IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, January 2026
Impact Factor: 7.67

E-Commerce User Activity Log Analysis Using Hadoop Data Lake

Bharath Guru[1], Chinmay L[2], Darshan A M[3], Channabasava[4], Mrs. Kavitha M. G.[5]
B E, CSE, Kalpataru Institute of Technology, Tiptur, India[1-4]
Assistant Professor, CSE, Kalpataru Institute of Technology, Tiptur, India[5]

Abstract: *The rapid growth of e-commerce platforms has resulted in the generation of massive volumes of user activity data such as clicks, searches, views, and purchases. Analyzing this large-scale data using traditional database systems is inefficient due to limitations in scalability, storage, and processing speed. Big Data technologies provide a promising solution to overcome these challenges. This paper presents an end-to-end system for analyzing e-commerce user activity logs using a Hadoop Data Lake architecture. User interaction data is collected from a web-based e-commerce application, stored in Hadoop Distributed File System (HDFS), and analyzed using Apache Hive. The proposed system enables efficient storage, querying, and analysis of large datasets to extract meaningful insights into user behavior, popular products, and engagement patterns. Experimental results demonstrate that the Hadoop–Hive-based approach is scalable, cost-effective, and suitable for batch analytics in e-commerce environments.*

Keywords: E-commerce, Big Data, Hadoop, HDFS, Hive, Data Lake, User Activity Logs

DOI: 10.48175/IJARSCT-30771