques are used to actually construct or assemble the parts. The major advantage of rapid prototyping is to make available the product prototype, and change it numerous times depending on the customer's design wishes, without the commitment of actual long production runs. Recently, machine learning and AI techniques have boosted the power of rapid prototyping. Garg and Tai (2014) show how the ensemble method of genetic programming (GP) and artificial neural networking (ANN) can be used to improve rapid prototyping by using these machine learning models to optimize the parameter settings that control the wear strength, tensile strength and other features of the prototype.

In many instances sales representatives are also tasked with making pricing decisions, at least in companies where pricing responsibility is delegated to the salesforce. A new company called *Optimizely* runs field experiments, often called A/B experiments in the online world, which involves experimentation with different price points and pricing formats. Using dashboards that incorporate pricing variables with other information on leads, contacts, IP addresses, etc., the machine learning based software can automate the experimentations process in real-time so that sales representatives can use data to figure out the best prices to quote to different segments of their customer base.

The sales profession has historically been interested in uncovering behavioral characteristics of sales representatives that can lead to better salesperson-customer interactions, thereby making their sales representatives more persuasive in influencing customer preferences and behavior. Surprisingly, as early as 1947, there were attempts to scientifically study how non-verbal cues like facial expressions and gestures, along with speech characteristics that could predict the quality of a sales encounter with a customer (Chapple & Gordon Jr, 1947). It is quite remarkable that such an early breakthrough attempt to explain sales performance was designed in such a sophisticated manner. The 'Interaction Chronograph' used in that study was operated by the experimenter, and to quote the authors, "Her responsibility is to watch primarily for facial activity, recording muscular contractions of speech and gesture as activity, and relaxation of the muscles as inactivity," (Chapple & Gordon Jr, 1947, pp. 173–174). Peterson, Cannito, and Brown (1995) investigated how a salesperson's voice characteristics affected selling effectiveness. The following is a brief discussion of work in marketing and sales that uses machine learning and AI to understand consumers and their buying behavior since this aspect of marketing provides fertile ground for salespeople to succeed.

Bonoma and Felder (1977) made an early investigation of non-verbal cues in marketing communications. A recent summary of how marketing academics and practitioners have been using analytics and big data through the years was presented by Wedel and Kannan (2016). The advent of machine learning and AI have given a fresh lease on life to attempts to understand the role of voice characteristics and all other non-verbal cues because researchers are now able to handle 'unstructured data' at much larger volumes and in real time. In marketing, text data has been used most frequently to understand consumers' behavior through their verbal expression of emotion, especially in product reviews which are known to considerably affect their product preferences (Schindler & Bickart, 2012; Yin, Bond, & Zhang, 2017). More recently, other types of data like video and pictures (Pieters & Wedel, 2004; Pieters, Wedel, & Batra, 2010) and non-verbal cues have been used. Thomas and Summers (2002) investigate how non-verbal cues like eye gaze, speech hesitations, gestures, clothing, and posture affect buyers' impressions of the salesperson in an industrial selling context.

Many of the most recent developments in analyzing text and other verbal and non-verbal cues depend on breakthroughs in 'computational linguistics' which is an interdisciplinary field that uses methods from computer science to perform statistical or rule-based analyses of natural language. Starting with the early work of Weizenbaum (1966), the creator of ELIZA, computer scientists have made tremendous strides in processing natural language using machine learning (Powers & Turk, 1989).

4.3.1. Research questions

An intriguing area of future research can be dynamic presentations. Researchers have addressed the issue of adaptive selling, changing the sales approach based on the needs of the customer. Can machine learning and AI help design dynamic sales presentations by incorporating customer feedback (verbal and non-verbal) and salesperson's goals while conducting a presentation? Dynamic presentations will need big data, machine learning and AI programming, and high processing platforms during the presentation.

4.4. Overcoming objections in closing stage

As in the case of pre-approach and approach, overcoming objections to the closing stage is emerging in research as is acclimation of sales management to the evolving stages as traditional face-to-faces sales give way to automation but the most important customers are given improved service by a more focused sales force at both a national and global level (Sheth & Sharma, 2008). Overcoming objections to AI intervention and especially to robo-advisors in the closing stage is likely to see the most involvement of sales representatives despite the advances in AI-based system communication with customers. Standard FAQs may not have the ability to address all the concerns of customers, especially those that arise from more complex sales of personalized products and solutions. Closing for simpler orders is facilitated by Internet based tools such as creating accounts and follow-up through emails, text, etc.

