Chapter 6
Strategic Integration of Machine Learning for Fraud Detection in E-Commerce Transactions

P. Vijayalakshmi
Knowledge Institute of Technology, Salem, India

Anand Anbalagan
Technical Vocational Training Institute, Addis Ababa, Ethiopia

K. Subashini
Tagore Engineering College, Chennai, India

N. Bharathiraja
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

B. Selvalakshmi
Tagore Engineering College, Chennai, India

Gaganpreet Kaur
https://orcid.org/0000-0002-3322-1315
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

G. Sudhakar
Sri Sai Ranganathan Engineering College, Coimbatore, India

ABSTRACT

The rise in internet users has led to an increase in online payments, but this also comes with a surge in online fraud. To combat this, e-commerce firms must adopt device intelligence for fraud detection. Machine learning (ML) is crucial for analyzing large datasets to identify suspicious patterns. This study explores the effective application of ML in detecting fraudulent activities, focusing on various approaches, challenges, and recommendations. It starts with an overview of the prevalence and impact of e-commerce fraud, highlighting the need for robust detection systems. Key

DOI: 10.4018/979-8-3693-5718-7.ch006

ML techniques, including supervised, unsupervised, and semi-supervised learning, are analyzed for their strengths and weaknesses. It emphasizes the importance of continuous monitoring and model adaptation to evolving fraud tactics, advocating for dynamic updates and feedback loops to enhance detection systems. By integrating ML algorithms effectively, e-commerce businesses can improve security, safeguard revenues, and build trust with consumers and partners.

1. INTRODUCTION

Sales of retail e-commerce are still rising fast (Aburbeian & Fernández-Veiga, 2024). A sizable e-commerce website processes daily queries from millions of visitors. Regrettably, while legal traffic from consumers has increased, e-commerce scams have also increased, endangering the industry's finances and reputation (Vinod et. al., 2023). Over 280,000 reports were filed with the Internet Crime Complaint Centre (IC3) in 2015, which ultimately resulted in over $1 billion in losses (Nama & Obaid, 2024). On e-commerce platforms, hijacking accounts and credit imitation are two prevalent methods of fraud. Fraudsters can create new accounts on the internet using stolen or counterfeit credit cards, or the customer can take advantage of an individual's account balances by stealing them from the person. In either scenario, the internet page and its visitors suffer losses. To prevent such behavior, it is imperative to develop efficient systems for identifying fraud (Garg & Sharma, 2024).

All of these algorithms rely on consolidated characteristics, such as the total number of products that someone has looked at in the past month, but personal behaviors are sometimes the only way to identify fraud. Furthermore, basic features or regulations soon become outdated as fraudulent behaviors evolve to evade detection. Therefore, feature extraction—the process of capturing user behavior in as much detail as possible—and algorithm selection—the process of identifying frauds from a large volume of data—are critical components of any fraud detection system (Shukla et. al., 2024).

The order in which a user clicks throughout a session, or their surfing behavior, is one of the most crucial pieces of data for identifying fraudulent activity. According to statistics, forged users' actions differ from those of genuine users. Actual users have a consistent way of browsing. The users will probably go through several products that are comparable to the one the user purchased for study (Arulkumar et. al., 2023). On the other hand, fraudulent individuals exhibit more consistent behavior, such as going straight to the virtual things the user wish to purchase or acting haphazardly. Although these devices are Apple, the customers are unrelated because these are tablets, phones, and PCs. As a result, it's critical to record every