MIT Open Access Articles

Soul and machine (learning)

Soul and Machine (Learning)

Soul and Machine (Learning)[1]

Davide Proserpio, John R. Hauser, Xiao Liu, Tomomichi Amano, Alex Burnap, Tong Guo,

Dokyun (DK) Lee, Randall Lewis, Kanishka Misra, Eric Schwarz, Artem Timoshenko, Lilei Xu,

Hema Yoganarasimhan[2,3]

---

[2] Authors affiliations: USC Marshall School of Business, MIT Sloan School of Management, NYU Stern School of Business, Harvard Business School, Yale School of Management, Duke Fuqua School of Business, CMU Tepper School of Business, Independent, UCSD Rady School of Management, University of Michigan Ross School of Business, Northwestern University Kellogg School of Management, Airbnb, UW Foster School of Business.

[3] Corresponding author: Davide Proserpio, email: proserpi@marshall.usc.edu

Abstract

Machine learning is bringing us self-driving cars, medical diagnoses, and language translation, but how can machine learning help marketers improve marketing decisions? Machine learning models predict extremely well, are scalable to "big data," and are a natural fit to analyze rich media content, such as text, images, audio, and video. Examples of current marketing applications include identification of customer needs from online data, accurate prediction of consumer response to advertising, personalized pricing, and product recommendations. But without the human input and insight—the soul—the applications of machine learning are limited. To create competitive or cooperative strategies, to generate creative product designs, to be accurate for "what-if" and "but-for" applications, to devise dynamic policies, to advance knowledge, to protect consumer privacy, and avoid algorithm bias, machine learning needs a soul. The brightest future is based on the synergy of what the machine can do well and what humans do well. We provide examples and predictions for the future.

1. Without a soul, machine learning is but a tool

In today's information-rich environment, firms use real-time data and sensor signals combined with predictions of consumer response to automate decisions. Accurately predicting consumer reactions and competitor responses to marketing strategies remains a fundamental challenge. If used judiciously, machine learning—a set of algorithms, both supervised and unsupervised, that apply to large data to inform decisions—can greatly improve actionable predictions.[4]

Machine learning has made significant advances in recent years. Today we see progress in areas such as self-driving cars, recommender systems, automated conversational agents, automated advertising allocation and auctions, machine translation, and financial fraud detection. Marketing practice has already benefited from many of these advances, and firms of all sizes employ production-level machine learning systems to improve targeted advertising campaigns, the products offered to individual consumers, prices and promotions. Marketing has just begun to leverage machine learning approaches to create new powerful applications, offer new insights, and generate new theories. In this paper, we take a step back and ask: How can we best integrate machine learning to solve previously untenable marketing problems?

Machine learning is evolving, and the integration of marketing and machine learning is still in its infancy. While machine learning improves many tasks, machine learning without a "soul" is just a series of computations mapping inputs to outputs. Soul is the human intuition, expertise, and institutional knowledge that crafts simple computations into art to provide valid and useful insights and actions. Consider identifying the voice of the customer from user-generated content (UGC) such as reviews, tweets, and other posts. Early seminal research used unsupervised and (mostly)

---

[4] Examples of machine learning algorithms include neural networks, gradient-boosted trees, variational autoencoders, probabilistic graphical models, and reinforcement learning.

automated methods such as Latent Dirichlet Allocation (LDA) which automatically places the words in a corpus into buckets (Blei, Ng, and Jordan 2013, Büschken and Allenby 2016). Without soul and intuition, analysts risk interpreting the "bags of words" as a psychology patient might interpret a Rorschach inkblot.

By merging human insight with advanced machine learning, Timoshenko and Hauser (2019) describe an approach that leverages UGC to identify customer needs. The machine is a neural-network based machine learning model, which identifies relevant content and removes redundancy from a large UGC corpus in a cost-effective and scalable manner. The final step integrates human analysts to review the selected, informative content and formulate customer needs. Someday, natural language processing might be ready to find the nuggets of customer needs in the identified sentences, but until that day comes, we must rely on the intuition of analysts and the ability of managers to interpret the results.

