Big Data Analytics in E-Commerce

P Manikanta, K Sri Trinesh, Sri Ram Chowdary, M Praneetha, Dr G Dinesh Kumar

Dept of Computer science and engineering, Koneru lakshmaiah Education Foundation, Vijayawada, India
Dept of Computer science and engineering, Koneru lakshmaiah Education Foundation, Vijayawada, India
Dept of Computer science and engineering, Koneru lakshmaiah Education Foundation, Vijayawada, India
Dept of Computer science and engineering, Koneru lakshmaiah Education Foundation, Vijayawada, India
Dept of Computer science and engineering, Koneru lakshmaiah Education Foundation, Vijayawada, India

ABSTRACT
As the world becomes more digital, e-commerce continues to grow at an unprecedented rate. With the amount of data generated by online shopping, it's no surprise that businesses are turning to big data analytics to drive growth and personalization. By analyzing consumer behaviour, purchasing patterns, and other data points, e-commerce companies can make informed decisions that improve the customer experience and increase sales. In this paper, we'll explore how big data analytics can help businesses in the e-commerce industry, and the benefits of leveraging this technology for growth and personalization. From understanding consumer behaviour to predicting trends, we'll provide insights that can help your e-commerce business stay ahead of the competition.
KEYWORDS: Big data analytics; E-commerce; Data driven Marketing; predictive analytics; Digital Marketing; 5 Vs in Big Data;Apriori Algorithm.

I.   INTRODUCTION

In the contemporary digital era, the e-commerce arena undergoes perpetual transformation, as businesses ardently seek to secure a competitive advantage and foster expansion. Among the most potent instruments within their reach, big data analytics reigns supreme. Big data refers to the enormous volumes of data generated through various online activities, such as website visits, customer interactions, purchase history, and social media engagement.

The role of big data analytics in e-commerce cannot be overstated. It provides businesses with valuable insights into consumer behaviour, preferences, and trends. By harnessing the power of this data, businesses can make informed decisions to optimize their operations, improve customer experiences, and drive revenue growth.

Big data analytics empowers e-commerce enterprises to attain a profound comprehension of their clientele. Through the meticulous analysis of extensive data sets, businesses can pinpoint patterns, trends, and correlations, enabling them to customize their offerings to align with customer requirements and preferences. This elevated degree of personalization significantly enriches the overall customer experience, culminating in heightened customer satisfaction, unwavering loyalty, and, ultimately, augmented conversion rates.

Moreover, big data analytics enables businesses to identify and target specific market segments effectively. Businesses can develop highly focused marketing campaigns that resonate with their target audience by looking at client demographics, purchase history, and browsing behaviour. This not only improves marketing effectiveness but also maximizes return on investment.

Furthermore, big data analytics can help businesses optimize their supply chain and inventory management. By analyzing historical sales data, customer demand patterns, and market trends, businesses can accurately forecast demand, manage inventory levels, and streamline their operations. This leads to reduced costs, minimized stockouts, and improved customer satisfaction through timely and efficient order fulfilment.

Big data analytics are essential to the e-commerce sector because they give companies insightful data that helps them grow and personalise their offerings. It empowers businesses to make data-driven decisions, optimize operations,

improve customer experiences, and ultimately, achieve a competitive advantage in the dynamic e-commerce landscape. Through wholeheartedly embracing the potential of big data analytics, enterprises can unlock vast opportunities for triumph in the digital sphere.

II. LITERATURE SURVEY
Big Data Analytics: Impacting Business in Big Way:

The paper delves into the evolving landscape of Big Data analytics, emphasizing its increasing relevance in understanding customer behaviour, purchasing patterns, and social media sentiment analysis. In today's business environment, data is generated abundantly through clickstreams on authentic websites, social media interactions, mobile device location information, and machine-generated data. The central components driving this transformation include Intelligent Connected Machines equipped with Internet connectivity and advanced sensors for data capture, automation controls, and software applications. This marks the dawn of a third revolution, building upon the industrial and Internet revolutions, now fueled by the potential of Big Data. Consequently, Big Data systems emerge as the next frontier for businesses.

