Is Deep Learning a Game Changer for Marketing Analytics?

Companies are already making sophisticated marketing decisions with data and analytics. Will Deep Learning enable a leap forward—or just marginal gains?

By Glen Urban, Artem Timoshenko, Paramveer Dhillon, and John R. Hauser

Glen Urban is the David Austin Professor in Marketing, Emeritus, at MIT's Sloan School of Management. Artem Timoshenko is an assistant professor of marketing at Northwestern University. Paramveer Dhillon is an assistant professor of information at the University of Michigan. John R. Hauser is the Kirin Professor of Marketing at MIT's Sloan School of Management.

October 2019

Sloan Management Review Vol. 61, No. 2 (2020): 71-76.

https://sloanreview.mit.edu/article/is-deep-learning-a-game-changer-for-marketing-analytics/

Deep learning is delivering impressive results in AI applications. Apple's Siri, for example, translates the human voice into computer commands that allow iPhone owners to get answers to questions, send messages, and navigate their way to and from obscure locations. Automated driving enables people today to go hands-free on expressways, and it will eventually do the same on city streets. In biology, researchers are creating new molecules for DNA-based pharmaceuticals.

Given all this activity with deep learning, many wonder how the underlying methods will alter the future of marketing. To what extent will they help companies design profitable new products and services to meet the needs of customers?

The technology that underpins deep learning is becoming increasingly capable of analyzing big databases for patterns and insights. It isn't difficult to imagine a day when companies will be able to integrate a wide array of databases to discern what consumers want with greater sophistication and analytic power and then leverage that information for market advantage. For example, it may not be long before consumers, identified via facial recognition technology while grocery shopping, receive individualized coupons based on their previous purchase behavior. In the future advertisements may be individually designed to appeal to consumers with different personalities and be delivered in real time as they view YouTube. Deep learning might also be used to design products to meet consumer's personal needs, which could then be produced and delivered through automated 3D printing systems.

Different types of organizations will try to harness the powers of deep learning in their own ways. An auto maker might use them to target new customers, revamp the buying process, or fine-tune product features a specific set of buyers will want. It could draw on a sea of relevant data to do all this, including auto repair data, consumer ratings on vehicle quality and reliability, car registrations, Twitter posts concerning the car-buying and user experience, Facebook posts showing people with their cars, manufacturers' consumer relationship management data files, and clicks on the Internet. A bank, meanwhile, could leverage deep learning to develop new products or services, and customize its promotions. By analyzing data on customer loan histories, credit card transactions, savings and checking account records, web site clicks, social media behavior, product ratings, and search histories, it could gain insights into the things certain customers value. What do 40-year-old professionals living in urban neighborhoods want most in a credit card? Do they prefer travel rewards, buyer protection, cash back, or low interest rates?

To be sure, a lot of managers already use analytics with statistical models and focused databases to track brand performance, schedule promotions, and make spending decisions. So, how is deep learning different? Is it a fundamental leap forward, or will it simply enable marginal gains? In this article we will examine these questions in relation to a study we conducted involving credit cards. In addition, we will consider what this research suggests about the future direction of analytics.

Although we are still in the early days of deep learning, it isn't too early to ask: What will it offer companies compared to existing analytics methods managers are accustomed to? Can it provide better predictions, and if not, how can it be improved? And what kinds of investments in data and technology will companies have to make in deep learning to take advantage of the latest and most powerful capabilities? Our research suggests that, while deep learning may not lead to large gains in predictive accuracy in every setting, there are reasons for optimism.

Our Experiment

To compare deep learning to traditional methods for marketing analytics, we studied a large database of click streams, demographics, and ad exposures relating to the credit card market from NerdWallet.com, a large online vendor of credit cards, based in San Francisco. We wanted to see if a multi-level deep learning model could predict credit card choices more accurately than traditional models.

The ability to predict customer choice is the first step toward improving decisions that go into product design, media resource allocation, how to promote the product (in this case, a credit card), and whom to target. Knowing what people value most requires experimentation and predictive choice modeling. Our hunch was that deep learning would provide a clearer, more useful picture than a simpler regression model. To test that assumption, we looked at the credit card selection processes of 260,000 individuals, taking into account 25 demographic factors (including obvious things like age, gender, and household income, and more detailed categories such as credit score, cards the consumer currently owns, and zip code); 132 attributes for each card (such as APR interest rate, reward points – miles, cash, point, and card fees – annual cost, transfer of balances); and the cards each person applied for.