The soul, whether it be intuition, creativity, grounded theory, or expertise, guides the transformation of raw data into actionable insights. Formalized intuition, i.e., theory in marketing, economics, psychology, statistics, engineering, and design science enhances the soul. Consider the threats to privacy coming from advertising networks that use extensive information to target individual consumers. Machine learning helps each firm target better, but does machine learning benefit all? A theory is necessary to formalize this normative question, and to understand the complex set of interactions among firms, consumers, and advertising networks. One can draw on economic and marketing theory to understand the endogenous competitive reactions and identify the types of information used to target. The latter is necessary because some information invades privacy and other information does not. The classification of information types requires human insight as does the choice of which effects to model explicitly (Rafieian and Yoganarasimhan 2020).

Theory cannot be automated and, despite machine learning advances, faces the same challenges that Little articulated in 1970—it must be simple, robust, easy to control, adaptive, complete on important issues, and provide clear insight. These judgments are the art derived from soul.

Even with good theory, "engineering" a machine learning solution requires soul. For example, Burnap et al. (2019) propose a machine learning approach to facilitate product design alongside human input. Researchers must decide how to integrate automation into a firm's product pipeline, and develop a model to leverage multiple datasets, typical for practical applications. This includes deciding model structure, e.g., types of layers in a neural network, how to balance competing objectives, and how to regularize models using domain knowledge. Engineering a machine learning algorithm tailored to a given domain is an art that requires soul.

In this paper, we present an overview of emerging machine learning areas that are being applied to marketing problems, give concrete examples that stem from the recent marketing literature, and discuss key benefits and limitations of these approaches. The theme throughout is that the manager's and analyst's soul remains in charge and is critical to successful applications.

2. Soul and Machine: Where are we heading?

We present an outlook on where marketing and machine learning (ML) are headed. In section 2.1 through 2.5, we introduce five promising areas of ML that we believe can shape the future of marketing. In each section, we describe the relevant marketing applications, the ML algorithms, the soul involved, the big data problem solved, the advantages and limitations of ML over traditional methods, and the importance of the soul. Table 1 provides an overview of these five areas.

Table 1: Overview of Emerging Machine Learning (ML) Areas for Marketing Applications

	Theory-Driven Machine Learning	Generative Models	Causal Inference	Reinforcement Learning	Algorithmic Transparency and Accountability
Marketing Application	Game theoretic strategy (e.g., targeting in auction environments)	Product design, Data fusion of consumer and market data.	Policy evaluation (e.g., personalization, disclosure)	Dynamic strategy (e.g., sequential pricing)	Content marketing, privacy protection, bias correction for fairness, hypothesis development
Machine (Learning)	Game theory model with empirical tasks done by ML	Variational autoencoders, generative adversarial networks, autoregressive Models	Causal Forest, Double/ Debiased ML	Bandits, Deep Q Learning	Salience extraction, concept extraction, attention, anomaly detection, interpretable ML
Soul	Researcher's choice of theories; Selection of which effects are important to model	Researcher's ability to integrate ML with human intuition and creativity of marketing tasks.	Researcher's ability to find exogeneity and model endogeneity; What can be "black box" and what cannot.	Researcher's choice of theories can inform where we want to focus learning	Researcher's objectives, preferences, judgment about whether interpretability has been achieved
Big Data 3V Problem Solved	Volume	Volume, Variety	Volume	Volume, Velocity	Volume, Variety
Traditional Marketing Example	Game theory models	Conjoint analysis	Structural models	Markov decision processes	Parameters identification
Advantages of ML over Traditional Methods	Scalability	Automation, scalability	Scalability, flexibility	Scalability and real-time deployment	Scalability, flexibility
Limitations of Off-the-Shelf ML Models	Do not include domain knowledge and structure	ML not tailored to marketing domain knowledge	Focus on prediction instead of policy evaluation	Low interpretability in complex environments	Algorithm & data bias, lack of appropriate objective function
Why We Need The Soul	The theoretical framework, the agent decision-making rules, and the phenomena to model explicitly are chosen by the researcher	Marketing tasks cannot yet be fully automated. Generative models must be integrated with humans in marketing processes.	The arguments to support the causal inference assumptions: SUTVA, Unconfounded-ness, and positivity are provided by the soul	The state, action, and reward are defined by the soul; The use of these can increase the rate of learning and lower the costs of experimentation	The fairness and unbiasedness criteria are often in the eye of the beholder. Interpretability is, by definition, a human judgment
Representative Papers	Rafieian and Yoganarasimhan (2020)	Burnap et al. (2019), Dew et al (2020)	Guo et al. (2017); Yoganarasimhan et al. (2020)	Schwartz et al. (2017); Misra et al. (2019); Liu (2020)	Proserpio et al. (2016) Liu et al. (2019); Lu et al (2020); Lee et al. (2020);