4.4.1. Research questions

An area of sales that requires attention but has not been addressed by firms in general is lost-order audits and incorporating lost-order audits into future sales approaches when handling objections. Thus, the question at hand is: Can big data, machine learning and AI be combined to serve, maintain and grow accounts, including those who have undergone lost orders? In other words, how can machine learning and AI be used to conduct a 'lost order audit' and help sales organizations better anticipate objections and handle them?

4.5. Follow-up

Follow-up is two different processes. The first one is filling the current order and the second one is follow-up after the current order is filled. Filling the current order broadly as comprises of recording the order, initiating order processing, inventory management and order fulfilment via supply chain and procurement systems. Singh and Huhns (1994) document the effectiveness of an early order processing system that uses AI to automate all work flows that are required for order processing. Some of the actions required for order processing range from simple things like routine paperwork to sophisticated inventory management, which in turn require accessing supply chain information, and processing payment securely. In the world of multi-channel marketing, orders are received through salespeople, email, fax, internet or a combination of the above. All orders need to be accurately identified, recorded, interpreted and transmitted to the firm's ERP system as fast as possible and this calls for automated processes. Most importantly, orders must be validated and then price, availability and adequate replacement products have to be accurately assessed. Customer service and/or maintenance must follow. The early order processing systems could carry out the more routine tasks associated with order processing, but modern AI backed systems allow for automation even in areas that call for judgment, for example in cases where the required information resides in different databases all over the company (at different branches and possibly at headquarters), the data can be recorded in different formats and have errors and missing values.

Machine learning and AI have made some inroads in the post order follow-up stage of the sales process. Gainsight, a company that offers software to manage sales and customer service, has linked up with the questionnaire company, Survey Monkey, to automatically alert salespeople of the need to invoice after the close. The proprietary software prompts salespeople on upsell and cross-sell opportunities, and ensures that all members of the sales team are on the same page as far a delivery timelines and schedules are concerned. In an early example of applying knowledge-based systems to the problem of cross-sales, Anand, Hughes, and Bell (1998) provide a data mining approach that emphasizes rule discovery and deviation detection. The approach identifies the best 'characteristics' that can be used to predict potential customer targets for cross-sales. Characteristics are pairs of attributes that exist in a particular group of records, which are the records of customers who are currently purchasing a product from the company. Once such characteristics of current customers of a product are known, the customers in the target data set who are likely to purchase some other product of the firm are then identified. The purchase data are augmented with lifestyle data and other survey data to increase predictive accuracy.

Electronics giant LG Electronics has very recently unveiled a plan to that will use AI, enabled by machine learning technology, to offer customized after sales services for the benefit of the users of its smartphone (Mu-Hyun, 2017). The company realized that 80% of user visits to its service centers are for simple and minor requests for software problems, and that these could be efficiently and effectively handled by automated processes. The customized remote service is expected to drastically cut down the number of visits to its service centers, which will then free up the time of after-sales service personnel to focus on major repairs (addressing product problems after the sale) that are knowledge intensive and time-consuming. For manufacturing firms, a major aspect of after-sales service is the continued availability of service parts and their pricing. Most manufacturers still rely on past practices of laboriously poring over detailed Excel spreadsheets to properly price their service parts. This results in parts being offered at different prices at different service centers at different times of the year, and often leads to customer frustration and complaints. A full optimization of after-sales service profits and revenue requires optimal setting of prices, and ensuring timely delivery.

4.5.1. Research questions

There is a huge gap in the sales and service functions of a firm (Rapp et al., 2017). An area of research interest can be: How can machine learning and AI better streamline sales-service links? Another area of concern from firms is that the functional areas do not talk to each other. Therefore, a research area can be: How can machine learning and AI create better advanced warning systems from the sales frontline to the supply chain and other related functions?