The NerdWallet website aggregates information and expert reviews about some 2,000 different credit cards from hundreds of banks so that customers can compare features and decide what's important to them. The products that users view during their online sessions are recorded. NerdWallet also notes when a person clicks to see more details or selects particular cards for comparison. After users decide which credit card they want from the NerdWallet site, they are directed to the corresponding bank's website to complete an application. For the purposes of our study, we treated this action as an indication of eventual choice.

Using NerdWallet's database, we compared three choice models.[1] The first model was a straight-forward linear regression of choice as a function of the individual user demographics and card attributes. Each variable had a simple direct effect on choice through one equation, with no interactions among variables. For example, a variable such as reward miles might use a coefficient times the number of reward miles granted per thousands of dollars charged by the individual on the card to predict the probability that a user might choose the American Express Green card.

The second model was a simple deep learning model. The unique feature of deep learning models is the use of hidden layers between the input variables and the probability of choice. Each input variable (for example, reward miles) is linked to a latent variable, which in turn is tied to the probability of choice. Latent variables aren't specifically identified – rather, they are combinations of input variables at the first level. The essence of deep learning is that latent variables link to other latent variables in the next model level; subsequent latent variables are combinations of *other* latent variables. This may seem unnecessarily complexity, but linking latent variables across hidden layers is what gives deep learning its power. The links connecting latent variables across multiple levels is why the name deep learning is used. In our model we used three levels. The first layer captured the effects of the observed card attributes and demographics on a set of latent variables and how those latent variables affected subsequent layers of latent variables. The hope was that with three hidden layers that allowed for complex interactions and non-linearities, we could to predict card choice more accurately.

Our third model was an enhancement of the simple deep learning model, where the dependent measures were the probability of choice and the particular cards consumers considered in their final choice. It took into account granular information about the buying process and the other credit cards people considered (based on their click streams). Our hope was that by adding this additional step of consideration to the buying process, we might improve on the accuracy of the choice predictions of the basic deep learning model.

What We Found

How successful were the different models in predicting the credit card choices people made? Based on our analysis, the two models that drew on deep learning were able to predict card choices with more precision than the traditional approach. (See **"How Deep Learning Can Outperform Traditional Marketing Analytics."**) But the improvements were not as large as we thought they would be.

The simple regression had an accuracy rate of 70.5%, meaning that in 70.5% of the cases we were able to predict correctly which card a particular consumer would apply for and which ones he or she would not apply for. The simple deep learning model was slightly more accurate, at 71.7%, and the more sophisticated model had an accuracy rating of 73.0%.

We had thought that adding more latent layers would result in more dramatic improvements. But our expectations may have been overblown. Well-designed linear regression models refined by experience can be robust and accurate. The many parameters—that is, the many demographic and card attribute measures--enabled good predictions from the traditional approach.

Costs and Benefits

So what did our study tell us about the potential for deep learning? Does it make sense for marketing organizations to invest in the technology and develop the critical capabilities now, given how small the improvements in choice prediction may be? In our opinion, the small gains in predictive accuracy over what's possible using traditional data are not likely to generate sufficiently high returns in most cases to justify significant investments .[2] Before simply diving in, organizations need to look beyond the potential gains and consider the challenges and costs of implementation.

A major drawback of deep learning is that it's difficult to identify which variables drive the biggest response. Was it the credit card travel reward levels or the low interest rates that had a bigger influence on consumer choice? Because the variables are processed through so many latent layers affecting choice, it's hard to measure the impact of one factor. We can simulate the effect of changes in one variable, but the effect depends on the levels of all the other variables used in the complex structure of latent variables in the model. It can be difficult to link intuitive judgments of marketing managers to the model results, and this will make implementation difficult.

There are other factors to take into account. There are costs to acquire deep learning technology, staff required to implement it, and the additional data costs that may need to be incurred. Finally, even with today's fast computers, deep learning models need huge amounts of computer power and may experience long run times. This becomes a real limitation if real-time implementation is required. The reality is that in most situations a compelling case for deep learning as an alternative to traditional data analysis can't be made unless the modest gains in prediction can generate big cost reductions or profit opportunities.

That said, we still think the prospects for deep learning are bright in some many new contexts. As we have noted, it has significant advantages over traditional regression for analyzing "rich" databases that include images and non-numerical data. Rich data could include user-generated content (such as Amazon reviews, Instagram posts, Facebook posts, and comments on company websites), and the potential value of such data could be significant. Although the NerdWallet sites contained highly granular search and product data, it didn't rise to the level of rich data, so the advantages didn't come into play. Having consumer verbal reviews of credit cards may have improved the predictive accuracy of choice.

