Proceedings of the 4th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/51/20241165

Research on the application of machine learning in business analytics: Cases of Amazon and eBay

Rongrong Zhang

School of Digital Economics and Management, Wenzheng College of Soochow University, Suzhou, China

lelem74038@student.napavalley.edu

Abstract. With the rapid development of the Internet and the rise of e-commerce, commercial enterprises are faced with a large amount of data and a complex market environment. In this situation, machine learning, as a powerful tool, is widely used in the field of business analysis. In this dissertation, we take Amazon and eBay as examples to study the application of machine learning in the company's business analytics, focusing on its role in market prediction, customer behavior analysis and operation optimization. By analyzing the relevant cases, we find that machine learning technology plays an important role in helping companies make more accurate decisions and improve efficiency. Studying the application of Amazon machine learning in business analytics can promote in-depth research on the application of machine learning in business in academia, and promote the application and development of machine learning technology in other business scenarios. Overall, the application of machine learning in business analytics can help companies understand customer behavior, optimize operations, and improve sales results. However, there are still some challenges, such as data quality, algorithm selection and privacy protection. Therefore, further research and innovation are necessary to advance the development of machine learning applications in business analytics.

Keywords: Machine Learning, Business Analytics, Amazon, Data Mining, Supply Chain Management.

1. Introduction

With the explosive growth of data, organizations need to utilize data more effectively to guide strategic decisions. Machine learning, as a powerful tool, can help enterprises extract valuable information from massive data and perform predictive analysis. Amazon, as one of the world's largest e-commerce companies, has a huge inventory of goods, massive user data and a complex competitive market environment. How to improve operational efficiency and market competitiveness by utilizing these data for business analysis has become the focus of Amazon's attention. Machine learning, as a data-driven analysis method, has a powerful ability to identify patterns, predict trends and optimize decisions.

Regarding research on e-commerce enterprises, the existing literature [1] Ma et al provides a comprehensive overview of machine learning applications in e-commerce, including recommender systems, fraud detection, demand forecasting, and advertisement placement, etc. Li et al systematically review the applications of machine learning in e-commerce research [2]. The authors organize the main trends, methods and application areas of related research [3]. Zhang et al. analyzed the research trends,

DOI: 10.54254/2755-2721/51/20241165

hotspots and author collaboration networks of machine learning applications in e-commerce in recent years through bibliometric methods. Sun et al. explored the application of machine learning in e-commerce enterprises, especially in recommendation systems, personalized marketing and user behavior prediction, using Amazon as a case study [4]. Research on machine learning, therefore, this paper will focus on two e-commerce giant enterprises, Amazon and eBay, based on the existing literature, to study their applications of machine learning in business analytics.

2. Analysis of machine learning application areas in business analytics

2.1. Amazon

Background: Amazon is one of the largest e-commerce and online retail platforms in the world, which has a large number of products and users. In order to enhance user experience and increase sales, Amazon has developed a personalized product recommendation engine as shown in Table 1.

Table 1. Areas of application and Applied value.

Areas of application	Applied value
Personalized Recommendation System	Data collection and processing, data mining and model building: building personalized recommendation models
Sales forecasting and demand forecasting	Optimize inventory management, purchasing plans, and pricing strategies to discover underlying market trends and consumer behavior patterns.
Fraud detection and risk control	Protect the rights and interests of consumers and sellers, identify and prevent fraud in a timely manner, and safeguard the security of users and businesses.
Supply Chain Optimization	Optimize supply chain management to optimize distribution routes and reduce transportation costs.
Marketing and advertising optimization	Provide advertisers with targeted ad placements to increase ad click-through and conversion rates.
Operations Optimization	Help companies optimize operational processes, reduce costs and improve efficiency.

2.1.1. Personalized Recommendation System

Data collection and processing: Amazon uses machine learning algorithms to analyze users' historical purchase data, clicking behavior, browsing records and other information to provide personalized product recommendations for each user, establishing a detailed user profile and product association information. Amazon's personalized recommendation system generates millions of recommendations for hundreds of millions of users every day to improve the user buying experience and sales. The business analysis team optimizes and adjusts the recommendation algorithm by regularly analyzing data such as user feedback, click-through rate, and purchase rate to improve the accuracy and effectiveness of the recommendations.