2.1. Theory-Driven Machine Learning

ML algorithms have proven to be successful at challenging games like Chess (Deep Blue by IBM), Shogi (Bonaza), and Go (AlphaGo by Google). Two key features form the foundation of this success. First, researchers started with a theoretical model of the structure of the game, which allowed them to break down the problem into a series of empirical tasks that could be solved using ML techniques. Second, the ML models were trained on a wide range of data, which allowed the model to learn optimal strategies for any board configuration. In this way, the ML algorithms learned to respond to the broad range of possible scenarios that they might encounter.

Many marketing problems also have a game structure or optimization objective that can benefit from a solution concept that mirrors those of AlphaGo or Bonaza (Taddy 2018). Marketers face two critically important types of problems: (1) substantive questions, e.g., what is the ROI of a marketing intervention? And (2) prescriptive questions, e.g., how can a two-sided platform design a selling mechanism to maximize its revenues? An important issue in these types of questions is policy evaluation (for substantive questions) or counterfactual "what-if" evaluation (for prescriptive questions), i.e., understanding how some outcomes of interest would evolve under a different treatment regime or a new data generating process than the one observed in the data.

To solve these types of problems, marketers use a theoretical framework (soul) to break down the problem into a series of small empirical tasks that are then solved using nonparametric ML methods that are scalable to extremely large datasets. These ML models have high predictive accuracy and, when combined with data that have sufficient exogenous variation, they are excellent tools for forming and evaluating counterfactual policies.

A small but growing stream of literature in marketing has adopted a combination of theory-driven frameworks and ML methods to answer important substantive and prescriptive questions. For

example, Rafieian and Yoganarasimhan (2020) adopt this approach to examine the incentives of mobile ad-networks to engage in micro-targeting. Rafieian (2019a) and Rafieian (2019b) study the general problem of adaptive ad-sequencing in non-strategic and strategic environments (auctions). Both papers start with a theoretical framework, where the researchers choose the appropriate theories and which outcomes are important to model.

In the examples above, the soul and ML are both indispensable: without the soul, ML would not be useful because ML lacks domain knowledge to formulate the problem with the proper structure; and without the highly scalable and accurate ML models, it would not be possible to model theories and measure outcomes in real-time. In the future, we expect more problems like real-time and dynamic optimal pricing and advertising to be solved at scale.

2.2. Generative Models

Product management is an essential marketing function. Due to its ability to handle different types of data (e.g., images, text), in large volumes (thousands of new designs), and that accumulates rapidly (velocity), ML can improve product management by changing how firms generate and test digital and physical products. Technology firms, such as Airbnb and Netflix, routinely conduct thousands of A/B tests to optimize their digital products and user experience. The design of these products can be generated automatically using ML algorithms that efficiently sample the space of possible designs and simultaneously personalize it to different market segments. Marketing theory is critical for product management, as off-the-shelf ML may optimize objectives that are mathematically elegant but counterproductive to the firm. Indeed, Hauser, Liberali, and Urban (2014) show that modeling and taking into account consumer behavior, and explicitly accounting for switching costs, improves the ability to personalize the look and feel of websites to individual consumers.

