IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Designing And Implementing A Cloud Based E-Commerce Recommendation System Model

Dasari Ashok
Assistant professor
Department of Computer Science & Technology
Madanapalle Institute of Technology & Science, Madanapalle

E. Venkata Teja
Department of Computer Science & Technology
Madanapalle Institute of Technology & Science, Madanapalle

T. Uday Kiran
Department of Computer Science & Technology
Madanapalle Institute of Technology & Science, Madanapalle

R. Sai Pramod
Department of Computer Science & Technology
Madanapalle Institute of Technology & Science, Madanapalle

N. Harish
Department of Computer Science & Technology
Madanapalle Institute of Technology & Science, Madanapalle

Abstract:

The ability to make excellent product recommendations to users is critical for improving user experience and boosting business success in the quickly changing e-commerce industry. The goal of this project is to give consumers personalized product recommendations by designing and implementing a cloud-based e-commerce recommendation system model. By utilising sophisticated recommendation algorithms, such as content-based and collaborative filtering, in conjunction with expandable cloud infrastructure, the system effectively handles massive amounts of data to produce precise recommendations. Scalability, performance, and security are guaranteed by the system architecture, which includes cloud deployment, preprocessing, recommendation engines, and data gathering. The system's efficacy in enhancing user engagement and conversion rates is demonstrated through thorough assessment and testing. The results of this study give a basis for the development of recommendation systems in e-commerce.

1.Introduction

The way people purchase has been completely transformed by e-commerce, which provides unmatched accessibility and ease to a wide range of goods. But because there are so many options accessible, consumers frequently experience information overload and decision fatigue. E-commerce platforms are depending more and more on recommendation algorithms to provide customized product recommendations based on customer preferences in order to overcome this difficulty and improve user happiness. The project's goal is to create and put into use a cloud-based e-commerce recommendation system model that uses scalable infrastructure and cutting-edge algorithms to deliver timely and accurate recommendations[1]. This system hopes to provide effective and scalable recommendations by utilising cloud computing, which would enhance user engagement and retention and eventually spur business growth. Our goal is to contribute to the advancement of recommendation systems in the e-commerce industry by exploring the design principles, implementation methodologies, and performance characteristics of such a system through this project[10].

2.Related & Existing work:

2.1.A Framework to Secure Medical Image Storage in a Cloud Computing Environment.

Thanks to recent advancements in imaging equipment, contemporary healthcare professionals are now producing enormous amounts of medical images on a daily basis. Generally speaking, this is because more patients are requesting medical care. As a result, there is a constant need for a lot of storage space. Regretfully, local data centres are still used by the healthcare industry to manage business procedures and store medical data[11]. This has a major detrimental effect on running costs related to maintenance and licence payments. Healthcare organisations are considering cloud storage as an alternative to on-premise hosted solutions in order to address these issues. The availability, affordability, and scalability of cloud services serve as the primary justifications for this. This model's main goal is to contract out data and assign IT calculations to a third party.The latter uses the Internet to provide the necessary storage systems in order to meet customer requests. While there are huge cost savings with this arrangement, adopting cloud storage presents security risks. This article outlines many strategies that have been suggested to guarantee data safety in order to achieve this goal[10]. Numerous restrictions plague the current implementations. The authors provide a system for cloud computing medical picture storage that is secure. They achieve this by employing watermarking and multi-region segmentation techniques, which preserve secrecy and integrity. Furthermore, they depend on an ABAC paradigm to guarantee cloud storage access control. Four key components make up this solution: (1) split data to safeguard privacy; (2) authentication to access medical datasets; (3) integrity checking; and (4) access control to impose security measures[8].