Within this context, the paper presents three case studies tailored to the cell phone industry, e-commerce, and online insurance selling. These case studies underscore the critical role of Big Data in these domains and introduce the concept of a Big Data Analyzing Engine designed to identify, collect, store, and analyze large datasets to drive business success.

The introduction of the paper sets the stage by defining Big Data as a rapidly growing torrent of information that represents the next generation of data warehousing and business analytics. Coined by Roger Magoulas in 2005, Big Data is characterized as data that surpasses the capacity of conventional processing methods. The McKinsey study further elaborates, categorizing Big Data as datasets exceeding the capabilities of typical database software tools in terms of capture, storage, management, and analysis. It distinguishes Big Data from small data by its substantial volume, velocity (from batch data to streaming data), variety, value, and complexity (ranging from structured to unstructured data).

The proliferation of Big Data is facilitated by the expansion of storage capacity, enhanced processing power, and the ready accessibility of extensive datasets. However, handling data at this scale necessitates a different approach encompassing unique techniques, tools, and architectural considerations, both for addressing existing challenges more effectively and for tackling new problems that arise in this data-rich era.[1]

Evolving Analytics for E-commerce Applications:

This study delves into the transformation of an analytics application in response to recent technological advancements and the widespread proliferation of data generated by popular platforms, such as mobile devices and social networks. The unprecedented surge in data, particularly in the form of unstructured social media data, disseminated across numerous computational systems, presents a plethora of both pertinent and extraneous information. In particular, this research zeroes in on how mentions of an e-commerce platform on social media channels serve as a form of promotional content, potentially driving traffic and bolstering revenue. The analytics application plays a pivotal role in quantifying both website traffic and revenue, thereby facilitating an assessment of the influence of each communication medium.

To expedite this procedure, the paper delineates the automatic generation and posting of content featuring direct links to preselected, distinct product collections, sets of merchandise, or exclusive promotions through the utilization of the Application Programming Interfaces (APIs) offered by social media platforms. Robust big data methodologies are harnessed for the real-time distribution of metrics, performance visualizations, statistical reports, and evaluations. Furthermore, the paper underscores the intricacies associated with data replication, data transmission, and data curation.

The paper offers an exposition of the evolutionary journey of a Log File Analyzer Application, adapting to the ever-expanding landscape of big data technologies to fortify its role in bolstering analytics applications. It underscores the paramount influence of social media platforms in the relentless surge of data. The paper accentuates the imperative of extracting pertinent data from diverse origins and seamlessly delivering it to analytics applications with minimal latency. A key aspect highlighted is the central role played by middleware, which harnesses specialized APIs from a multitude of social media providers, in orchestrating this intricate orchestration.

Furthermore, the paper delves into the intricacies associated with amalgamating diverse data streams and origins stemming from various applications hosted on disparate servers. This complex process necessitates the utilization of high-throughput distributed messaging systems. The imperative role of data cleansing is underscored, as it is pivotal in extracting pertinent information essential for e-commerce applications. The wealth of information derived holds the potential to augment decision-making capabilities and furnish valuable comparative metrics, especially when scrutinizing the origins of website visitors identified through the access log files of web servers.

To sum up, the paper underscores the significance of precision in gauging the impact of social media on both revenue and website traffic. The Analytics Application assumes a pivotal role in appraising the influence of each social media platform utilized, and in optimizing the frequency of automatically generated postings accordingly. [2]

Using Big Data Analytics in E-Commerce A Review from the Viewpoints of Vendors and Customers:

This paper explores the challenges posed by the information revolution in e-commerce due to the vast amount of data that needs processing and analysis. It emphasizes the role of Big Data Analytics (BDA) in enhancing decision-making processes through the analysis of significant data sources, including messages and social media posts. BDA is positioned as a crucial element in e-commerce, focusing on increasing vendors' revenues and attracting customers.