Deep learning is particularly well suited to identifying patterns in heterogeneous and unstructured data. Recently, for example, some of us developed a deep learning-based model to help automotive designers predict how consumers would react to prototypes on the drawing board while also generating ideas for new products.[3] In another case, researchers created a model from customers' news consumption behavior to predict future viewing habits.[4] Lately, companies have begun to use deep learning to analyze pictures of branded apparel that consumers post on social media and anticipate which items will have the highest customer return rates and which coupons will generate the most profit.[5] This allows managers to develop better product-return policies and more effective discount strategies. A prior study by some of us showed that deep learning methods helped companies pre-screen user-generated content more efficiently, thereby focusing analysts' efforts on highly-informative reviews that identified customer needs at substantially lower cost than the traditional methods. [6]

New opportunities for deep learning applications are emerging all the time. Deep learning encourages experimentation and enables real-time adaptability of A/B experiments and tweaking as results are observed. We already have developed a simple AI approach that adapts the images and content of online ads to match a consumer's cognitive style and designs real time effective experimentation systems to adaptively learn the best ad copy for each group of customers. Existing statistical models let companies target ads to online consumers. But the AI system the system allows marketers to tailor ad copy to individuals' cognitive and communication styles, customize pictures and graphics to suit their visual preferences, and incorporate stats and other evidence to align with their decision-making styles. The applications are showing promising results and are likely to deliver even better results as new deep learning-based algorithms and more comprehensive data bases become available.[7]

Implications for Companies

What will all this mean for companies? Based on our experience, we see several takeaways for marketing managers:

1. **Be alert to the future opportunities.** Established, statistics-based systems will still play a critical role in marketing analytics, but it's not too early to imagine how deep learning can enhance current operations and enable solutions to new problems. High potential areas include more sophisticated promotion budgeting and scheduling, more advanced targeting of customers, and more refined new product development.
2. **Invest in rich data bases**. Integrating many sources and types of data will likely produce major advances in response analysis and the optimization of marketing resources. The biggest costs of deep learning are apt to involve data acquisition and database creation and integrity. In our study of credit card selection, more than 50% of the costs were associated with data collection, labeling, and cleaning. So managers should invest in software to gather consumer comments from their web site, Amazon, Twitter and Instagram along with individual click stream records of search and purchase decisions by customers.

3. **Build capability through training and new hires**. To take advantage of opportunities in deep learning, you will need in-house experts capable of working with your data bases. They will need to know how to maximize the use of rich data, employ it to solve previously unsolved marketing problems, and find new strategic insights. Processing rich data requires advanced knowledge in AI and budgets to hire consultants to augment existing internal staff.
4. **Develop plans for experimentation**. Beyond using deep learning to conduct better and more efficient A/B experiments, you can also use it to test new marketing variables. The greatest potential may lie in focused experiments that systematically employ different variables in order to measure the response. The state of the art would be to use different real-time stimuli for each user. The experiments could then by analyzed using deep learning methods.

Thoughts About the Future

Despite the costs and challenges of implementation, we are optimistic about deep learning's potential to advance state-of-the-art of marketing practice by enabling analysis of rich data bases and real-time experimentation. This application in marketing analytics can: 1.) improve estimates of market response to help maximize profit and marketing ROI; 2.) reveal new opportunities for product development; and 3.) allow for more targeted product design, distribution, promotion, and media optimization. Although only marginal gains are available with old data and traditional analytics, in many new analytic applications, the improvements in accuracy and insight will justify the investment in deep learning technology and data.

Acknowledgements

We would like to thank Suruga Bank and MIT's Initiative on the Digital Economy for their funding support, and NerdWallet and Comscore for providing the data used in the research.

Sidebar: How Deep Learning Can Outperform Traditional Marketing Analytics

Estimating sales response to marketing inputs can be a relatively straight-forward analytics exercise. Many organizations use a variety of classic statistical methods and tools to simulate the effects of things like price changes and shifts in advertising, promotions, and distribution. It's common for brand managers to track their progress with simple models and online dashboards. In some cases, they use A/B testing: If you have a statistical sense of how customers are likely to react to a marketing offer, you can look for ways to adjust the variables to improve the outcomes.

Deep learning offers distinct advantages over traditional statistical modeling for predicting customer attitude changes and shopping and purchasing behavior, which can lead to improved profit opportunities.