2.2.Outlining the Issues of Cloud Computing and Sustainability Opportunities and Risks in European Organizations

Since cloud computing and sustainability help firms and individuals become distinctive and exclusive in their work and study, they have become integral parts of fundamental strategies in organisations both locally and globally. In order to save expenses and boost productivity, businesses and people should incorporate sustainability into their plans and use cloud computing technology as a tool for sustainable work, particularly in information technology (IT) departments[6]. Using an online survey aimed at 56 information technology managers in Europe, this research investigates the awareness of cloud computing and sustainability benefits and hazards among European organisations. To assess the survey results, a conceptual model of cloud computing was created using structural equation modelling (SEM). The study's findings demonstrated that there are technological advantages to cloud computing, including sustainability in the organization's strategy, will enhance their job performance and job satisfaction, use and awareness; however, security, privacy and risks are still a major concern[5].

2.3.About the services of big data and cloud computing for e-commerce.

An option for on-site devices is cloud hosting. Keeping information on multiple devices is made possible via cloud technology, as opposed to storing it all on one machine. These gadgets are dependent on different servers that are situated far apart. All machines, however, are linked together and function as a single unit. A hosting server is used by the majority of eCommerce companies to combine the processing power of multiple devices. As a result, they establish a server cluster. And we refer to this grouping as a cloud. The functionality of online stores differs from other websites. To keep databases updated, safeguard personal information, handle payments, and monitor sales, they require more assistance and adaptability. As a result, moving to the cloud or adopting it is a great way to tackle issues that other devices can't[10]. Let's review the five main methods that online retail stores can use cloud computing to their advantage[3].

2.4.Information Systems:

Without a question, there has been a significant shift in business and organisational methods due to recent advancements in information technology, particularly information systems. Because the global financial systems are interdependent, even a slight drop in interest rates in the US could result in a disproportionately significant increase in South East Asia's stock market values[6]. The information revolution has given rise to completely new marketing strategies, which have drastically altered the dynamics between manufacturers, distributors, and consumers. Additionally, it has caused modifications to the organisational structure, management style, and methods for handling goods and services. Specifically, it presents a variety of opportunities and difficulties that impact the process by which data is transformed into information, and that information into knowledge. It presents numerous chances for environment management for these transitions, including raising individual and organisational productivity. This essay explores the relationships and points of convergence between information technology, information systems, and organisations. It also looks at how innovative information system utilisation affects organisational contexts[7].

2.5.Research on e-commerce data caching based on cloud computing :

Based on the current circumstances, it appears that China's e-commerce businesses are growing at an increasingly rapid pace, and e-commerce platforms' sales figures are consistently breaking new ground. However, a number of issues have surfaced as a result of this quick development[4]. The dangers encountered by traditional e-commerce companies and traditional organisations differ fundamentally due to the differences in B2C e-commerce business management methods from traditional business management methods. The risks that e-commerce businesses confront are more complicated and unclear than those encountered by traditional businesses. As a result, risk management strategies have not kept up with the rapid development of B2C e-commerce and are currently in the stage of ongoing investigation[2]. The need for e-commerce data management model study is pressing due to the growing popularity of e-commerce and its greater application. Theoretical research and the commercial application of cloud computing are examined in detail through literature analysis, comparative analysis, graphical analysis, and case analysis in order to explore the e-commerce data management model. My aim is to broaden the study of cloud computing platforms' applications in e-commerce. In the end, it demonstrates that e-commerce will have a strong competitive advantage going forward[3].

3.Proposed Model

3.1.User Interface:

The main point of contact between users and the e-commerce platform is the user interface. It consists of:Product browsing allows users to look at information, search for particular products, and explore products. Personalised recommendations: Recommendations are shown to users in a prominent manner and offer pertinent suggestions based on their browsing history and interests[7]. User profiles:Allows users to view order history,modify preferences, and manage accounts. Smooth navigation: To guarantee a seamless user experience across devices, the interface is made with ease of use in mind. It features responsive design and intuitive navigation.

