10-3-2024

Big Data Analysis Applied to the Retail Sector

Luis Barata
Instituto Politécnico de Castelo Branco, luis.barata@ipcb.pt

Gonçalo Rosa
Instituto Politécnico de Castelo Branco, goncalo.rosa@ipcbcampus.pt

Tiago Martins
Instituto Politécnico de Castelo Branco, tiagomartins1@ipcbcampus.pt

Pedro Silva
Instituto Politécnico de Castelo Branco, pt@ipcb.pt

Eurico Lopes
Instituto Politécnico de Castelo Branco, eurico@ipcb.pt

BIG DATA ANALYSIS APPLIED TO THE RETAIL SECTOR

Research full-length paper.

Barata, Luís M., Instituto Politécnico de Castelo Branco, Castelo Branco, Portugal, luis.barata@ipcb.pt

Rosa, Gonçalo, Instituto Politécnico de Castelo Branco, Castelo Branco, Portugal, goncalo.rosa@ipcbcampus.pt

Martins, Tiago, Instituto Politécnico de Castelo Branco, Castelo Branco, Portugal, tiagomartins1@ipcbcampus.pt

Silva, Pedro T., Instituto Politécnico de Castelo Branco, Castelo Branco, Portugal, pt@ipcb.pt

Lopes, Eurico, Instituto Politécnico de Castelo Branco, Castelo Branco, Portugal, eurico@ipcb.pt

Abstract

Big Data and Artificial Intelligence are rapidly evolving and transforming the world and e-commerce. To explore the most used techniques for building personalized recommendation systems and inventory management, a systematic review was conducted. It was found that hybrid Machine Learning techniques and the Hadoop system are widely used in these areas, delivering satisfactory results while maintaining scalable and efficient systems.

Keywords: Big Data; Data Analytics; Personalization of Recommendations; Stock Management.

1 Introduction

The digital economy can be characterized as business conducted via the Internet (Riley, 2023). With the great proliferation of AI (Artificial Intelligence) systems, sectors such as retail, healthcare, and education benefit from these technologies. The field of AI is extremely vast, encompassing areas such as ML (Machine Learning), where mathematical models are optimized to extract patterns from data; computer vision, which allows the machine to have spatial knowledge; and NLP (Natural Language Processing), which allows the machine to analyse and generate cohesive text. Furthermore, the analysis of Big Data permits the examination of vast quantities of data from prominent technological entities that are situated within the digital economy. E-commerce is managed by these algorithms, from personalized chatbots to systems that help with inventory management. With the growth of the digital market, the complexity of managing the business also increases, with worldwide deliveries and an extensive variety of products. Planning inventory becomes more difficult, and it is necessary to predict which products are needed, and where and when they need to be sent. It is also possible to reduce costs by allowing AI to analyse internal processes and correcting errors, and defects. Furthermore, as explained by (Campbell et al., 2020), AI can be used to understand customer behaviour and, in a way, predict what may happen, based on facial expressions, body gestures, voice, eye movement, and heart rate, being able to adapt service and what is shown online using virtual data created by the user.

Pappas et al., 2023, showed that AI can be very efficient in the defensive part of sales, that is, providing support and analysis, keeping humans in the offensive part, and serving and interacting with customers.

Table 1 shows the systematic reviews conducted over the years and, although this review is related to the others in the table, the specific objective of this review is to find the main techniques for recommendation systems and inventory management. Therefore, the next section will show how the articles were selected, while sections 3 and 4 will summarize and discuss the results of the articles, respectively.

REF	YEAR	TYPE	DESCRIPTION	THEME
(Ko et al., 2022)	2022	Literature analysis	The article analyses trends in recommendation systems.	Recommendation systems
(Govindan et al., 2022)	2022	Systematic Analysis	This article conducts a systematic review to identify specific Industry 4.0 technologies that improve various supply chain measures.	Supply chain management
(Tan & Sidhu, 2022)	2022	Systematic Analysis	IoT applications in the supply chain to understand current trends and future developments.	Supply chain management
(Fildes et al. , 2022)	2022	Literature Analysis	The article reviews the literature on retail demand forecasting, addressing strategic and operational challenges.	Supply chain management
(Chew et al., 2020)	2020	Literature analysis	The article explores the role of recommender systems in managing the wide range of online options, emphasizing the need for effective filtering to meet users' needs.	Recommendation systems
(Isinkaye et al., 2015)	2015	Literature analysis	The article aims to address information overload on the Internet and how recommendation systems can help by offering personalized content.	Recommendation systems
(Sivapalan et al., 2014)	2014	Literature analysis	The article addresses the impact of the Internet on daily tasks, highlighting the need for customization of products and services due to the diversity of customers.	Recommendation systems
(Nabizadeh et al., 2013)	2013	Literature analysis	The article highlights the importance of recommendation systems for personalizing suggestions amid online information overloads.	Stock management
(Corsten & Gruen, 2005)	2005	Systematic Analysis	The article highlights stock shortages in the FMCG (Fast- moving) industry. consumer goods), highlighting the importance of product availability.	Stock management
(Corsten & Gruen, 2003	2003	Systematic Analysis	The article challenges the perception of reduced stock-outs in retail, highlighting retailers' persistent struggle with this problem.	Stock management