Data mining and model building: Based on the collected data, Amazon applies data mining techniques, such as collaborative filtering and machine learning algorithms, to build personalized recommendation models. These models are able to predict the products that users may be interested in and generate customized product recommendations for each user.

2.1.2. Sales Forecasting and Demand Forecasting

By analyzing historical sales data, market trends, promotions and other information through machine learning algorithms, Amazon is able to predict the sales volume and demand trends of different products, so as to optimize inventory management, purchasing plans and pricing strategies. Machine learning technology is used to mine and analyze large amounts of sales data to discover potential market trends and consumer behavior patterns. By analyzing users' shopping preferences and purchase history,

DOI: 10.54254/2755-2721/51/20241165

Amazon is able to better understand market demand and adjust its merchandise strategy and pricing strategy.

2.1.3. Fraud Detection and Risk Control

Amazon uses machine learning technology to detect potential fraud and abnormal transactions to protect consumers and sellers. The machine learning model can analyze user behavior patterns, transaction data, reviews and other multi-dimensional information to identify and prevent fraud in a timely manner to protect the safety of users and businesses.

2.1.4. Supply chain optimization

Amazon uses machine learning to optimize supply chain management, including inventory management, order delivery and logistics planning. By analyzing large amounts of supply chain data and market demand, Amazon is able to provide more accurate delivery date forecasts, optimize delivery routes and reduce transportation costs.

2.1.5. Marketing and Advertising Optimization

Amazon uses machine learning to analyze user behavior and purchase preferences to provide advertisers with targeted advertising. The machine learning model can optimize ad display and recommendation based on users' interests and behavioral characteristics, increasing the ad click rate and conversion rate.

2.1.6. Operation Optimization

Machine learning can help companies optimize operational processes, reduce costs and improve efficiency. Machine learning is used to optimize delivery routes, thereby reducing fuel consumption and transportation time.

Value and results:

Enhanced user experience: a personalized recommendation engine makes it easier for users to discover products that match their interests and needs, increasing user satisfaction.

Increased sales: with accurate recommendations, users are more likely to purchase additional products, which increases Amazon's sales.

Data-driven decision-making: By analyzing the recommendation engine, the business analytics team is able to understand user behavioral patterns and trends, which guides the development of product and marketing strategies.

Continuous Improvement: Amazon continuously optimizes its recommendation algorithms and improves its recommendation system by analyzing data and incorporating user feedback to maintain its competitive edge in the highly competitive e-commerce market.

This case demonstrates how business analytics can help companies improve user experience, increase sales, and gain an edge in the competitive marketplace through data analysis, modeling, and optimization

2.2. eBay

eBay is one of the world's largest online shopping and auction sites, it also widely used machine learning techniques for business analysis and performance improvement, specifically used in risk management, search management, sales forecasting and other areas, as shown in Table 2.

Table2. Areas of application and Applied value.

Areas of application	applied value
risk management	Identify potential fraud and security threats
Search and Recommendation	Optimize search results and recommendation lists based on users' interests and behavioral characteristics to increase user satisfaction and sales.
Customer service and support	Better handling of user issues and feedback
Commodity pricing and sales forecasting	Forecasting sales volume and price trends for different commodities

DOI: 10.54254/2755-2721/51/20241165

2.2.1. Risk Management

eBay uses machine learning algorithms to analyze transaction data, user behavior, and history to identify potential fraud and security threats. eBay's machine learning models can monitor abnormal behavior during transactions and stop potential fraudulent activity in a timely manner.

2.2.2. Search and Recommendation

eBay uses machine learning algorithms to analyze search queries and product browsing records to provide users with personalized search results and recommended products. eBay's machine learning models can optimize search results and recommended lists based on users' interests and behavioral characteristics to increase user satisfaction and sales.

2.2.3. Customer Service and Support

eBay utilizes machine learning technology to build an automated customer service and support system to better handle user questions and feedback. eBay's automated system can match the best answers and provide solutions based on the questions and keywords entered by users.

2.2.4. Merchandise Pricing and Sales Forecasting

eBay uses machine learning technology to analyze historical sales data, market trends, and competition in order to predict the sales volume and price trends of different items. In this way, merchants can develop more effective pricing strategies and sales plans based on the forecast results to increase profits and market share.