3.2.Data Collection:

Obtaining different kinds of data from consumers and the e-commerce platform is known as data collection. User interactions: Views, purchases, add-to-cart operations, clicks, and so forth.Features of the product: Title, description, price, category, pictures, etc. Transaction history: Shipment addresses, payment details, order data, etc. User input includes reviews, ratings, questionnaires, and more. Real-time data collection and archiving takes place in a data warehouse or centralised database for later examination[8].

3.3.Recommendation Engine:

The fundamental component of the system is the recommendation engine, which provides users with customised recommendations. It makes use of several different recommendation algorithms, such as:

objects are recommended using collaborative filtering, which considers similarities between users or objects. It finds users who share similar interests and makes recommendations for products that these users have either enjoyed or bought[4].

Content-based filtering suggests products based on what the user has enjoyed or engaged with in the past. To identify products that align with the user's interests, it evaluates both product attributes and user preferences[9].

To increase recommendation relevancy and accuracy, the recommendation engine constantly examines user input and data.

3.4.Real-time Recommendation Generation:

Real-time recommendations are produced as users engage with the platform: Personalised suggestions are updated dynamically in response to user activity, including searches, purchases, and browsing habits.

In order to provide prompt and pertinent recommendations and guarantee a flawless user experience, the recommendation engine analyses user data and uses recommendation algorithms[8].

3.5.User Feedback and Iterative Improvement:

In order to increase user happiness and suggestion accuracy, user feedback is essential: User preferences and satisfaction levels can be learned a great deal from ratings, reviews, and feedback surveys[5].

In order to continuously improve recommendation algorithms and adjust to shifting customer preferences, the recommendation engine takes user feedback into account.

Iterative experimentation and testing aid in determining areas for improvement and validate the efficacy of suggestion tactics[3].

3.6.Security and Privacy: To safeguard user privacy and preserve user data, security procedures are put in place: Data encryption: To prevent unwanted access, sensitive data, including payment details and user passwords,is encrypted.Controls over access: Based on user responsibilities and permissions, role-based access controls restrict access to sensitive data and system operations[6].

Compliance with rules: To preserve user privacy and

guarantee legal compliance, the system conforms with data protection regulations, such as GDPR[9].

3.7.Scalability and Performance:

Large data volumes and user requests can be handled by the system while retaining peak performance[6]. Cloud infrastructure: Scalable and adaptable infrastructure resources can be made available to meet increasing demand by utilising cloud services.
Distributed computing: By enabling parallel processing and effective resource utilisation, distributed computing technologies enhance the scalability and performance of systems.Load balancing: In order to avoid overloading and maintain steady performance even under high load circumstances, load balancing solutions divide incoming traffic among several servers[5].

4.Algorithm

4.1.Data Preprocessing: To start, make sure the product dataset is organised and tidy by performing preprocessing. If necessary, take care of any missing values and remove any unnecessary columns[3].

4.2.Feature Extraction:

Take relevant attributes that describe each item and extract them from the product dataset. Product qualities including title, description, category, brand, price, and any other metadata that may affect consumer preferences might be included in these features.Transform text-based features (such product descriptions) into numerical representations by utilising word embeddings like Word2Vec or term frequency-inverse document frequency (TF-IDF) methods. As a result, feature vectors that may be utilised in machine learning models are created from the textual data[3].

4.3.User Preferences:

Record user choices according on how they engage with the online store. This could contain information on products that have been seen, bought, loved, or rated, as well as search history.User preferences should be represented as a feature vector, with each element denoting the user's affinity for a particular feature or attribute[2].

4.4.Similarity Calculation: Determine how similar each product's attributes are to the user's preferences (feature vector) for each product in the dataset.Depending on the features and the desired behaviour of the recommendation system, common similarity measurements include cosine similarity, Euclidean distance, Pearson correlation coefficient, or Jaccard similarity[5].

4.5.Recommendation Generation: Calculate the similarity score between each product in the dataset and the user's preferences.Sort the products into groups of N according to how similar they are, then choose which of the N products are most similar to the consumer.To offer new suggestions, you can choose to exclude things that the user has already engaged with or bought.