- The first advantage is that it enables companies to look at more interactions and non-linearities, thereby allowing for a more comprehensive view of the customer. Instead of using one equation to model response, deep learning models use multiple layers of interlinked real and latent variables. In a deep learning model, for example, sales might flow from marketing promotions, but the responses could be directly tied to advertising copy and the social media impact of the brand through latent variables in hidden layers in a neural network. This network allows variables to interact in complex ways. Although this sometimes makes interpretation of effects more difficult, the interactions offer opportunities to improve predictive accuracy.
- A second advantage is tied to the technical criteria used for estimating success. Traditional statistical methods use the ability to describe historical relationships as the criteria for success. Deep learning, by contrast, uses the ability to predict consumer choices and responses through new data. This improves the model of consumer response and the probability that predictions will actually occur.
- A third advantage is the way deep learning can handle verbal and visual data simultaneously with numeric inputs. This is particularly valuable in today's Internet-infused environment, where available data can include user comments and other non-numerical content. Deep learning models can merge large databases such as Instagram and Twitter postings with more traditional data bases like sales information and marketing expenditures to examine broad questions like, "What are customers really looking for, and how much should we spend to reach them?"

[END OF SIDEBAR]

References

[1] See "Technical Note for Analysts" for details.

[2] See technical note for a description of state-of-the-art methods for monetizing gains in predictive accuracy.

[3] A. Burnap, J. R. Hauser, and A. Timoshenko, "Generating and Testing Product Aesthetics: A Human-Machine Hybrid Approach," (MIT Sloan School of Management, working paper series, MIT Cambridge, MA, 2019).

[4] P. Dhillon and S. Aral, "Modeling Dynamic User Interests: A Neural Network Approach," (MIT Sloan School of Management, working paper series, MIT Cambridge, MA, 2019).

[5] L. Liu, D. Dzyabura, and N. Mizik, "Visual Listening In: Extracting Brand Image Portrayed on Social Media," SSRN Working Paper, 2019.

[6] A. Timoshenko and J. R. Hauser), "Identifying Customer Needs from User-Generated Content," Marketing Science (2019), 38(1):1-20.

[7] G. Urban, J. Hauser, G. Liberali, M. Braun, and F. Sultan, "Morph the Web to Build Empathy, Trust and Sales," MIT Sloan Management Review (Summer 2009): 53-61.

Technical Note on Deep Learning

In this note we describe the differences between deep learning and traditional methods in more technical terms and outline how gains in accuracy can be monetized. We assume that the reader has rudimentary understanding of statistics.

Traditionally, consumer response has been modeled by single-equation models relating sales (outcome variable) linearly to the price/promotion, advertising and distribution measures (explanatory variables). The state-of-the-art today uses choice data (buy, not buy, how much) in a single equation (logistic) model, where outcome variables are choices made by individual consumers. Explanatory variables include a range of marketing measures such as exposure to ads on the Internet, search efforts, traditional demographics, loyalty, price, advertising and promotion spending, distribution, and product attributes. Data is used to find the model parameters that best reproduce past consumer choices. Statistics determine whether the results (and parameters) are "significant" (that is, unlikely to have occurred by chance).

Deep learning uses a different approach for estimating consumer response from these variables. Instead of relying on a single equation, it uses a neural network of latent variables and hidden layers. For example, if there are three layers in the network and if ten input variables are linked to ten latent variables at the first layer, there will be 100 parameters to estimate in the first layer. The next layer can be linked non-linearly to the next one; intermediate layers might have more or fewer latent variables. In fact, there might be hundreds of coefficients to estimate. In order to estimate these many coefficients, deep learning algorithms divide the data into three sections.

- The first section is used to "train" the model using algorithms that minimize the losses that measure how well the model fits the data. This is manipulating the data to describe well past choices. To avoid overfitting, analysts can configure models so they are less sensitive to noise.
- The second section of the data is used to determine the model's overall architecture (e.g. the number of hidden layers and the type of non-linear transformations between them).
- The final section is used to test the predictive accuracy.

Translating the gains in prediction accuracy into monetary gains is an area of active research. The standard approach is to examine the payoff from decisions with improved accuracy. This will vary according to the specific decision and company margins. A smally percent reduction in loan delinquencies may produce a high payoff for a credit card issuer. A similar improvement in brand choice may not generate more profit from better specification of advertising budgets.

Modeling based on deep learning is relatively new, and the algorithms that estimate the parameters are improving rapidly. For example, recently "dueling adversarial" models that compete with one another to generate the best DL parameter estimates have been used recently in state of the art applications of deep learning. New algorithms, together with data fusion and ever larger and more diverse/rich data, will enable improved response analysis.

For further information, see:

I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning (Cambridge, MA: MIT Press, 2015).

A. Geron, Hands on Machine Learning with Scikit-Learn and Tensor Flow (Sebastopol, CA: O'Reilly Media, 2017).

R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction (Cambridge, MA: MIT Press, 2018).