The objective of this investigation is to ascertain the advantages of employing Big Data Analytics (BDA) in the field of e-commerce, benefiting both sellers and customers alike. To gauge the impact of BDA in this domain, the study conducts an analysis of fifteen carefully selected research studies. E-commerce vendors employ BDA to gain a competitive advantage, enhance their comprehension of consumer behaviour, and foster customer loyalty. Furthermore, BDA-driven recommendation systems serve to personalize customers' search and shopping experiences, further amplifying the benefits of this analytical approach.

However, there are known drawbacks to using BDA in e-commerce, including the risk of shopping addiction and issues with the expense of BDA practitioners and equipment. In conclusion,

the paper underscores that while BDA enhances the electronic shopping experience for both consumers and vendors, the rapid growth of data remains a significant challenge.

In the introduction, the paper sets the stage by highlighting the massive amount of data generated by internet-based technologies and its categorization as Big Data. It explains how BDA has become essential for businesses to gain insights from this data. Furthermore, the paper delves into the surge of e-commerce, propelled by the proliferation of internet usage, and elucidates how e-commerce enterprises have harnessed Big Data Analytics (BDA) to streamline operations, elevate customer satisfaction, and boost revenues. The study endeavors to probe the repercussionsof Leveraging Big Data in E-Commerce: Revealing Strategic Insights for Optimizing Big Data Analytics in E-Commerce, Benefiting Both Vendors and Customers.[3]

Managing Applications with Big Data-Driven Precision Revolutionizing E-Commerce:

This paper explores the impact of internet applications on the generation of large and complex datasets, often referred to as big data. It highlights the role of Big Data Analytics (BDA) in extracting patterns, trends, and associations from this data. The central focus is on the potential revolution that BDA can bring to e-commerce applications. The paper discusses the importance of BDA tools and addresses various issues and challenges associated with big data analytics.

In the introduction, the paper emphasizes the exchange of information via the internet and the role of e-commerce. It underscores how the industry is shifting toward global leadership, solving business problems using big data. The rapid growth of e-commerce is intimately linked with the utilization of big data, spurred by internet technologies, social networking platforms, and advanced data capture methods. The paper underscores the transformative influence of big data, which engenders cost efficiencies, enhancements, and the provision of superior services.

Big data is crucial to e-commerce because it makes it possible to manage the enormous amounts of data created around the world.E-commerce generates both structured and unstructured data, making BDA essential. BDA has significantly impacted the e-commerce industry, enhanced self-service customer experiences and improving decision-making for competitive advantages. The paper highlights the various

characteristics of big data (Volume, Velocity, Variety, etc.) and its importance in e-commerce, particularly in handling unstructured data from social media.

In conclusion, the paper underscores the critical role of BDA in managing and extracting valuable insights from large datasets, particularly in the context of e-commerce. The automation of big data is increasingly regarded as a pivotal element in shaping the future of e-commerce, adeptly addressing the myriad challenges stemming from the sheer volume and intricacy of data. Within this landscape, social media emerges as a paramount data source, assuming a pivotal role in the domain of e-commerce, particularly in the realms of product promotion and marketing.[4]

E-commerce Marketing Based on Big Data Analysis and Processing:

This paper addresses the growing challenge of information overload on the internet and its impact on consumer energy and information processing abilities. It discusses how e-commerce enterprises are utilizing big data for personalized shopping guidance. The paper focuses on analyzing the e-commerce industry's development in the context of big data and proposes methods for improvement. It introduces a novel model for e-commerce marketing based on the analysis and processing of big data.

In the introduction, the paper highlights the heterogeneity of big data, which encompasses data from various physical sources with multi-modal characteristics. This includes social media data distributed widely on the internet, such as micro-blog data containing user attributes, natural text, images, videos, and audio. The paper emphasizes that even data with the same attributes from different physical spaces may not share the same distribution characteristics, making big data analysis complex.

The rise of social networks and the influx of unstructured data types, including user-generated content, audio, text, videos, and photographs, have ushered in the era of big data, which is characterised by its inherent data opulence. Additionally, the internet of things and mobile internet contribute to the wealth of data available. However, the paper notes that hardware infrastructure has struggled to keep pace with this rapid data growth.