4.6.Presentation of Recommendations:

Using the user interface of the e-commerce platform, show the consumer the recommended products. This could manifest as email alerts, recommendation widgets, or a customised site.To assist visitors in making educated selections, include more details about each suggested product, such as a title, image, price, and a succinct description.

5.Block Diagram

5.1.User: After supplying the required data, users register with the system[8]. Users can use the features of the system after registering.

5.2. Admin Module: Admins oversee account activation and user registration. Users can utilise the system and log in once it has been activated.
5.3.Data Collection: The system gathers information from multiple sources, including user interactions and product properties.

5.4.Data Preprocessing: Preprocessing involves the extraction of features, integration, cleaning, and normalisation of raw data. This gets the data ready for analysis and creating recommendations.

5.5. User Interests: Based on how users engage with the system for example, by seeing or purchasing items the system records their preferences and interests.

5.6. Product Representation: To facilitate comparison and similarity calculations, product attributes are represented as feature vectors.

5.7.Recommendation Generation: User preferences and product attributes are compared using a content-based recommendation system. The system provides customers with customized product recommendations based on this resemblance.

5.8. User Interface: The system's user interface presents suggested products to consumers, enabling them to explore and engage with the recommendations.

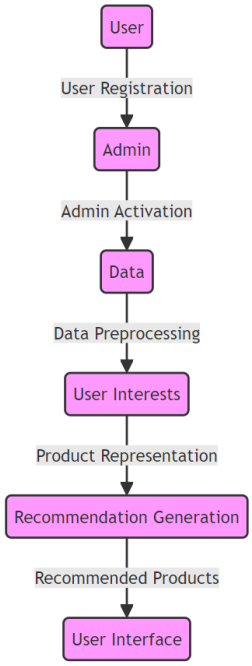

Fig 5.1.Block Diagram

This block diagram shows the entire workflow of the cloud-based e-commerce recommendation system with content-based filtering algorithms, from user registration to recommendation creation and display[19].

6.System Architecture

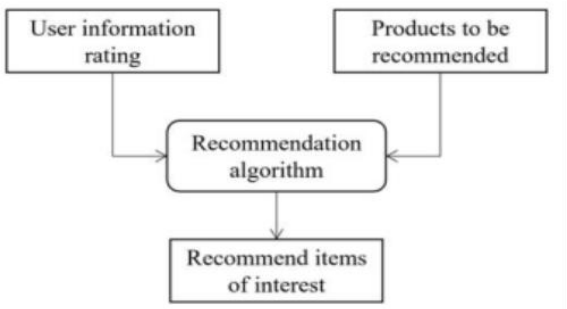

Fig 6.1.System Architecture

A cloud-based e-commerce recommendation system's system architecture handles user data, product data, recommendation algorithms, and the recommendation process through a number of important components. An description of each part is provided below:

6.1.User Information:

The term "user information" describes the data connected to specific users of the online store. Typical information in this data is: User demographics: Age, gender, location, etc. User interactions include product views, purchases, reviews, ratings, and more[9]. User preferences: Interest in certain product categories, favoured brands, financial limitations, etc. In order to provide personalised suggestions, user data is gathered and kept in a database or data warehouse. From there, it is fed into recommendation algorithms.

6.2.Rating Data:

User reviews and ratings are examples of user feedback on products that make up rating data. These evaluations enhance the precision of recommendation systems and offer insightful information about user preferences[10].

Users' interactions with the e-commerce platform yield rating data, which is then gathered and kept in a database. In collaborative filtering algorithms, it is frequently employed to find comparable users or objects according to their ratings[11].

6.3.Products to be Recommended:

The list of goods that are on sale in the online store is referred to as "products to be recommended." Included in this are the following product attributes: name, description, category, cost, etc.Product pictures are the items' visual representations[3].Usually, a database or catalogue management system is used to hold product data[18].