5. Conclusions

So far, the greatest impact of automation and technology in sales has been, and continues to be, on all routine, standard and repeatable activities. In these cases, technology acts as a supporting role to make the selling functions more efficient. Going forward, perhaps the greatest impact of digitalization in sales will be in all the activities and efforts that go into understanding customer behavior in order to design and deliver highly customized offerings. Thus, in the future, technology will act as an active decision-facilitator, maybe even a decision maker in some cases, that can act in close collaboration with the salesperson to enhance the latter's effectiveness. Some examples of customer behavior are development of consideration sets, development of preferences and utilities from consumption, social influence, and, buying patterns. This understanding is critical to the success of sales strategies.

In business-to-business contexts, the salesperson or the selling team is also required to understand the needs and interactions of the buying center in order to obtain sales. These include, all the activities in the buying process, the sequence in which they are performed, the identity of the actors performing them, and the importance of the various actors throughout the buying process. The complexity of buying centers, which includes the size and the variety of stakeholders in it, continues to increase and it is very difficult for salespeople to interact successfully with buying centers in this complex environment.

We suggest that with the advent of the Fourth Industrial Revolution and the increased use of machine learning and AI algorithms and models can simulate the buying center with all its complexities, and can inform salespeople about the data to use, allowing salespeople to anticipate roadblocks and potential pitfalls when engaging with the buying center. In this paper, we discuss machine learning and AI in the context of the seven steps of selling. We suggest that the impact of machine learning and AI on the personal selling and sales management function will be profound. This paper provided exemplars and featured areas of future research, which we hope will be an impetus for future research in these areas.

Acknowledgements

This article was an invited article and did not undergo the traditional IMM review process. The paper was blind reviewed by three reviewers.

References

Abdul-Kader, S., & Woods, J. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications, 6*(7), 72–80.

Anand, S. S. A. R. P., Hughes, J. G., & Bell, D. A. (1998). A data mining methodology for cross-sales. *Knowledge Based Systems, 10*, 449–461.

Artificial intelligence (2017). Merriam Webster.com. Retrieved November 1, 2017, from https://www.merriam-webster.com/dictionary/artificial%20intelligence.

Baumgartner, T., Hatami, H., & Valdivieso, M. (2016, June). Why salespeople need to develop machine intelligence. *Harvard Business Review*. Online only and can be accessed at: https://hbr.org/2016/06/why-salespeople-need-to-develop-machine-intelligence.

Ben-Hur, A., Horn, D., Siegelmann, H., & Vapnik, V. (2001). Support vector clustering. *Journal of Machine Learning Research, 2*, 125–137.

Blank, S. (2013). *Four steps to the epiphany* (2nd edition). K and S Ranch.

Blum, A., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence, 97*, 245–271.

Bonoma, T. V., & Felder, L. C. (1977). Nonverbal communication in marketing: Toward a communicational analysis. *Journal of Marketing Research, 14*, 169–180.

Brooks, G., 2017a, Three major impacts of machine learning on manufacturing today, https://insidebigdata.com/2017/06/01/3-major-impacts-machine-learning-manufacturing-today/.

Brooks, R., 2017b, The Seven Deadly Sins of AI Predictions, https://www.technologyreview.com/s/609048/the-seven-deadly-sins-of-ai-predictions/

Carbonell, J. G., Michalski, R. S., & Mitchell, T. M. (1983). An overview of machine learning. *Machine learning* (pp. 3–23). Berlin Heidelberg: Springer.

Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research, 184*, 1140–1154.

Chapple, E., & Gordon, D., Jr. (1947). An evaluation of department store salespeople by the interaction chronograph. *Journal of Marketing,* 173–185.

Dooley, B. (2017), Why AI with augmented and virtual reality will be the next big thing, https://upside.tdwi.org/articles/2017/04a/04/ai-with-augmented-and-virtual-reality-next-big-thing.aspx

Dubinsky, A. J. (1981). A factor analytic study of the personal selling process. *Journal of Personal Selling and Sales Management, 1*(1), 26–33.

Florez-Lopez, R., & Ramon-Jeronimo, J. M. (2009). Marketing segmentation through machine learning models: An approach based on customer relationship management and customer profitability accounting. *Social Science Computer Review, 27*(1), 11–96.

Friedberg, B. (2017) 6 New Investing and Robo-Advisor Fintech Startups from Finovate, https://www.roboadvisorpros.com/investing-robo-advisor-fintech-startups-finovate/

Gareth, J., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning.* New York: Springer Science + Business Media.