3. Examples of machine learning applied to specific business scenarios

3.1. Scanning shopping sites for counterfeit or stolen goods

Machine learning can be used to search for stolen or counterfeit goods with customer branding on shopping websites. In order to search for counterfeit or stolen goods in the millions of product listings on a shopping site, a crawler algorithm is used to crawl the web pages containing relevant keywords and store the web page images in a specific folder. The images need to be crawled and viewed on a daily basis by logging in to the relevant portal to determine if they are infringing on copyrights. This is a very expensive and boring process, expensive in terms of the labor cost to view the images, and boring in terms of the time consuming, but not very technical, task of viewing the images. But this problem can be greatly improved with a little machine learning [5].

Deciding which machine learning algorithm to use to solve a problem is often complex. On the one hand, there is a large amount of data labeled as infringing or non-infringing, in which case it is easy to see that this is a supervised classification problem. So, consider trying a neural network algorithm? On the other hand, retrieving negative information is very time-consuming due to the design of the portal (it does not store an image, just a link), so there is only a limited dataset to acquire. In addition, corporate logos often appear in images alongside logos of other companies (copyright infringement is often subtle), and a deep enough neural network will likely recognize the presence of similar trademarks rather than the logo of a particular company alone (just as a neural network typically recognizes all breeds of dogs, not just German Shepherds). In this case, the key is to train a shallower convolutional neural network to perform target detection rather than classification [6]. This prevents the network from learning logo-general features and forces it to learn logo-specific features.

The CNN (Convolutional Neural Network) algorithm itself is mature, then the key to this project is to generate a training set for training target detection. It needs to consist of a large number of images from the original training set with different angles and alignments, with various tilts and alignments, each with a wraparound frame around the object we want to detect (the logo). Thereby, scanning for counterfeit goods or stolen goods using machine learning can be achieved.

DOI: 10.54254/2755-2721/51/20241165

3.2. Predicting whether a retail store will be burglarized

For multinational retail chains, branch stores are located all over the world, including some less secure areas. It is therefore necessary to construct an algorithm that generates a risk score for burglary based on historical instances of burglary.

Due to the nature of the business, the risk of burglary increases with crime in the surrounding area, rather than due to any particular feature of the store (fragile windows, etc.). The purpose of the algorithmic tool is not to directly prevent burglary, but to assess how a fixed budget can be allocated across the entire portfolio of stores to reduce burglary losses. Therefore, the algorithmic output must be the benefit of updating each security feature, or recommending which ones to update, a bit like a recommendation engine [7].

Organizations can address this issue through machine learning. First, build separate classifiers (using XGBoost) for each factor the customer wants to identify, setting a fixed training set date range for each period to address the fact that predictions for longer periods are more likely to overestimate long-term risk. Next, these factors are imported into a regression model to calculate the predicted loss for each retail store, and that metric (combined with a geographic metric) is used to generate specific recommendations. From there, the use of machine learning to predict whether a retail store will be burglarized can be achieved.

3.3. Recommending Content to 150 Million Visitors

Infinite Scroll, generally known as the infinite scroll mode, can be used for pages with many similar entries to be displayed with an infinite dropdown to avoid users from getting more content by clicking on the next page. For example, on the Today's Headlines website, when you open it, you can scroll all the way down, and when you scroll to the bottom of the page, it will load more news items. However, there is confusion about which article to display next after pulling down to the bottom, and machine learning can be a good solution to this problem, using word embedding and user embedding to build a collaborative filtering recommendation engine that provides users with articles that match their interests, not just the most popular ones. One of the biggest challenges is not so much the question of which algorithm to choose, but the handling of massive amounts of data: one must ensure that the system returns recommendations very quickly - the essence of the infinite scroll experience.

Another challenge was data burial and backup, as there are very many features that can be extracted for each user behavior, so a way needed to be found to store this information without exponentially increasing storage costs. The analytics team decided to split the model by site, category and user information to reduce the size of the embedding matrix. Various downscaling techniques were also deployed to make it more manageable, and the system was closely monitored to determine the ideal cutoff date for backups [8]. It can be noticed that very often, algorithm selection and parameter tuning are not the most important phases of a machine learning project in an organization. Instead, customer requirements, data types, computational speed, predictive usage, and even database optimization can affect the success of a project.