Generating physical products, however, is more challenging than generating digital products; but marketing ML methods are making headway. Burnap et al. (2019) show that ML can be used to

improve the aesthetic design process for automobiles, including automatically generating new appealing products with Generative Adversarial Networks (GAN), a class of ML algorithms that, given a training set, can generate new data with similar properties (e.g., a GAN trained on images can generate new images that look authentic to the human eye).

In the product design settings, the soul is the researcher's ability to integrate ML with human intuition, and the creativity of designers to go beyond what is in the training set. The soul is critical because it provides modeling assumptions that explicitly account for highly-iterative and asynchronous design processes within the firm.

Looking forward, we expect to see generative approaches with ever-increasing sizes and diversity of data sources. Dew et al. (2019), for example, develop a deep generative model (i.e., variational autoencoder) that combines several data sources containing both text and images to help marketers generate new logo designs. Their work is an example of how ML techniques augment human intelligence and creativity during product management.

2.3. Causal Inference

Causal questions are important to marketing. For example, how much will sales increases if we increase advertising spending by 20%? Or how much total revenue would we earn if we launch a new product? Because causal inference involves assumptions and logic that are normally domain-specific, theoretical grounding is necessary to encode causality in ML.

Despite developments in ML for causal inference, the soul is key to define the model in which a causal claim is made, and to evaluate the critical assumptions that are necessary for identification. For example, recent developments in economics and statistics allow researchers to obtain individual-level treatment effects, e.g., generalized random forest (Athey et al. 2019) and double/debiased ML (Chernozhukov et al. 2018). These frameworks allow researchers to obtain robust, non-parametric individual-level estimates with good statistical properties for inference at scale, allowing for causal analysis that does not require prior knowledge about the structure of heterogeneity. However, despite the power of ML to model heterogeneous treatment effects from a

large volume of data, the researcher's institutional knowledge—the soul—is key as it recognizes exogeneity and ensures that assumptions of causal inference such as unconfoundedness are satisfied.

Modeling heterogeneity facilitates a richer mapping between theory and empirical data, and a more nuanced understanding of mechanisms that is at the core of evaluating and designing marketing policy (such as targeted interventions). Guo et al. (2017) evaluate the consequences of disclosing firms-to-physician payments and find that firms differentiate in their physician-targeting strategies without cutting the total expenditure. Yoganarasimhan et al. (2020) compare different ML-methods that can be used to design and evaluate personalizing targeting policies based on causal frameworks. The area of ML for causal inference is growing rapidly and we expect more work in the area of causal inference, policy design, and policy evaluation at scale in the future.

2.4.   Reinforcement Learning

Among the greatest strengths of ML is the ability to continuously learn and improve predictions and optimize actions through repeated interactions with the environment. In the ML literature, these algorithms are called reinforcement learning. In marketing, as well as in economics or psychology, this process is known as *learning by doing*.

Many ML algorithms seek to make predictions or recommend actions based on the data, e.g., predict whether customers will buy in the next quarter, or optimally select customers for targeting. Learning by doing (or optimal or near-optimal experimentation) expands the focus of ML to evaluate whether and how we should collect more data to improve future predictions or recommendations. The soul provides marketing theories, identifies what needs to be learned, and how we might experiment or collect data to learn. ML algorithms leverage large volumes of data to generate predictions and inferences conditional on learning and (near) optimize how to learn. New data update predictions, inferences, and subsequent recommendations in real-time, and identify what data, if any, are needed. The process continues.