Table 1 – Related work

2 Methodology

To conduct the systematic review, the PRISMA methodology was applied (Page et al., 2021) to ensure quality and a standardized structure for this systematic review. In this way, some questions were formulated that we wish to answer, aiming to achieve the objectives previously mentioned in Section 1. The questions are:

- What are the most used techniques for recommending products?
- What techniques are employed for inventory management, with a particular focus on machine learning algorithms and pipelines that are designed to handle large volumes of data?

2.1 Search formula

To obtain articles relevant to the designated objectives and questions, it was decided to use three reliable databases, namely ScienceDirect, IEEE, and ACM. The search equation was built to capture Big Data articles with a focus on inventory management and recommendation systems, ensuring that all these topics are mentioned in the article titles and/or abstract. A final equation is as follows: ("Big Data" OR "Data Analytics") AND ("Retail" OR "Retail Sector") AND ("Product Recommendations" OR "Personalization") AND ("Inventory Management" OR "Stock Optimization”) AND "algorithms". With this equation, 153 articles were obtained, without any duplicate articles.

2.2 Inclusion and exclusion criteria

As illustrated in Figure 1, the initial step in selecting articles is to evaluate the titles and abstracts to identify those that align with the research objectives. In order to ascertain the articles that would be suitable for inclusion in the study, a set of inclusion criteria was defined. These criteria included the presence of recommendations and/or personalisation of products or services, or the topic of inventory management. Consequently, any article that contained or conveyed the idea that at least one of the aforementioned topics was explored in the article was retained for further consideration. Conversely, any article that did not meet these criteria was excluded from the study, resulting in the exclusion of 106 articles.

Of the 47 articles that remained, some exclusion criteria were used, such as: being written in a language other than English or Portuguese, not being able to obtain a free copy, and having an informative focus on general areas. Consequently, 33 articles were excluded, leaving only 14 articles suitable for review.

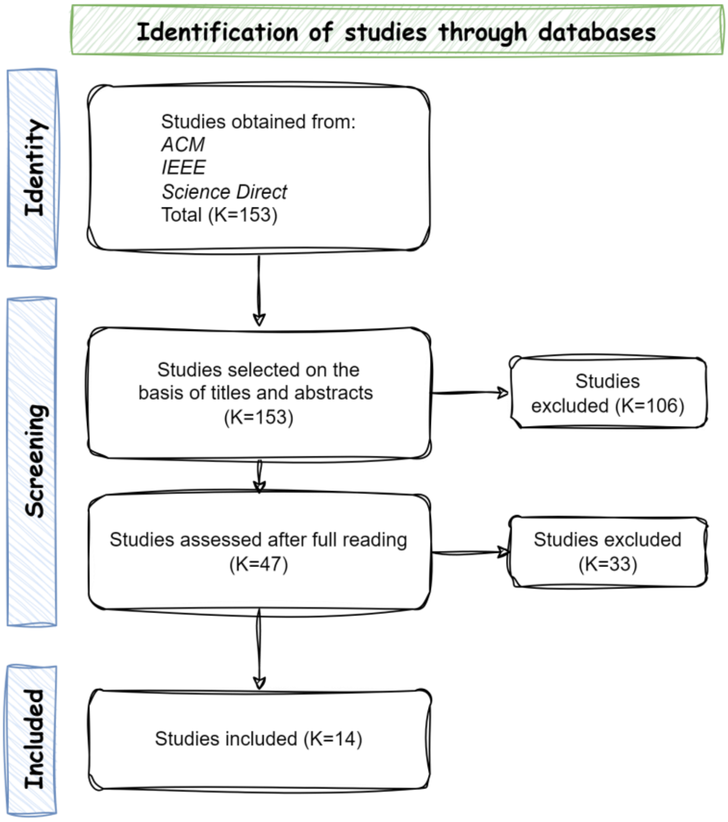

Figure 1 – Article selection diagram based on (Page et al., 2021).