Garg, A., & Tai, K. (2014). An ensemble approach of machine learning in evaluation of mechanical property of the rapid prototyping fabricated prototype. *Applied Mechanics and Materials, 575*, 493–496.

Gath, I., & Geva, A. B. (1988). Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*(7), 773–781.

Gupta, S., Keen, M., Shah, A., & Verdier, G. (2017). *Digital revolutions in public finance.* International Monetary Fund (ISBN 9781484315224).

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *Elements of statistical learning.* Springer Science + Business Media LLC.

Heath, D., Kasif, S., & Salzberg, S. (1993). Learning oblique decision trees. In: In R. Bajcsy (Ed.). *Proc., 13th Int. Joint Conf. on Artificial Intelligence* (pp. 1002–1007). San Francisco: Morgan Kaufmann.

Homburg, C., Müller, M., & Klarmann, M. (2011). When does salespeople's customer orientation lead to customer loyalty? The differential effects of relational and functional customer orientation. *Journal of the Academy of Marketing Science, 39*(6), 795–812.

Hruschka, H., & Natter, M. (1999). Comparing performance of feedforward neural nets and K-means for market segmentation. *European Journal of Operational Research, 114*, 346–353.

Huang, J.-J., Tzeng, G.-H., & Ong, C.-S. (2007). Marketing segmentation using support vector clustering. *Expert Systems with Applications, 32*, 313–317.

Jacobs, P., & Rau, L. (1988). Natural language techniques for intelligent information retrieval. *Association for Computing Machinery,* 85–99.

John, G. H., & Kohavi, R. (1995). Feature subset selection using the wrapper model: Overfitting and dynamic search space topology. In: *Proc., 1st Int. Conf. on Knowledge Discovery and Data Mining* (pp. 192–197). Menlo Park, CA: AAAI Press.

Kim, Y., Street, W. N., Russell, G. J., & Menczer, F. (2005). Customer targeting: A neural network approach guided by genetic algorithm. *Management Science, 51*, 264–276.

Knight, W., (2017) Put humans at the center of AI, https://www.technologyreview.com/s/609060/put-humans-at-the-center-of-ai/

Landt, F. W. (1997). *Stock price predictions using neural networks.* Leiden University.

Lawrence, S., Tsoi, A., & Gilles, C. (1996). *Noisy time series prediction using symbolic representation and recurrent neural network grammatical inference.* University of Maryland, Institute for Advanced Computer Sciences.

Lee, J. A. N. (1995). *Computer Pioneers/J.A.N. Lee.* Los Alamitos, CA: The Institute of Electrical and Electronics Engineers, Inc. (Arthur Lee Samuel bio from book also available as pdf at http://history.computer.org/pioneers/pdfs/S/Samuel(2017.pdf).

Loten, A., (2017), AI to Drive Job Growth by 2020: Gartner, https://blogs.wsj.com/cio/2017/12/15/ai-to-drive-job-growth-by-2020-gartner/

Malthouse, E., & Li, H. (2017). Opportunities for and pitfalls of using big data in advertising research. *Journal of Advertising, 46*(2), 227–235.

Marr, B. (2016) Why everyone must get ready for the 4th industrial revolution, https://www.forbes.com/sites/bernardmarr/2016/04/05/why-everyone-must-get-ready-for-4th-industrial-revolution/#26be9e2f3f90.

Marr, B. (2017). What is the Internet of Things—a complete beginner's guide in 2017, a Forbes Tech/Big Data post, 04/2017. Available at https://www.forbes.com/sites/bernardmarr/2017/04/10/what-is-the-internet-of-things-a-complete-beginners-guide-in-2017/#2e878a665982, Accessed date: 6 November 2017.

Mu-Hyun, Cho (2017) LG to adapt machine learning into after-sales service," http://www.zdnet.com/article/lg-to-adapt-machine-learning-into-after-sales-service/.

Mukherjee, S., Osuna, E., & Girosi, F. (1997). Nonlinear prediction of chaotic time series using support vector machines. *Neural Networks for Signal Processing [1997] VII. Proceedings of the 1997 IEEE Workshop. Added to IEEE Explore in 06 August 2002*http://dx.doi.org/10.1109/NNSP.1997.622433.