In summary, the paper addresses the challenges posed by information abundance on the internet and the role of big data in personalized shopping guidance within the e-commerce industry.

It acknowledges the complexity of heterogeneous data and highlights the limitations of current hardware infrastructure in handling this data deluge.[5]

Cutting-Edge E-Commerce Analytics: In the Big Data Era, Real-Time Clickstream Data Ingestion:

This paper's core focus lies in the monitoring of visitor browsing behaviors on e-commerce websites, aiming to gain a deeper understanding of their intentions and preferences in order to enhance online marketing strategies. It specifically delves into the intricacies of collecting and processing real-time clickstream data, which comprises a sequence of page requests triggered by user clicks, each leading to unique URLs. The paper introduces a model for real-time clickstream data ingestion within the e-commerce domain, utilizing established Big Data tools such as Kafka, Flume, Spark, and Cassandra.

The primary emphasis of this study is on building high-velocity, fault-tolerant streaming data acquisition pipelines in a distributed setup, rather than delving into the analysis of data patterns. Clickstream data records user interactions with web content by capturing the URL generated for each click event. For instance, in an online shopping portal, a clickstream data entry might encompass user details, the product they clicked on, its price, and their geographical location.

The paper underscores the importance of efficiently capturing vast volumes of high-velocity clickstream data through the use of Big Data ecosystem tools. It highlights potential applications for clickstream data in e-commerce, including understanding user navigation paths, predicting product associations and future purchases, and optimizing the overall user experience.

Contributions made by this paper include presenting a data ingestion framework using Big Data tools and addressing the handling of late-arriving data through heuristic watermarking. The paper also demonstrates the framework's horizontal scalability and introduces an event chaining scheme for aggregating click events.

In conclusion, this paper provides valuable insights into the real-time capture and ingestion of clickstream data in the context of e-commerce, with a strong focus on the architecture and scalability of the data pipeline. It also explores potential applications for clickstream data analysis, ultimately aiming to enhance the e-commerce user experience and optimize online marketing strategies.[6]

SURVEY ONBIGDATA ANALYTICS:

This paper addresses the challenges posed by the generation of terabytes of data daily from Cloud Computing, the Internet of Things (IoT), and modern information systems. It underscores the complexity and high dimensionality of data generated by e-commerce transactions, highlighting the limitations of traditional methods in storing and analyzing this vast volume of data. The paper explores the field of big data analysis, its challenges, issues, and various advancements.

In this era of Big Data, the paper defines big data as large amounts of complex and multifaceted data that have transformed traditional technologies. Organizations are increasingly utilizing critical data in various fields, such as monitoring object movements, sensor deployments, and data tracking. Big data plays a crucial role in uncovering hidden information and gaining advantages. Researchers are actively developing novel data analysis techniques for big data, leading to the continuous evolution of various algorithms and platforms.

The paper discusses specific scenarios where organizations need to analyze data from their websites to understand customer feedback and tailor services accordingly. It emphasizes that decision-makers often rely on the analysis of collected data to make informed decisions.

The Information Discovery Process, known as Knowledge Discovery in Databases (KDD), is introduced as a methodology for extracting valuable information from unstructured datasets through detailed analysis and interpretation.

The paper mentions government initiatives to encourage research in Big Data analytics due to the rapid growth in data generated by organizations like Google and Facebook, with vast amounts of data, such as petabytes, being generated and processed regularly.

In conclusion, the paper provides insights into the challenges and opportunities presented by Big Data in the context of e-commerce and modern information systems. It highlights the need for advanced data analysis techniques and the growing importance of extracting valuable insights from vast datasets.[7]

Big data analytics & its application in E-commerce:

The importance of big data analytics in the field of e-commerce is covered in this paper. It demonstrates how the internet and technological progress have produced huge quantities of data that can give firms insightful information. Both small and large e-commerce companies are utilizing big data analytics to gain a competitive edge by deeply understanding their markets, products, and clientele.