3 Results

All selected articles are represented in Table 2, which contains the year of publication, the database where the article was found, as well as the techniques used in the type of problem that was inferred after reading. It can be observed that many articles were obtained from the ScienceDirect database, followed by IEEE and finally ACM. The articles were ordered descending by year.

REF	YEAR	TYPE	TECHNIQUES	DATABASE
(Hameed et al, 2023)	2023	Stock management	Apriori	IEEE
Ding et al, 2023	2023	Stock management and recommendation	SD (System Dynamics), AB (Agent-Based Modelling)	ScienceDirect
(Dutta et al., 2023)	2023	Stock management and recommendation systems	K- Means, Logistic Regression, and EOQ (Economic Order Quantity)	IEEE
(Taj et al., 2023)	2023	Stock management	RFID (Radio Frequency Identification), Blockchain	IEEE
(Sulova et al., 2022)	2022	Stock management	Hadoop, SEAM (Stacked Ensemble All Models)	ACM
(Tan & Sidhu, 2022)	2022	Stock management	RFID (Radio Frequency Identification), EOQ (Economic Order Quantity)	ScienceDirect

(Raga et al., 2021)	2021	Recommendation Systems	K-NN	IEEE
(Pereira & Frazzon, 2021)	2021	Stock management	K- Means, NARX (autoregressive model)	ScienceDirect
(Malhotra & Rishi, 2021)	2021	Recommendation Systems	HDFS, MapReduce	ScienceDirect
(He, Xue, & Gu, 2020)	2020	Stock management	Smart Connected Products	ScienceDirect
(Tsagkias et al., 2021)	2021	Recommendation Systems	LtM (Learning to Match), LtR (Learning to Rank)	ACM
Marchand & Marx, 2020)	2020	Recommendation Systems	-	ScienceDirect
(Bag et al., 2019)	2019	Recommendation Systems	Linear regressions, Statistical analysis	ScienceDirect
(Fu et al., 2017)	2017	Recommendation Systems	FS (Fuzzy System), HFS Hierarchical Fuzzy System	ScienceDirect

Table 2 – Information about the articles found.

Figure 2 shows that most selected articles were published in the years 2021 and 2023, followed by 2020 and 2022. Possibly because the amount of data and the need for automation of multiple processes is continuously growing rapidly.

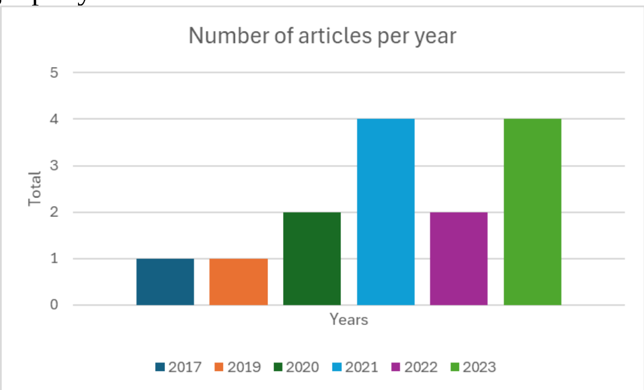

Figure 2 – Count of articles per year.

Figure 3 shows that from 2017 to 2021, there is always an article about recommendation systems, as these are common systems that need to be continually reinvented. Stock systems, more specifically, inventory management, contain articles from 2020 to 2023, and some articles address both topics, published in 2023.

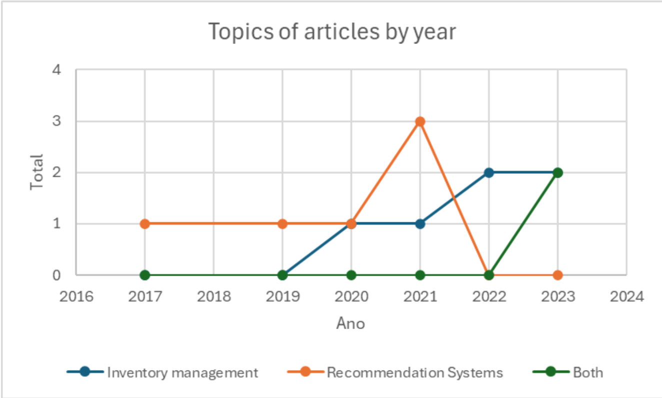

Figure 3 – Count of themes per year.

The number of articles by type of systems are divided equally as can be seen in Figure 4, ensuring that the review does not prioritize and explore one type more than the other.