Murthy, S. K. (1998). Automatic construction of decision trees from data: A multi-disciplinary survey. *Data Mining and Knowledge Discovery, 2*, 345–389.

Node (2017) Node knows your next markets of opportunity, http://hello.node.io/.

Peterson, R., Cannito, M., & Brown, S. (1995). An exploratory investigation of voice characteristics and selling effectiveness. *Journal of Personal Selling and Sales Management, 15*(1), 1–15.

Pieters, R., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing, 68*(2), 36–50.

Pieters, R., Wedel, M., & Batra, R. (2010). The stopping power of advertising: Measures and effects of visual complexity. *Journal of Marketing, 74*(5), 48–60.

Plouffe, C. R., Williams, B. C., & Wachner, T. (2008). Navigating difficult waters: Publishing trends and scholarship in sales research. *Journal of Personal Selling and Sales Management, 28*(1), 79–92.

Porter, Kyle (2017) Why AI won't replace (great) salespeople, https://www.entrepreneur.com/article/292162

Potharst, R., Kaymak, U., & Pijl, W. (2001). Neural networks for target selection in direct marketing. *ERIM Report Series Research in Management* ERS-2001-14-LIS.

Power, B. (2017, May). How Harley-Davidson used artificial intelligence to increase New York sales leads by 2,930%. *Harvard Business Review*. Online only access at: https://hbr.org/2017/05/how-harley-davidson-used-predictive-analytics-to-increase-new-york-sales-leads-by-2930.

Powers, D. M. W., & Turk, C. C. R. (1989). *Machine learning of natural language.* Springer-Verlag ISBN 978-0-387-19557-5.

Rapp, A. A., Bachrach, D. G., Flaherty, K., Hughes, D. E., Sharma, A., & Voorhees, C. M. (2017). The role of the sales-service interface and ambidexterity in the evolving organization: A multilevel research agenda. *Journal of Service Research, 20*(1), 59–75.

Ratnaparkhi, A. (1997). *A simple introduction to maximum entropy models for natural language processing.* University of Pennsylvania Institute for Research in Cognitive Science Technical Report No. IRCS-97-08.

Sapankevych, N., & Sankar, R. (2009). Time series prediction using support vector machines. *IEEE Computational Intelligence Magazine, 4*(2), 24–38.

Schindler, R., & Bickart, B. (2012). Perceived helpfulness of online consumer reviews: The role of message content and style. *Journal of Consumer Behavior, 11*(3), 234–243.

Sheth, J. N., & Sharma, A. (2008). The impact of the product to service shift in industrial markets and the evolution of the sales organization. *Industrial Marketing Management, 37*(3), 260–269.

Singh, M., & Huhns, M. (1994). Automating workflows for service order processing: Integrating AI and database technologies. *IEEE Expert, 6*(5).

Srihari S., 2010, "Machine Learning: Generative and Discriminative Models," http://www.cedar.buffalo.edu/wsrihari/CSE574/Discriminative-Generative.pdf (Accessed 1 Sept 2017).

Suh, K.-S., & Lee, Y. E. (2005). The effects of virtual reality on consumer learning: An empirical investigation. *MIS Quarterly, 29*(4), 673–697.

Syam, N. (2017). *Analytics of staged pipeline processes: A continuous time Markov chain model of win prediction.* Working PaperUniv of Missouri.

Thomas, L., & Summers, J. (2002). An initial evaluation of industrial buyers' impressions of salespersons' nonverbal cues. *Journal of Personal Selling and Sales Management, 22*(1), 41–53.

Verbeke, W., & Dietz, B. (2011). Drivers of sales performance: a contemporary meta-analysis. Have salespeople become knowledge brokers? *Journal of the Academy of Marketing Science, 39*(3), 407–428 June.

Wedel, M., & Kamakura, W. (2000). *Market segmentation: Conceptual and methodological foundations* (2nd edition). New York: Springer Science + Business Media.

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing, 80*, 97–121.

Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36–45.

Yin, D., Bond, S. D., & Zhang, H. (2017). Keep your cool or let it out: Nonlinear effects of expressed arousal on perceptions of consumer reviews. *Journal of Marketing Research, 54*(3), 447–463 June.