Learning-by-doing has been used for several tactical decisions such as promotion and pricing. For example, Hauser et al. (2014) and Schwartz et al. (2017) consider the online promotion

question—which ads are most effective at attracting new customers? —using multi-armed bandits. Misra et al. (2019), instead, address the pricing question for an e-commerce retailer point of view— what is the profit-maximizing price for e-commerce products? Finally, Liu (2020) creates dynamic personalized pricing strategies using Deep Q-learning.

In Schwartz et. al. (2017), the soul is in the form of a preference structure where advertisement conversion rates are a function of the underlying attributes, and this facilitates learning across advertisements; in Misra et. al. (2019), the soul is which economic theory to impose and which revealed preference axioms matter; and in Liu (2020), the soul leverages pricing theories to define states and actions. The researcher's insights about the critical features of the problem and consumer behavior significantly increase the rate of learning and reduce the cost of experimentation.

Progress has been substantial in this area, but many challenges lie ahead. For example, reinforcement learning raises the issue of algorithm collusion. If all competing firms use reinforcement learning, these actions might result in supra-competitive prices (Calvano et. al. 2018 and Hansen et. al 2020). The prices generated by these algorithms will be sub-optimal for all firms as each firm has a unilateral incentive to set lower prices; moreover, these prices could be potentially harmful to consumer welfare. Research on competition laws is needed to account for autonomous pricing agents and the, potentially negative, externalities they create (Harrington 2018). In this case, the soul is the perspective that goes beyond the individual application to model the joint application of ML by many or all firms in the market.

2.5.   Algorithmic Transparency and Accountability

While ML techniques applied to unstructured data have evolved over the last decade, the marketer's ability to extract actionable insight remains elusive. Many ML algorithms lack transparency and interpretability, and, hence, struggle to generate theories and new insights. ML can provide accurate classifications or predictions but rarely provide intuitive explanations. In response to this problem, a new stream of research called eXplainable Artificial Intelligence (XAI) has

thrived. XAI research seeks to 1) create inherently interpretable and high performing ML algorithms (Chen et al 2019); 2) create transparent post hoc algorithms to mimic black box outputs with high fidelity (Ribeiro et al 2016); or 3) define characteristics of good explanation (e.g., Lu et al 2020). XAI could aid marketers in creating new theories by building ML algorithms that model consumer behavior. A recent example is the Focused Concept Miner (Lee et al. 2020), an algorithm based on a deep learning text-mining method that automatically extracts interpretable concepts from text that are highly correlated with user-provided business outcomes. Similarly, any unstructured and structured consumer data can be processed at scale through XAI methods to discover interesting patterns in the data. But these patterns must still be interpreted by analysts (soul) to explore new hypotheses, and ultimately develop new theories.

A related application of XAI in marketing is interpretable variable generation for hypotheses testing. Traditionally, marketing researchers generated testable hypothesis variables using prior knowledge or personal experience. However, when dealing with massive datasets containing unstructured data, it is more difficult to generate interpretable hypothesis variables. For example, Liu et al. (2019) use XAI to investigate which content information in product reviews affect consumer purchase behaviors. Prior literature suggests that price and quality information in reviews is likely to affect purchases; however, extracting a variable representing quality from text content is difficult because the definition of quality is not universally defined and often subjective and vague. To solve this problem, the authors develop a three-step approach to generate interpretable hypothesis variables. First, they create an end-to-end deep learning model that uses the unstructured review data as input and sales conversion as output. Second, after estimating this model, the authors perform a post-hoc analysis to extract salient n-grams that affect conversion the most. Third, the salient n-grams are clustered using topic modeling (Blei et al. 2003). The generated topics create sensible hypotheses, which are then formally tested using regression models.

XAI, however, is not a panacea, and applying XAI techniques requires researchers' judgment on whether interpretability has been achieved, as well as domain knowledge to audit these systems for any inconsistencies or bias. By generating and testing more hypotheses, humans create more

theories and accumulate more knowledge, and this iterative process cannot (at least for now) be automated, but it requires the soul.