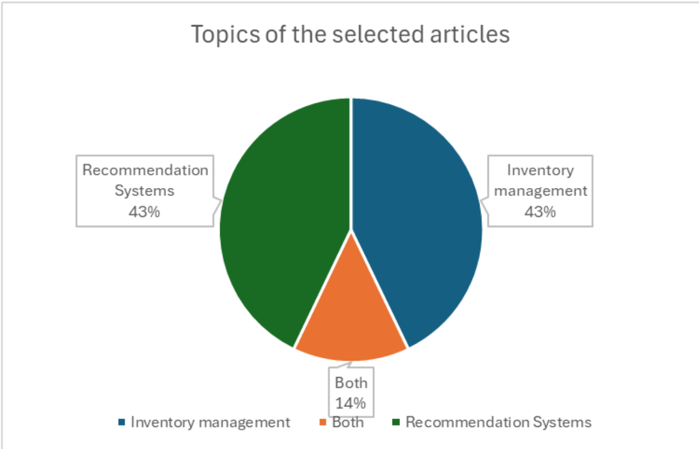

Figure 4 – Count of themes per year.

All articles were duly summarized and grouped by type so that each type of system could be analysed continuously. However, articles that address both types were grouped separately. Within each type, the articles were also ordered descending by year.

3.1 Recommendation Systems

In (Raga et al., 2021), it was found that local businesses do not have the resources to fight against large commercial areas, since they do not have the technological advantage that they have. This frequently results in customers being compelled to await the availability of the desired product, or alternatively, to seek the product elsewhere. To combat this problem, algorithms can be used, such as K-NN (K-Nearest Neighbours), to group stores by score, that is, based on a customer. Subsequently, it is checked whether the product exists in that store and, if so, the average price of the products is checked, and a score is assigned to that establishment. This prevents the customer from having to go to an establishment without knowing whether the product exists and whether the price of that product is the cheapest among the various commercial areas.

As for (Malhotra & Rishi, 2021), the authors explore the limitations present in current product and page search systems rank in the E-Commerce environment. The main objective of this is to improve this search by analysing the preferences of a given user and their search behaviour. The authors make use of HDFS (Hadoop Distributed File System), MapReduce, and Relevance Vector Machine. This allows users to “equip” with a robust ECMRS system (E-Commerce Metasearch and Ranking System).

The authors (Tsagkias et al., 2021), address the challenges of search and recommendation systems in e-commerce, which are essential for improving user experience, increasing sales, and promoting customer loyalty in the highly competitive online commerce sector. One of the main challenges highlighted is the management of extensive and varied product catalogues. Organizing and presenting products on e-commerce platforms can present challenges due to the extensive range of categories. Furthermore, this article addresses the issue of user intent ambiguity, where users may vaguely express their preferences, making it difficult to understand and satisfy their needs. The importance of contextual information is also highlighted to improve the relevance and personalization of search results and recommendations. Factors such as user location, device type, browsing history, and temporal trends can have a substantial impact on user preferences and purchasing decisions. Therefore, it is imperative to integrate contextual signals into recommendation algorithms. The article explores the ethical considerations surrounding e-commerce search and recommendation systems, including concerns about privacy, transparency, and algorithmic bias. It highlights the importance of developing recommendation systems that are fair and responsible, respect users' privacy rights, and avoid perpetuating discriminatory results. The author proposes exploring advanced techniques, such as NLP , ML, and DL (Deep Learning), to improve understanding of user intent and increase the relevance of search and recommendation results.

In (Marchand & Marx, 2020), the authors developed a framework for personalized recommendations. It starts with a base model that classifies items based on user interaction with each item. However, as some users tend to implicitly give high ratings to items, as well as certain items tend to be more visited because they are popular, it is necessary to consider these unusual interactions, the model will also consider these situations. Users' tastes also change over time, and as the model is based on time series, these changes in tastes must be modelled, thus adding a temporal variable, which is dynamic and not static, as its value will change over time. Simple linear regressions optimized by the least squares method are used to estimate each parameter, due to lack of data. However, the user must have classifications that are at least 120 days old for the model to estimate correctly. Regressions are also used to ensure that the classification considers other non-static effects. As there are users who do not rely on the product's characteristics to interact with it, the regressive model will not work. To do this, a collaborative filter is used to check similarities between the user's old ratings and items they have never rated, allowing their ratings to be estimated. Thus, a hybrid model was formulated, which decides which of the two algorithms to use, based on the difference in errors.

In (Bag et al., 2019), it is found that consumers' purchasing intentions play a crucial role in determining the success of durable goods in the market. Understanding the factors that influence these intentions is essential for companies to effectively guide their marketing strategies. The authors performed statistical analyses, including regression analyses and factor analyses, to identify the main attributes that significantly influence consumers' purchasing intentions. Regression analysis helped quantify the relationship between each attribute and purchase intentions, while factor analysis helped reduce the dimensionality of the data and identify underlying patterns.

