
Photo by Jezael Melgoza on Unsplash

Case Studies

Professor Yanzhi (David) Li

Unlock the Full Data Potential

Analytics is becoming increasingly important to business success. To create the right strategy, design the right product, and deliver the right service, all analytics need to be empowered by insights. After all, analytics is a tool that helps make more informed business decisions.

Through analytics, companies derive valuable insights into their customers, operations, and marketing. The people who play critical roles in this process are data scientists, and they typically receive their training in computer science - especially machine learning and artificial intelligence, as well as statistics and applied mathematics.

In some cases, technical ability plays a differentiating role, a matter of whether a problem can be solved or not, especially when the problem is large-scale or complex. Imagine an e-tailer who needs to deliver parcels as required by its customers in specified time windows. Failing that would mean unhappy and possibly lost customers. Solving the problem at the scale of hundreds of thousands of customers with hundreds of delivery staff in a city is challenging.

We have accumulated a substantial and comprehensive knowledge database about the behaviour of consumers and decision-makers, financial systems, and market dynamics. Such understanding should play a guiding role in exercising analytics and should be incorporated into the analytics process. In this way, the full potential of data can be unlocked. Neglecting this can lead to undesirable consequences.

Case Study One (1)

Optimizing Tmall Supermarket

Professor Li and his team worked with *Alibaba* on a recommendation problem for their Tmall Supermarket. Different from its traditional business, Tmall Supermarket runs a direct selling business model: *Alibaba* buys and owns the inventory and sells to its customers. To increase customer stickiness, Tmall Supermarket runs a channel called "今日疯抢" (Today's Best Deals), which is positioned to sell a selection of the most popular products from the tens of thousands of products on offer. A simple-minded approach for selecting the products would be to take those with the highest sales. A more sophisticated approach is to use the many established recommendation algorithms in machine learning, such as collaborative filtering.

Previously, Tmall Supermarket used a customized approach that combined some business rules, such as "A and B cannot be sold together," and machine learning algorithms to select products. Typical recommendation algorithms rank items based on their relevance and recommend the items with the highest relevance to customers. The team examined the problem and the current practice and identified some drawbacks. First, the objective of the problem was not clearly defined: Was it to maximize the clicks, the revenue, or the profit? Maximizing total relevance may not optimize any of these, though possibly positively related. Second, the previous approach ignored the mutual influence of products contained in the channel. For example, when two competing products were offered, their sales would be diminished by each other. Thirdly, customers' substitution behaviour, switching to another product when one's most preferred product is unavailable, has long been established as a trait in business research but was also ignored in the previous approach.

In view of these drawbacks, the team proposed a new approach that clearly reflected Tmall's objectives and captured the customer behaviour and cross-product influence. In particular, the team used the preference-ranking customer choice model, which is well researched in marketing literature. Marketing research shows that customers often follow a two-step process in their shopping process. In the face of overwhelming choices, they will first quickly scan the products, boil them down to a smaller group, called the consideration set, and then deliberate more carefully on this narrowed selection to decide what to buy. This customer choice model allows us to capture customer substitutions for any given selection of products. The analytics model we built, and the solution process are simplified due to incorporating consumer behaviour knowledge.

The outcome of our business insight-driven effort was impressive. The implementation of our approach on Tmall Supermarket brought significant improvements: a 7.4% improvement in the conversion rate (the proportion of customers who purchased among those who clicked the products) and a 16.9% improvement in the number of goods bought.

Case Study Two (2)

There Are Better Revenue Models Than Auction

Professor Li and his team worked with an advertising firm globally to propose an optimal allocation of ad resources on a global scale. Since the inception of the internet economy, advertising has been the most critical revenue model. Despite all the controversies about user privacy, the internet allows advertisers to target their customers effectively. There are potentially many interested advertisers for each search keyword or user impression, and a classical way of deciding the ad resource allocation is auction; simply speaking, whoever brings the highest payoff to the website wins. However, event-based auction has a few well-known drawbacks. For example, the first time you are exposed to an ad, suppose you are aware of it, there is a higher chance you may click it. If you have clicked it, the chance you will click it again is extremely low. This case applies to a targeted group as well. Therefore, the number of clicks you will receive from a targeted group for a particular ad first increases sharply, then increases slowly with the number of times you display the ad. In other words, the number of clicks is an S-curved function of the number of impressions.

In marketing theory, this diminishing marginal effect is termed "advertising wearout effect." Such dynamic effects are hard to capture by per-event based biddings, although the industry has proposed various measures to correct this; for example, dynamic updating of the click-through rate (CTR), which measures the proportions of individuals who see an online ad and subsequently click on it. Another major drawback associated with the auction mechanism is that it may induce fraudulent behavior. For example, advertisers may intentionally set very low prices at the beginning of the day and wait to win the bids until the later part of the day when other advertisers' budgets have been used up. Again, the advertising firm must introduce some smoothing-out scheme into the supposedly event-based auctions to prevent such behavior.

In view of the drawbacks of the current selling mechanism, the team proposed to plan the advertising delivery by a globally optimal allocation of ad resources. The team no longer relied simply on event-based auctions, which by nature are short-sighted and cannot produce a coordinated strategy. Instead, we made use of existing advertising theory, namely, the characteristic S-curved function of ad clicks.

In order to get the number of clicks, one may feel tempted to simply take the product of CTR and the number of impressions. This strategy, however, would be wrong since the CTR is changing dynamically. In other words, it is the insight that advertising will wear out over time that has inspired us to think about a planning approach, which further creates the need to predict the number of clicks for a given number of impressions. In effect, we have produced an analytics task that is new to the community.

The team developed a prediction model and, subsequently, an optimization model to allocate the ad resources. We tested our method at a worldwide leading advertising platform and found an improvement of over 10% in revenues, a considerable boost to profitability.

Publications

Mao, Huiqiang; **Li, Yanzhi**; Li, Chenliang; Chen, Di; Wang, Xiaoqing; Deng, Yuming. (2020). PARS: Peers-aware Recommender System. *The Web Conference 2020: Proceedings of The World Wide Web Conference WWW 2020*, 2606-2612.

Shen, Huaxiao; **Li, Yanzhi**; Guan, Jingjing; Tso, Geoffrey K.F. (2021). A Planning Approach to Revenue Management for Non-Guaranteed Targeted Display Advertising. *Production and Operations Management, 30*(6), 1583-1602.