A stream of ML called Fairness, Accountability, and Transparency (FAccT) explicitly recognizes the soul. Researchers realized that the output of ML algorithms might be biased due to intrinsic bias present in the training data; the potential harm of this bias is exacerbated by the scale and scope of these algorithms. For example, this bias can be particularly detrimental for work in the area at the intersection of marketing and public policy where researchers leverage alternative data sources to create economic measures and indexes to complement, and eventually replace, existing ones.

Many measures used by governments to make decisions (economic indices, consumer sentiment, etc.) are survey driven; but ML can soon augment or substitute for survey-based techniques. Researchers can scrape data at high frequency to collect publicly available information about consumers, firms, and social media, which can then be used to generate indices in real-time. For example, Nikhil et al (2014) developed "Streetscore", a scene understanding algorithm that predicts the perceived safety of a streetscape using Google Street View data. In terms of macro-economic indexes, both Antenucci et al (2014) and Proserpio et al. (2016) use Twitter data to predict labor market outcomes such as the unemployment rate. This research suggests a future where urban resource allocation is more efficient by using "big data." But is this allocation fair, unbiased, and equitable? ML algorithms cannot answer this question and this is why the soul—researchers and stakeholders (e.g., policymakers, governments, NGOs, advocacy groups—is extremely important in these settings. The soul must guarantee that the predictions generated by ML are fair, unbiased, and equitable. This is especially true for these types of ML applications because the input data (social media or user-generated content) are potentially representative of a small part of the population, but decisions and policies are applied to the entire population. Despite these challenges, recent research suggests that, with the right data and careful development, these ML

measures could be more precise and able to better predict the economic conditions at high granularity and in real-time.

3.   Conclusions

Managers and researchers are increasingly turning to ML tools to solve and optimize marketing decisions. But without a soul (i.e., human insight), the capabilities of machine learning will be limited or, worse, misused. The need for a soul tells us that, as of this writing, there is no magic answer. Algorithm A might work well for identifying metaphors while Algorithm B might be best for identifying product returns. It would be a mistake to denote any algorithm as "best." Many can be crafted for the job. The soul also tells us to avoid dogma—ML can implement theories or help to identify causality, but no method is perfect. All can be challenged and improved. Finally, the soul tells us to embrace diversity of perspectives whether they come from marketing, economics, psychology, design, computer science, or engineering. Each perspective helps customize an algorithm for the best possible outcome.

True advances come from the combination of the soul and machine learning. The soul is the hard-to-quantify and hard-to-pin-down creativity and insight of the marketing scientist who tunes the theory to direct the path of the development, uses insight to identify which algorithm is best, tunes the structure and parameters, and chooses the regularization that keeps the model in check and avoids overfitting. And, the soul is the marketing manager who must know when to trust machine learning outputs, when to question them, and when to use instinct (the sum of past experience) to make the final decisions.

In this paper, we discussed five areas of ML that are beginning to affect the discipline of marketing and that will likely shape the future of marketing. A common theme that emerged from

this discussion is that the soul and machine learning are iterative: the soul informs the design of the ML implementation which, in turn, informs the soul. As the field of ML matures, we anticipate many more of these iterations to take place.

References

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. The Annals of Statistics, 47(2), 1148-1178.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." *Journal of machine Learning research* 3, no. Jan (2003): 993-1022.

Burnap A, Hauser JR, Timoshenko A (2019) Design and evaluation of product aesthetics: a human-machine hybrid approach (Social Science Research Network, Rochester, NY).

Chen, C., Li, O., Tao, D., Barnett, A., Rudin, C., & Su, J. K. (2019). This looks like that: deep learning for interpretable image recognition. In Advances in NIPS (pp. 8930-8941).

Chernozhukov, V., D. Chetverikov, M. Demirer, et al. (2018). Double/debiased machine learning for treatment and structural parameters. The Econometrics Journal 21(1), C1–C68.

