Wenrui Li 23043574 uceewl4@ucl.ac.uk
Msc IMLS ELEC0136: Data Acquisition and Processing Systems 23/24

Mini case study 1: Enhancing Customer Insights in E-commerce (Data Sources and Strategies)

eBay is a typical e-commerce platform offering a vast variety of goods for citizens' life. As a data analyst for eBay, I'm going to identify data sources and related strategies from the following perspectives.

Internal Databases

Top e-commerce companies may have their specially-designed databases including information of products, users, orders, payments, discount, etc., which match various platform functionalities [1].

Transaction records includes different items bought by the user in specific time with certain quantities. The latent correlation and features of these items can be analyzed to identify customer's purchasing pattern and infer information like age, gender, habits and occupation. In this case, items always bought simultaneously could be recommended to customers jointly and customer information can be associated with searching history to provide more precise recommendation. For instance, if beer and baby diaper are often bought together by someone, then maybe he became a dad recently. Association rules like apriori algorithm can be used to calculate confidence and support of these items, and more related goods like cigarettes, milk powder can be recommended to him. This helps develop a better understanding of customer needs, probe into underlying purchasing behaviors and enhance overall sales by associated recommendation.

transaction id	items bought
10	Beer, Nuts, Diaper
20	Beer, Coffee, Diaper
30	Beer, Diaper, Eggs
40	Nuts, Eggs, Milk
50	Nuts, Coffee, Diaper, Eggs, Milk
support(Beer -> Diaper) = 60%; confidence(Beer -> Diaper) = 100%	

Table 1: Association rules with support and confidence.

Ranking information like *searching history ranks* and *best-sellers* is an important metric needed to be utilized by data analysts [2]. Searching history ranks indicates concentration and interests for certain goods, and best-sellers shows favorites from a majority of customers. These information can be presented to customers to arouse attention and boost purchasing power. For example, data analysts may use RANK OVER() and some aggregation functions (SUM(), COUNT(), etc.) in SQL to obtain ranks for certain categories and recommend the best-sellers/top searching ones to customers. Combined with the filtering and sorting systems in eBay, this reduces customers' time for making comparisons and enhance customer loyalty with convenience.

Other internal databases can also provide comprehensive insights in decision-making

process. *Inventory and sales data* in previous seasons can be used to regulate storage, supplying and demands. For example, average quantities of fresh fruits and perishable items sold in different quarters can be used to make future estimation, which can reduce loss from large-scale waste. This can also be cooperated with information from related suppliers and vendors. Moreover, in "My eBay" section, there is s*hopping cart records* and *browsing history*, which can be analyzed for underlying purchasing behaviors. For example, some e-commerce platforms send emails like *"Do you still want to buy...."* if the items were saved in customer's cart for a long time, which improves customer loyalty as well as boost purchasing power. Besides, *payment information* deserves special attention from data analysts for identifying possible fraudulent transactions and online crisis. For example, we can use anomaly detection algorithms like LOF, DBSCAN, etc., to expose abnormal payment behaviors, which can improve platform security and reduce customers/profit loss.
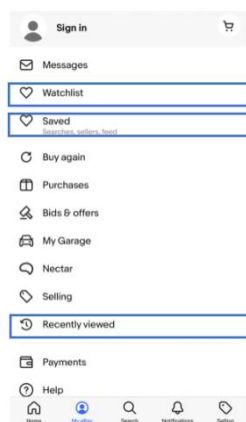

Figure 1: My eBay section in eBay application platform.

Users

Reviews and customer feedback are important data sources for improving understanding of customer needs. eBay provides seller score for customers to make comparison. If a seller has a score of 99.5%, it means that 99.5% of the buyers who left reviews for that seller had a positive experience [3]. In this case, data analysts can simply use statistical analysis for seller scores or apply sentiment analysis model (like triple classification with positive, negative, neutral emotions) on customers' reviews to identify seller credits. Besides, customer feedback from periodic questionnaires and complaints records are also valuable. As a user of eBay, the platform occasionally sends me emails like *"For you: Finds that deserve another look"* for recommendation and emails requesting for leaving a feedback on items. Through numeric scoring and text analysis, eBay can therefore grasp knowledge in customers' requirements, build more suggestive recommended systems and enhance customer loyalty.
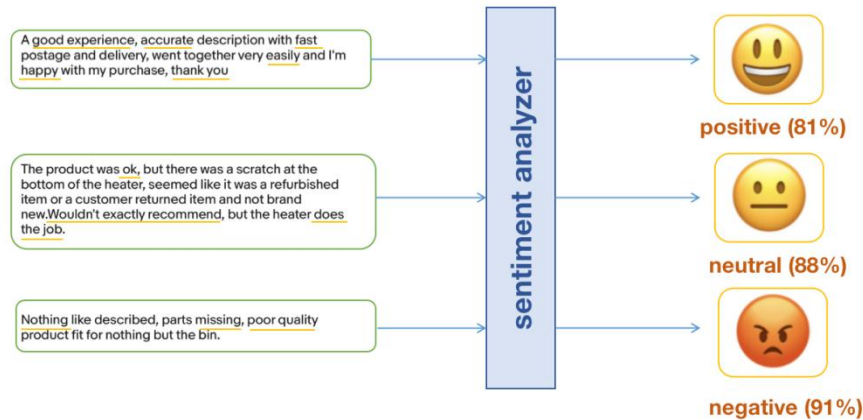
Wenrui Li 23043574 uceewl4@ucl.ac.uk
Msc IMLS ELEC0136: Data Acquisition and Processing Systems 23/24

Figure 2: Sentiment analysis model for customer reviews.

Third Party

Many third party data sources can provide constructive insights as well. For some cases, various third party platforms reached agreement on sharing specific information from internal databases for platform improvement. Firstly, *location information* in your phone background or from navigation applications like Google Maps can be used to identify purchasing behaviors of certain regions. For example, if there is a high-volume transactions for rain gear (umbrella, raincoats, etc.) of customers from the same region, the information can be associated with weather forecasting systems for better recommendation and prediction. When the next rain season is coming, related rain gear will automatically be recommended to your main page. This can also be applied for certain demographics. For example, supplements, walking sticks and medical devices can be recommended to customers living in serious-aging region. Moreover, eBay can cooperate with social media like Instagram, Twitter, etc., to learn and recommend items with customer interests. For example, stationary and digital devices can be recommended to customers who are recently skimming educational courses on Youtube. Besides, data information like clicking rate of advertisement cast by eBay on other media platforms can be used to improve channel utilization. More advertisements could be put on platforms with high amount of user clicks.

Measures above can increase overall sales to some extent. However, data sources from third party may not always be available from web scrapping. More importantly, mutual contract should be reached by both parties in a legitimate format. The acquisition of third party information should highly obey regional law and regulations, which leads to some practical restrictions.

References

[1] 5 Data Sources That Ecommerce Companies Should Excavate for Their AI Efforts.
https://dataconomy.com/2020/05/14/5-data-sources-that-ecommerce-companies-should-excavate-for-their-ai-efforts/
[2] 7 eCommerce Data Sources You Must Scrape.
https://www.blog.datahut.co/post/ecommerce-data-sources-to-scrape

[3] eBay Customer Service.

https://www.ebay.co.uk/help/buying/resolving-issues-sellers/seller-ratings?id=4023