The paper emphasizes that e-commerce companies can harness big data capabilities and various analytic techniques to comprehend customer purchasing behaviour. By doing so, they can tailor their marketing strategies to directly align with customer preferences, develop products based on customer demands, and ensure a high level of customer service. The research focuses on qualitative analysis and includes relevant use cases to demonstrate the practical applications of big data analytics in e-commerce.

In the introduction, the paper acknowledges that a significant portion of the digital universe consists of e-commerce data, encompassing various aspects like web browsing histories, social media activities, geolocation services, and online shopping records. E-commerce transactions, whether single or recurring, occur online through platforms like Flipkart, Amazon, and Myntra. Data collected by e-commerce firms falls into four categories: business activity or transaction data, click-stream data, audio data, and video data.

The paper underlines the importance of tracking consumer shopping behaviour using data accumulated over time from browsing histories. It suggests that e-commerce businesses need to adapt to evolving consumer behaviour, and effectively analyzing the gathered data is crucial for their success. Given the exponential growth in data volume, traditional data management tools are inadequate for storing and processing this data. Big data analytics, with its 3Vs model (volume, velocity, variety), and later extended to the 5Vs model (volume, veracity, velocity, value, variety), is presented as a solution to effectively analyze large datasets to reveal patterns and trends in human behaviour.[8]

BIG DATA ANALYTICS: HADOOP ANDTOOLS:

This paper delves into the pivotal role of information technology in managing extensive data quantities, notably encompassing both unstructured and structured data derived from diverse sources like emails, web logs, and popular social media platforms such as Twitter and Facebook. The document underscores the hurdles encountered when dealing with big data, encompassing data capture, storage, retrieval, sharing, and analysis.

The introduction acknowledges the exponential growth of data due to technological advancements and the rise of social networking sites. The amount of data produced continues to increase, requiring efficient processing methods. Big data technologies are essential for accurate analysis, leading to informed decision-making, operational efficiency, cost reduction, and risk mitigation for businesses.

The paper emphasizes the need for infrastructure capable of managing and processing large volumes of both structured and unstructured data in real-time while ensuring data privacy and security. Introducing Hadoop, an open-source framework coded in Java, as a solution that underpins parallel and distributed data processing while providing dependable data storage. Hadoop's chief advantage lies in its cost-efficiency, rendering it particularly apt for managing vast data files. Hadoop harnesses the Hadoop Distributed File System (HDFS) for storage and employs Map-Reduce for processing.

The paper briefly mentions various Hadoop-related projects, including Avro, Pig, Hive, Chukwa, HBase, MapReduce, HDFS, and Zookeeper. These tools and frameworks contribute to the effectiveness of Hadoop in managing and processing big data.

The focus of the paper is on the use of big data analytics in the e-commerce sector, where companies utilize data from website traffic and user navigation patterns to understand individual and group preferences, interests, and dislikes based on previous purchases. Additionally, the paper mentions the comparison of some commonly used data analytic tools.[9]

Platforms for Big Data Analytics: Trend towards Hybrid Era:

The paper focuses on the challenges and opportunities presented by Big Data in the field of data analysis. It highlights the importance of choosing the right hardware and software platforms to effectively analyze Big Data, especially in terms of scalability and adaptability. Various Big Data platforms are discussed, and their strengths and weaknesses are examined.

The introduction emphasizes the role of Big Data in transforming traditional data analysis methods. To choose the appropriate platform, users must consider factors such as their application needs, time constraints, data size, the nature of the processing model (iterative or single iteration), future data processing requirements, data transfer speed, data type, and hardware failure handling. These considerations align with the five Vs of Big Data: Value, Volume, Velocity, Variety, and Veracity.
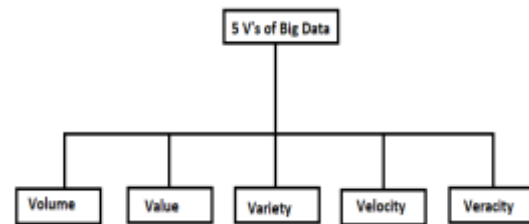

Figure:1 5Vs in Big Data

Big Data is defined as datasets that are large, diverse, and high velocity. Handling such data with conventional tools and platforms is challenging. The rapid growth of data is attributed to advancements in networking, data storage, and data gathering across various domains.