6.4.Recommendation Algorithm:

The main part of the recommendation system that provides customers with customised product recommendations is the recommendation algorithm. Typical algorithms for recommendations consist of: objects are recommended using collaborative filtering, which considers similarities between users or objects. Content-based filtering suggests products based on what the user has enjoyed or engaged with in the past[7].

Hybrid methods: Integrate several algorithms to take use of their advantages and raise the precision of recommendations[6].To provide recommendations that are specific to each user's preferences, the recommendation system analyses user information, rating data, and product data.

6.5.Recommend Items of Interest:

The functionality of the recommendation system that finds and suggests items of interest to users based on their preferences and behaviour is referred to as this component.To ascertain the applicability and relevance of products to recommend, it examines rating information, product features, and user interactions.Users can browse and explore the recommended items through the user interface of the e-commerce platform, which presents recommendations to them.

A cloud-based e-commerce recommendation system's system architecture combines these elements in order to efficiently gather user preferences, examine product information, and provide tailored recommendations that improve user experience, encourage interaction, and increase sales[14].

7.Modules

Every module in the cloud-based e-commerce recommendation system has a distinct function that helps to maximise the system's overall effectiveness and functionality[5].

7.1. User modules: Users can sign up and gain access to the system through the User Module. Users give their phone number and email address while registering in order to facilitate communication. The user cannot log in until the admin has activated their account once they have registered[2]. After logging in, users have access to submit datasets, see the outcomes of algorithm execution, and use the system's capabilities, such clicking on the Artificial Neural Network (ANN) to calculate accuracy and loss[1].

7.2.Admin Module:
offers administrative tools for controlling user profiles and system information. Using a browser interface, administrators can activate users who have registered, log in, and view system data overall[2]. In order to guarantee authorised access to the system, user activation is essential. Administrators can monitor how the system is run and have the ability to log out after completing their required administrative duties[6].

7.3.Data Preprocessing module: It is in charge of getting raw data ready for algorithms used in machine learning. In order to guarantee data quality and suitability for analysis, it entails procedures including data cleansing, integration, normalisation, and reduction. For instance, utilising the Mean technique to handle missing values guarantees that the data is complete and prepared for additional processing[10]. This module makes sure that the input data that is supplied into the machine learning models is consistent and formatted correctly[17].

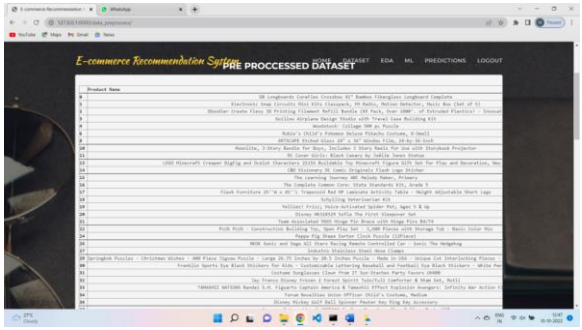

Fig 7.1.A snippet about pre-processing from the recommendation system

7.4.Machine Learning Module:
Based on the preprocessed data, the Machine Learning module creates predictive models using a variety of methods. Machine learning is used in the suggested model to improve the relevance and accuracy of the recommendation system. In e-commerce, methods like categorization methodologies are used to forecast consumer preferences and behaviour[15]. In order to maximise model performance and parameter selection, data is preprocessed using techniques like the mean/average method, z-score normalisation, and information gain before being formed into machine learning models[20]. This guarantees that, based on customers' past interactions and interests, the recommendation system offers them tailored and useful product recommendations.
To put it briefly, the suggested model combines these components to make it easier for users to connect with one another, manage administrative tasks, prepare data, and generate recommendations based on machine learning in cloud-based e-commerce settings. Every module is essential to maintaining the system's accuracy, efficiency, and functionality when it comes to providing consumers with personalized recommendations.