4. Discussion

Amazon and eBay, as e-commerce giants, have applied machine learning techniques in business analytics. However, there are some research limitations associated with these applications:

Data privacy and security: Amazon and eBay process large amounts of user data for business analytics, but this also poses risks to data privacy and security. Due to the sensitive information of individual users involved, researchers may not have access to complete datasets for their studies or may only be able to use data that has been anonymized. Thus, research on machine learning applications for such companies is limited by data access and privacy protection.

Trade Secrets and Competitive Advantage: Amazon and eBay, as commercial companies, their practices in machine learning applications are usually trade secrets and are not shared publicly. This makes it difficult for researchers to obtain detailed use cases and algorithmic details, thus limiting the depth and feasibility of the research.

DOI: 10.54254/2755-2721/51/20241165

Lack of public datasets and standardized benchmarks: To study Amazon and eBay's machine learning applications in depth, a large amount of data support is required. However, commercial data from these companies is often private and not publicly available. In addition, the lack of recognized standard datasets and benchmarks to evaluate and compare the performance of machine learning algorithms in business analytics limits the progress of related research [10].

Dependence on company-specific environments and infrastructures: machine learning applications at Amazon and eBay typically rely on the company's own environment and infrastructure. These applications may be tightly coupled with the organization's internal systems and data architecture, making it difficult to directly migrate to other business or research environments. As a result, there are challenges in generalizing these machine learning applications to other business scenarios.

Despite these limitations, it is still important to study Amazon and eBay machine learning applications for business analytics. By overcoming data access limitations, establishing partnerships, using proxy data, and exploring publicly available datasets, these limitations can be partially addressed and further research and development of machine learning in business analytics can be promoted.

5. Conclusion

The use of machine learning in business analytics is becoming indispensable. By deeply analyzing data, predicting trends, and optimizing decisions, businesses can achieve greater success.

5.1. Improving Business Efficiency

Enterprises use machine learning technology to analyze massive amounts of data, including sales data, user behavior data, and supply chain data, to help optimize business processes and improve operational efficiency. Studying Amazon, eBay's application of machine learning can reveal cases of how it utilizes data analytics tools and algorithms to achieve efficient business operations, thus providing lessons and inspiration for other enterprises [9].

5.2. Improve user experience

Machine learning techniques are used to personalize product recommendations, optimize search results and improve customer service to provide a better user experience. Studying the application of machine learning in Amazon, ebay can help understand how to use machine learning algorithms to understand user needs, predict user behavior and optimize products and services accordingly.

5.3. Data-driven decision making

Research on machine learning applications can provide insights into how a data-driven mindset can be applied to business decisions. Amazon uses machine learning algorithms to support the decision-making process based on data to improve the accuracy and effectiveness of decisions. Studying Amazon's machine learning applications can help us understand how data analytics and machine learning methods can be integrated with the decision-making process to optimize business decisions.

5.4. Advancing the use of machine learning in business

Amazon, eBay as a large e-commerce company, has accumulated rich experience and practice in the research and application of machine learning. Studying the application of Amazon machine learning in business analysis can promote the academic community to conduct in-depth research on the application of machine learning in business and promote the application and development of machine learning technology in other business scenarios.

References

[1] Ma, X., Chen, H., Wang, H., Li, T., & Liu, C. (2020). Machine learning for e-commerce: A comprehensive survey. *Information Fusion,* 58, 1-25.

[2] Li, S., et al. (2019). A systematic review of machine learning in electronic commerce research.

DOI: 10.54254/2755-2721/51/20241165

[3] Zhang, Y., Ma, T., Duan, Y., & Shi, Y. (2021). Machine Learning Applications in E-commerce: A Bibliometric Analysis. *Computers in Human Behavior*, 124, 106963.
[4] Sun, Z., et al. (2017). Machine Learning Application in E-commerce: A Study of Amazon.
[5] Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78-87.
[6] Chen, X., & Lin, X. (2014). Big data deep learning: challenges and perspectives. *IEEE Access*, 2, 514-525.
[7] Varian, H. R. (2014). Big data: new tricks for econometrics. *The Journal of Economic Perspectives,* 28(2), 3-28.
[8] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. *MIT press*.
[9] Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big data*, 1(1), 51-59.
[10] Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42(8), 30-37.