The central goal of this paper is to offer a comprehensive examination of platforms well-suited for the processing of Big Data. It conducts a survey of diverse software frameworks designed for Big Data analytics, meticulously evaluating their merits and shortcomings. Furthermore, the document delves into data mining algorithms and their applicability to real-world challenges within the realm of Big Data analysis. The authors propose that the effective implementation of established data mining algorithms, taking into account the strengths of available software frameworks and platforms, can facilitate the forecasting of future trends in Big Data processing and analytics. The paper also contemplates hybrid approaches that amalgamate multiple platforms, with a specific focus on their adaptability and performance advantages in the context of real-time data processing.
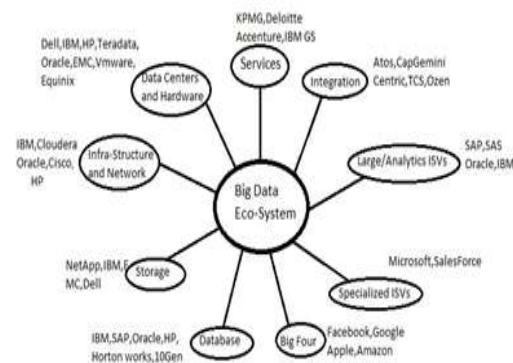

Figure:2 Big Data Eco-system

The paper doesn't explicitly mention e-commerce, but the discussed concepts and platforms have broad applications, including in e-commerce, where the analysis of large and diverse datasets is crucial for business intelligence and decision-making.[10]

Case studies: Success stories of e-commerce companies using big data analytics:

Case studies offer priceless insights into how e-commerce businesses have used big data analytics to fuel growth and offer individualised consumer experiences. These success stories serve as inspiration and guidance for businesses looking to harness the potential of big data in their own operations.

One notable case study is the e-commerce giant, Amazon. With its vast array of products and millions of customers, Amazon has utilized big data analytics to fuel its growth and success. By analyzing customer browsing and purchasing patterns, Amazon can recommend personalized product suggestions, tailor advertisements, and even optimize its supply chain and inventory management. This level of personalization has played a significant role in enhancing the customer experience and driving customer loyalty.

Another compelling example is Netflix. As a leading streaming service, Netflix collects massive amounts of data on user viewing preferences and behaviour. By applying advanced analytics and machine learning algorithms to this wealth of data, Netflix can curate personalized recommendations, create targeted content, and optimize its user interface. This data-driven approach has been instrumental in attracting and retaining subscribers, keeping them engaged, and ultimately driving business growth.

Furthermore, clothing retailer Stitch Fix has successfully utilized big data analytics to revolutionize the online shopping experience. By analyzing customer data, including style preferences, sizes, and feedback on previous purchases, Stitch Fix can provide individualized clothing selections to its customers. This personalized approach has not only increased customer satisfaction but also significantly improved conversion rates and revenue for the company.

These case studies demonstrate the transformative impact of big data analytics on e-commerce businesses. By leveraging the power of data, companies can uncover valuable insights, make data-driven decisions, enhance customer experiences, and ultimately drive growth. As more

businesses recognize the potential of big data analytics, we can expect to see further success stories emerge in the e-commerce industry.