8.Conclusion and Discussion:

In today's world, e-commerce has emerged as the most popular business model.Personalised recommendation systems have garnered increasing attention in the era of big data in e-commerce due to the rapid expansion of information. People may purchase and consume goods on an increasing number of channels thanks to e-commerce[7]. E-commerce websites can increase customer turnover by using recommendation algorithms, such as data mining, to suggest products to customers based on their preferences. The amount of information in the e-commerce system is growing daily due to the quick development of commercial websites on the internet, and information overload is becoming a major issue[3].
Considering that user behaviour forms the basis of the recommendation system, malicious user behaviour ought to be filtered out. Some examples of malicious behaviour are when individuals create a huge number of user behaviours, browse specific products through programmes on a regular basis, or when corporations modify commodity attributes based on algorithms[8]. The evolution of e-commerce recommendation engines has given rise to a new mode of business recommendation thanks to the cloud's robust storage, operation, and security features, as well as its optimal mode of resource allocation and sharing. Regarding data integration, the business system's source data can be imported into the data analysis middleware layer for analysis, and the business system's data can also be directly saved in the distributed file system layer for administration and access[5].
Below you can find some snippets of the recommendation system:

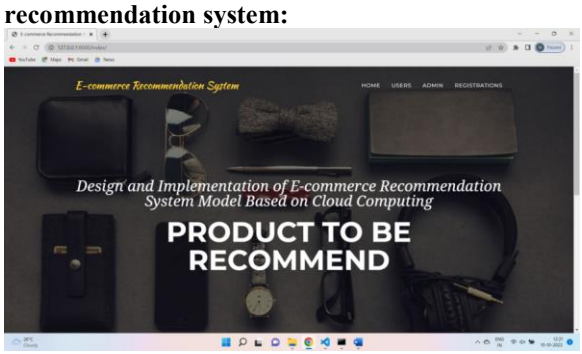

Fig 8.1. User Interface

The visual and interactive platform that allows consumers to interact with a cloud-based e-commerce recommendation system is called the user interface (UI). It includes components including websites, mobile applications, and voice interfaces that are intended to provide consumers with product recommendations in an easy-to-use and entertaining way. With features like personalized recommendations, simple navigation, and interactive elements like sorting and filtering, the user interface (UI) aims to improve the user experience. It is essential for encouraging user interaction, increasing conversion rates, and eventually enhancing the recommendation system's overall efficacy.

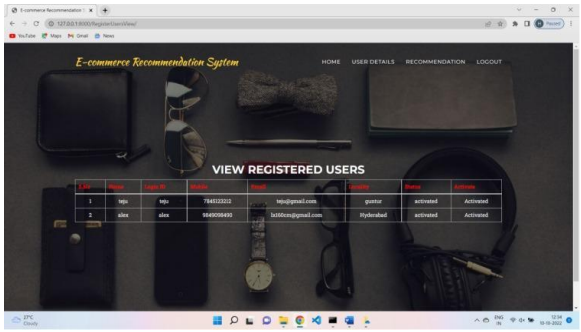

Fig 8.2.Registered Users

In the context of an e-commerce recommendation system, registered users are people who have set up profiles or accounts on the site. These users offer useful data, including their preferences, past purchases, demographic information, and possibly even outright criticism. Through registration, users use the system to customize their experience by receiving recommendations that are specifically suited to their own tastes and habits. Additional services like wish lists, order monitoring, and saved searches may also be advantageous to registered users, improving their overall purchasing experience. Additionally, registration enables the system to establish enduring connections with users, encouraging loyalty and maybe resulting in recurring transactions.

Evaluation scores are crucial in machine learning as they gauge the effectiveness of models trained with various methods, including cosine similarity, sigmoid kernel, and linear kernel.

