

Letter from the Special Issue Editor

The global e-Commerce market size is valued at USD 9.09 trillion with an annual growth rate of 14.7 percent. The 2020 pandemic dramatically changed people's lifestyles. E-Commerce will further accelerate its growth and penetration into people's daily lives. E-Commerce websites and apps are among the top visits of everyone's daily routine. Customers want E-Commerce websites and apps as their personal assistant that finds the exact products they are searching for, provides recommendations when they are not sure which products to buy, and answers questions about product details. E-Commerce presents a diverse set of data mining challenges such as product attributes parsing, learning to rank product lists, product clustering, and classification, personalization.

This special issue selects six example works in data mining for E-Commerce from industry companies such as Walmart, Etsy, and Amazon and academia institutes. In recent years, Deep Learning has become the new norm for AI and Machine Learning. We certainly see the trend in the submissions that many old E-Commerce problems benefit from the advanced deep learning algorithms. We selected a few papers in this issue to showcase the applications of Deep Learning in E-Commerce applications. The first paper presents academic research to improve hierarchical product classification with transformer models. Transformer models such as BERT are state of the art for many NLP tasks. However, E-Commerce often has its unique vocabulary and domain-specific texts. Fine-tuning and self-supervised continuous pre-training with domain-specific data is the trend to apply the BERT style models in E-Commerce. The authors demonstrated the effectiveness of this approach in the Common Crawl data set. The second paper works on a similar problem, but it is from a real-world application by the Amazon team. The Amazon team demonstrated the effectiveness of domain-specific fine-tuning of BERT models and shared valuable tips and tricks to get it working in a real production such as negative sampling, soft label with temperature scaling, bootstrap learning, and leveraging in-domain knowledge augmentation. The third paper is also from an Amazon product knowledge team. It presents the application of graph neural networks in parsing and understanding product descriptions. E-Commerce data comes as a big heterogeneous graph of queries, customers, sellers, products, product entities (brand, franchise, etc.) and thus is an ideal place for graph neural networks to shine.

E-Commerce has the advantage over brick-and-mortar stores in their ability to personalize the experience based on user profile. The fourth paper is an example of such personalization work in airline ticket marketing. The interesting part is their approach that leverages knowledge graph embeddings rather than rule-based knowledge engineering to better target the right audience. The fifth paper presents the work in Etsy's personalized product ranking. As the buyer continues on their shopping mission and interacts with different products in an online shop, their model learns which attributes the buyer likes and dislikes, forming an interpretable user preference profile and improving re-ranking performance over time within the same session. E-Commerce should make a generational lift from a simple Information Retrieval engine to a more proactive mission-aware shopping assistant. Etsy's work demonstrated encouraging user experience improvements as a result of mission-aware personalization.

The last paper in this special issue is from Walmart. The authors re-visited an old problem in learning to rank for web search, the positional bias in the displayed result list. E-Commerce websites often have a more diverse UI layout, such as a grid view instead of a simple list view in web search. The author addressed the inefficiency of applying the classic position bias removal method to E-Commerce. The paper is interesting since it demonstrates some unique challenges in E-Commerce than the general web search.

Working on this issue has been a privilege for me, and we would like to thank the authors for their contributions.

Bing Yin and Sreyashi Nag
Amazon.com