III.   METHODOLOGY

Big data analytics is the process of collecting, cleaning, and analyzing large datasets to extract meaningful insights. E-commerce businesses can use big data analytics to improve their operations in serval ways, they are:

- Recommendation Systems: Recommendation algorithms, such as collaborative filtering, content-based filtering, and hybrid methods, help e-commerce platforms suggest products or services to customers based on their preferences, browsing history, and behaviour. These algorithms are crucial for enhancing the customer shopping experience and increasing sales.
- Market Basket Analysis: Market basket analysis uses association rule mining algorithms like Apriori and FP-Growth to identify patterns in customer purchase behaviour. This helps in cross-selling and upselling products by suggesting complementary items to customers based on what others have purchased.
- Customer Segmentation: Clustering algorithms like k-means or hierarchical clustering are used to group customers with similar characteristics together. This allows e-commerce companies to tailor marketing strategies, promotions, and product recommendations to specific customer segments.
- Predictive Analytics: Machine learning algorithms like regression, decision trees, and random forests are used for predictive analytics. They help forecast future trends, demand, and sales, enabling businesses to optimize inventory management and supply chain operations.
- Natural Language Processing (NLP): NLP techniques, including sentiment analysis and text mining, can be used to analyze customer reviews, social media comments, and product descriptions. This helps e-commerce companies understand customer sentiment, identify emerging trends, and improve product listings.

Now we are representing one of the algorithms which is used for recommendation system.

Apriori Algorithm:The term "Apriori algorithm" pertains to the computational method employed for deducing association rules between

objects, elucidating the relationships among two or more entities. In simpler terms, it can be described as an association rule learning technique that discerns patterns like "people who purchased product A also tend to purchase product B."

The primary aim of the Apriori algorithm is to establish these association rules among various objects, effectively delineating how multiple objects are interconnected. This algorithm is often referred to as "frequent pattern mining." Typically, the Apriori algorithm is applied to databases containing a substantial volume of transactions.Pseudo-code:

C_k: Candidate itemset of size k
L_k : frequent itemset of size k
L_1 = {frequent items};
for (k = 1; L_k !=Ø; k++) do begin
C_k +1 = candidates generated from L_k;
for each transaction t in database do increment the count of all candidates in C_k+1 that are contained in t
L_k +1 = candidates in C_k+1 with min_support
end
return ?$_k$ L_k;

Figure:3 Apriori pseudo-code

Let's implement the algorithm using sample data as example, Problem:For the following given Transaction Dataset, Generating the final data setbyApriori. And consider the values as Support=50% and Confidence=75%.

Transaction ID's	Items Purchased
1	A, B, C, D
2	A, B, D
3	A, E, F
4	A, D, E
5	B, D, E

Step-1 Finding the Frequent Item set and their Support:

In this step we will be finding the frequency count for all the distinct items from the given transaction data and finding the support value for all the distinct items.

Support = (number of transactions containing item) / (total number of transactions)

Item	Frequency	Support (%)
A	4	80%
B	3	60%
C	1	20%
D	4	80%
E	3	60%
F	1	20%

Step-2 Removing all items with support levels below the required minimum.
Remaining items after removing which has less than the given minimum support 50%.

Item	Frequency	Support (%)
A	4	80%
B	3	60%
D	4	80%
E	3	60%

Step-3 Now forming the two items candidate set and writing their frequencies.

In this step we will be forming two items candidate set for the remaining items which are greater than the given support value. And finding their frequency and support values.

Items	Frequency	Support (%)
A, B	2	40%
A, D	3	60%
A, E	2	40%
B, D	3	60%
B, E	1	20%
D, E	2	40%

Step-4 Removing all items with support levels below the required minimum.
Remaining items after removing which has less than the given minimum support 50% but here the items are two candidate set items.

Items	Frequency	Support (%)
A, D	3	60%
B, D	3	60%

Step-5 Generating Rules for finding confidence
In this step we will be generating rules for the item sets for finding the Confidence value.
Confidence (A->B) = support (AUB)/support (A)
For Rules we consider item pairs as:
a) (A, D)
A->D and D->A
b) (B, D)
B->D and D->B

Rules	Confidence	Confidence in %
A->D	3/5 * 5/4 = ¾	75%
D->A	3/5 * 5/4 = ¾	75%
B->D	3/5 * 5/43= 1	100%
D->B	3/5 * 5/4 = ¾	75%

Because each rule's confidence is higher than or equal to the minimal confidence of 75% specified in the problem, all of the above pairs are strong candidates.