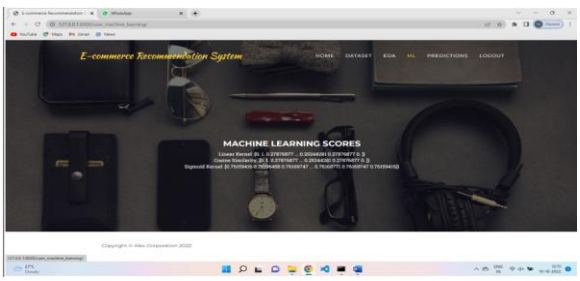

Fig 8.3.Machine Learning Scores

Linear Kernel: Models trained with a linear kernel are often evaluated using measures such as accuracy, precision, recall, and F1 score. These metrics work well for classification tasks, especially if there is linear separability in the data.

Sigmoid Kernel: Models trained with a sigmoid kernel are subject to comparable evaluation metrics. Furthermore, the non-linear structure of the sigmoid kernel makes measurements like ROC AUC useful, particularly in binary classification problems.

Cosine Similarity: Metrics unique to recommendation systems or similarity-based tasks are frequently used in the evaluation of models that use cosine similarity. Mean Average Precision (MAP), recall, and precision are frequently used to assess.

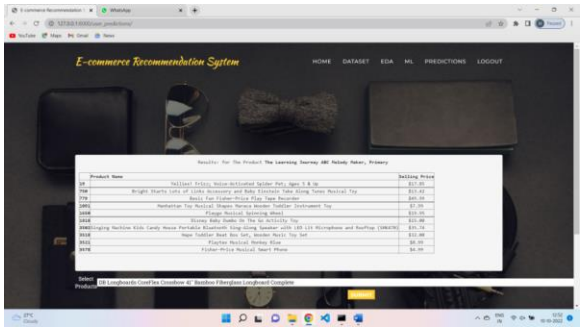

Fig 8.4.Recommendations made by our recommendation system

Our recommendation system provides individualized product or item recommendations based on each user's preferences, usage history, and behavior on the platform. Our system uses sophisticated algorithms, including content-based filtering, collaborative filtering, or hybrid techniques, to analyze large volumes of data and produce recommendations that are pertinent and accurate. By showing consumers things they are likely to be interested in, these recommendations hope to improve user experience, boost engagement, and eventually drive conversions. Furthermore, our system is always learning and adapting in response to user feedback and behavior, which guarantees that recommendations are current and accurately represent the evolving tastes of users over time.

9.Future Enhancements:

To increase suggestion accuracy, user engagement, and system scalability in the future, the cloud-based e-commerce recommendation system may concentrate on a number of important aspects[16]. To improve personalisation and adaptability, this involves investigating and putting into practice sophisticated recommendation algorithms, such as models based on deep learning or strategies based on graphs. Furthermore, the integration of contextual data and dynamic pricing tactics may provide more pertinent and prompt suggestions that are customised to the individual's inclinations and present situation. To

improve user experience and trust, the system could also benefit from multi-modal suggestions, interactive user interfaces, and transparent explainable AI[7]. The system's efficacy and relevance could be further increased via ongoing model training and adaptation, integration with social media and outside data sources, and privacy-preserving measures. Prioritising these upcoming improvements will help e-commerce companies keep up with changing user needs and preferences, driving increased user satisfaction and business success[5].

10.References:

[1] Marwan M, Kartit A, Ouahmane H. A Framework to Secure Medical Image Storage in Cloud Computing Environment. Journal of Electronic Commerce in Organizations, vol. 16, no. 1, pp. 1-16, 2018.

[2] Isaias P, Issa T, Chang V, et al. Outlining the Issues of Cloud Computing and Sustainability Opportunities and Risks in European Organizations: A SEM Study. Journal of Electronic Commerce in Organizations, vol. 13, no. 4) :1-25, 2015.

[3] Somalee Datta, Keith Bettinger, Michael Snyder. Corrigendum: Secure cloud computing for genomic data. Nature Biotechnology, vol. 34, no. 10, pp. 588-591, 2016.