Emilio Calvano, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello. 2018. Artificial intelligence, algorithmic pricing and collusion. (2018)

Dew R, Ansari A, Toubia O (2019) Letting logos speak: leveraging multiview representation learning for data-driven logo design (Social Science Research Network, Rochester, NY).

Guo, Tong, S. Sriram, and Puneet Manchanda (2017). "The effect of information disclosure on industry payments to physicians." Available at SSRN: https://ssrn.com/abstract=3064769.

Harrington., Joseph E. Developing competition law for collusion by autonomous artificial agents. *Journal of Competition Law & Economics* 14, 3 (2018), 331-363.

Hauser JR, Liberali G, Urban GL (2014) Website morphing 2.0: switching costs, partial exposure, random exit, and when to morph. *Management Science* 60(6):1594-1616.

Lee, Dokyun (DK) and Manzoor, Emaad and Cheng, Zhaoqi, Focused Concept Miner (FCM): Interpretable Deep Learning for Text Exploration (Written May 20, 2018, Revised May 19, 2020). Available at SSRN: https://ssrn.com/abstract=3304756.

Little, JDC. 1970. Managers and models: The concept of a decision calculus. Management Science 16(8) B466-485.

Liu, X., Lee, D., & Srinivasan, K. (2019). Large-Scale Cross-Category Analysis of Consumer Review Content on Sales Conversion Leveraging Deep Learning. Journal of Marketing Research, 56(6), 918-943.

Liu, X. (2020). Dynamic Personalized Pricing Using Batch Deep Reinforcement Learning: An Application to LiveStream Shopping. New York University. Working paper.

Lu, Joy, Dokyun (DK) Lee, Tae Wan Kim, and David Danks. "Good Explanation for Algorithmic Transparency." Proceedings of the 2020 AAAI/ACM Conference on AI, Ethics, and Society. 2020.

Misra, Kanishka, Eric M. Schwartz, Jacob D. Abernethy (2019). Dynamic online pricing with incomplete information using multi-armed bandit experiments. *Marketing Science*, 38(2), 226-252.

Naik, Nikhil, Jade Philipoom, Ramesh Raskar, and César Hidalgo. "Streetscore-predicting the perceived safety of one million streetscapes." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 779-785. 2014.

Proserpio, Davide, Scott Counts, and Apurv Jain. "The psychology of job loss: using social media data to characterize and predict unemployment." In *Proceedings of the 8th ACM Conference on Web Science*, pp. 223-232. 2016.

Rafieian, Omid. 2020a. Optimizing user engagement through adaptive ad sequencing. Cornell University, Working paper.

Rafieian, Omid. 2020b. Revenue-optimal dynamic auctions for adaptive ad sequencing. Cornell University, Working paper.

Rafieian, Omid, Hema Yoganarasimhan. 2020. Targeting and privacy in mobile advertising. *Forthcoming, Marketing Science*.

Ribeiro MT, Singh S, Guestrin C (2016). Why should I trust you?" Explaining the predictions of a classifier, arXiv:1602.04938v3.

Schwartz, Eric M., Bradlow, Eric T., and Fader, Peter S. (2017). Customer acquisition via display advertising using multi-armed bandit experiments. *Marketing Science*, 36(4), 500-522.

Taddy, Matt. 2018. The technological elements of artificial intelligence. Tech. rep., National Bureau of Economic Research.

Timoshenko, Artem, John R Hauser. 2019. Identifying customer needs from user-generated content. *Marketing Science* 38(1) 1-20.

Urban GL, Timoshenko A, Dhillon P, and Hauser JR (2020), Cutting through the hype of deep learning, *MIT Sloan Management Review*, Winter, 71-76.

Yoganarasimhan, H., Barzegary, E., and Pani, A., "Design and Evaluation of Personalized Free Trials." University of Washington, Working Paper.