And confidence for the B and D is 100% means who ever purchases the B they will also purchase D.

IV.   RESULTS

Within this section, a concise examination of big data analytics in the realm of E-commerce elucidates the fact that E-commerce can achieve substantial growth by harnessing insights derived from big data analytics. The algorithms employed culminate in the provision of tailored recommendations to customers, offering products closely associated with their prior purchases.

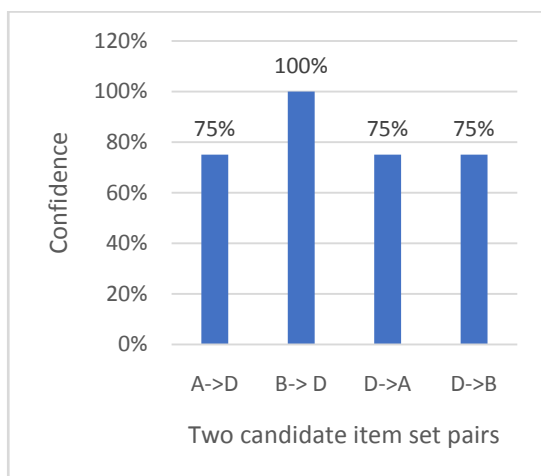| Items	Confidence
A->D	75%
B->D	100%
D->A	75%
D->B	75%

Figure:4 Result analysis

All the items are likely to be related and this is the pattern found from the past transactions. This way of recommending is so reliable and effective.

V.   CONCLUSION

The fusion of Big Data Analytics and e-commerce presents a potent strategy for developing digital products that align with the ever-changing demands of users in a swiftly evolving market. Whether it's tailoring personalized product recommendations or swiftly adapting pricing strategies in real-time, the influence of big data is profoundly extensive.It equips e-commerce enterprises with the agility to respond swiftly to market dynamics and consumer preferences. The result is a heightened customer experience and improved business outcomes.

While there's no one-size-fits-all approach to big data analytics, it is a critical asset that no e-commerce business can afford to ignore. The future of e-commerce belongs to those who embrace and harness the power of big data, as it continues to drive innovation and success in this digital age.

REFERENCES

[1] Neha Sharma1, Deepali Sawai2, Ganesh Surve3,2017, "Big Data Analytics: Impacting Business in Big Way", IEEE.

[2] Constantine J. Aivalis, KleanthisGatziolis, Anthony C. Boucouvalas,2016 ,"Evolving Analytics for E-commerce Applications Utilizing Big Data and Social Media Extensions", IEEE.

[3] SARAH S. ALRUMIAH 1 AND MOHAMMED HADWAN 2,2021, "Implementing Big Data Analytics in E-Commerce: Vendor and Customer View", IEEE.

[4] Sakshi Painuly1, Sachin Sharma2, Priya Matta3, 2021 ,"Big Data Driven E-Commerce Application Management System", IEEE.

[5] HONGSHENG XU, GANGLONG FAN, Ke LI, 2017 ,"Novel Model of E-commerce Marketing Based on Big Data Analysis and Processing", IEEE.

[6] Gautam Pal, Gangmin Li, Katie Atkinson,2018, "Big Data Real-Time Clickstream Data Ingestion Paradigm for E-Commerce Analytics", IEEE.

[7] Saba Ahmad, Paramdeep Singh, Dr. Anil Kumar Sagar,2018," A SURVEY ON BIG DATA ANALYTICS", IEEE.

[8] Anagha Ramkumar1, Prasanna Kulkarni1, Ahmed J. Obaid2, Azmi Shawkat Abdulbaqi3, Ahmad Al Yakin4,2023, "Big data analytics & its application in E-commerce", AIP Publishing.

[9] Mrunal Sogodekar, Shikha Pandey, Isha Tupkari, Amit Manekar, 2016 ,"BIG DATA ANALYTICS: HADOOP AND TOOLS", IEEE

[10] Alka Londhe, PVRD Prasada Rao, 2017 ,"Platforms for Big Data Analytics: Trend towards Hybrid Era", IEEE.