[4] Datta S, Bettinger K, Snyder M. Erratum: Corrigendum: Secure cloud computing for genomic data. Nature Biotechnology, vol. 34, no. 10, pp. 1072-1072, 2016.

[5] Liu Mingyue. About the services of big data and cloud computing for e-commerce . Information and Computer (Theory Edition), vol. 352, no. 6, pp. 150-151, 2016.

[6] Chandra J, March S T, Mukherjee S, et al. Information Systems Frontiers. Communications of the Acm, vol. 43, no. 1, pp. 71-79, 2016.

[7] Jovanovic J, Bagheri E. Electronic Commerce Meets the Semantic Web. It Professional, vol. 18, no. 4, pp. 56-65, 2016.

[8] Ghosh N, Ghosh S K, Das S K. SelCSP: A Framework to Facilitate Selection of Cloud Service Providers. Cloud Computing IEEE Transactions on, vol. 3, no. 1, pp. 66-79, 2015.

[9] Mi Rensha Aini. Research on the Application of Cloud Computing in E-commerce. Computer Knowledge and Technology, vol. 11, no. 22, pp. 186-188, 2015.

[10] Bauer E. Improving Operational Efficiency of Applications via Cloud Computing. IEEE Cloud Computing, vol. 5, no. 1, pp. 12-19, 2018.

[11] Zhang Lan. Research on e-commerce data caching based on cloud computing. Computer Knowledge and Technology, vol. 12, no. 23, pp. 249-250, 2016.

[12] Chang Xueqin. Research on e-commerce logistics distribution model based on cloud computing. Logistics Technology, vol. 34, no. 11, pp. 152-155, 2015.

[13] Herzfeldt A, Wolfenstetter T, Ertl C, et al. The Role of Individualization and Project Learning for Cloud Service Profitability. Journal of Electronic Commerce in Organizations, vol. 16, no. 2, pp. 64-82,2018.

[14] Y. Choi and M. Rhu, “PREMA: A Predictive Multi-task Scheduling Algorithm For Preemptible Neural Processing Units,” in Proceedings of the International Symposium on High-Performance Computer Architecture (HPCA), 2020.

[15] B. Hyun, Y. Kwon, Y. Choi, J. Kim, and M. Rhu, “NeuMMU: Architectural Support for Efficient Address Translations in Neural Processing Units,” in Proceedings of the International Conference on Architectural Support for Programming Languages and Operation Systems (ASPLOS), 2020.

[16] H. Jang, J. Kim, J.-E. Jo, J. Lee, and J. Kim, “MnnFast: A Fast and Scalable System Architecture for Memory-Augmented Neural Networks,” in Proceedings of the International Symposium on Computer Architecture (ISCA), 2019.

[17] U. Gupta, C.-J. Wu, X. Wang, M. Naumov, B. Reagen, D. Brooks, B. Cottel, K. Hazelwood, M. Hempstead, B. Jia, H.-H. S. Lee, A. Malevich, D. Mudigere, M. Smelyanskiy, L. Xiong, and X. Zhang, “The Architectural Implications of Facebook’s DNN-based Personalized Recommendation,” in Proceedings of the International Symposium on HighPerformance Computer Architecture (HPCA), 2020.

[18] J. Dean, D. Patterson, and C. Young, “A New Golden Age in Computer Architecture: Empowering the Machine-Learning Revolution,” in IEEE Micro, 2018.

[19] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T. Chua, “Neural Collaborative Filtering,” in Proceedings of the International Conference on World Wide Web (WWW), 2017.

[20] M. Campo, C.-K. Hsieh, M. Nickens, J. Espinoza, A. Taliyan, J. Rieger, J. Ho, and B. Sherick, "Competitive Analysis System for Theatrical Movie Releases Based on Movie Trailer Deep Video Representation," in arxiv.org, 2018.