In (Fu et al., 2017), it is demonstrated a FS (Fuzzy System), a system that creates rules considering uncertain data, allowing an output to be estimated based on this data. Typically, Fuzzy sets require intervention from experts to modify and correct failures due to uncertainties. In the article, an FS was created to negotiate prices that are estimated through transaction history. The grouping of inputs is done manually by experts, as clustering algorithms tend to group irrelevant attributes, affecting the intermediate variables of the HFS (Hierarchical Fuzzy System). The standard FS uses membership functions to transform inputs into their respective membership value. This is because we work with fuzzy sets, sets that work with uncertain data, and data may or may not belong to the set depending on the interpretation and data provided. To optimize these values, a recursive least squares method is used. SFS-SISOM (FS Single Input and a Single Output Module) uses weights for each input to help reduce the size of the data, or in this case, its importance. HFS uses several standard FS, where the output of some are the intermediate variables of the FS higher up in the hierarchy. To optimize the parameters, present in the different FS, it is necessary to minimize a cost function that calculates the price estimate error taking the real value and back-propagating it along the hierarchical layers.

3.2 Stock management

In (Hameed et al., 2023), the authors used the Apriori algorithm to find frequent items and, subsequently, association rules in a transaction dataset from a gift shop in the United Kingdom. This allowed for better management of the store's inventory and to place promotions on items based on association rules, that is, if they bought an item, they must buy another item, so that item is put on sale. It was found that the Apriori algorithm has better performance and presents more accurate results than other algorithms, especially in small datasets. The metrics used to select frequent items are normally the support value, which is the ratio between the appearances of a given item in transactions and the total number of transactions. When using these frequent items to obtain association rules, confidence is used, this being the ratio between the support of an association and the support of the association's antecedent, thus representing the strength of the association, that is, having the antecedent then follows the consequent. This algorithm examines several sets of items and checks whether they are frequent, based on a certain support threshold. If it is not, it will be discarded, leaving only the frequent items. The results were presented, but the association rules used Lift as a metric. This measures the probability of both the preceding and consequent being in the same transaction. If the value is greater than one, it means there is a high probability.

In (Taj et al., 2023), based on the systematic review, some techniques are suggested that can help with stock management. The use of RFID (Radio Frequency Identification) for tracking products, expiration dates, and identifying their category, in addition to allowing the integration of products into a blockchain, ensures security and financial risk control. Cloud computing technologies can also be used for inventory management, ensuring centralized databases that record information about inventory-level operations, processing data from sensors, and replicating these databases to ensure information security. At a sensory level, gyroscope sensors can be used to track the location of the product, proximity sensors to know when a product has been labelled or unlabelled, humidity, and luminosity.

In (Sulova et al., 2022), the authors describe a framework for processing large volumes of data, which are used in predictive models that help minimize the delivery time of a product. Thus, a four-stage framework was constructed. The first involves creating a business plan, thus having a plan and objectives outlined for predictive analysis. The second stage refers to the identification of data sources and their processing for the predictive models. The third stage is the use of different predictive algorithms and verification of the results obtained, choosing the algorithms that provide quality knowledge. Finally, the last stage is applying the models to production and identifying possible corrections that should be applied to them. H20 is an open-source platform that allows the use of the most varied Machine Learning Algorithm(s), allowing its integration with Hadoop systems.

In (Tan & Sidhu, 2022), some of the techniques used for inventory management can be seen. The use of RFID to identify and track IoT (Internet of Things) to communicate and exchange data. Another technique is the implementation of warehouse management systems, making it possible to find the optimal configuration where each product should be delivered, considering the quantity, warehouse capacity, and where it is on the delivery route. The use of removable shelving technologies to divide storage space between types of products, thus optimizing the space used, is also an explored technique. Real-time sensors can also be used within the warehouse to collect information about products, track the location and status of products, and analyse inventory parameters such as minimum stock level, reordering waiting time, and EOQ (Economic Order Quantity). Finally, use predictive algorithms to estimate which products are in greatest demand, thus allowing you to manage product inventory based on estimates.

In (Pereira & Frazzon, 2021), the authors propose to present a way, focused on data, of connecting the digital part with the physical part in terms of retail supply management. Using omni-channel management, that is, bringing the two worlds together can be a complementary force, attracting more customers, optimizing inventory management, and accelerating distribution. Methodologies such as genetic algorithms and swarm intelligence were used to try to minimize costs and optimize deliveries considering users' uncertainty. However, there are several types of information available, something that these algorithms do not normally use and/or optimize. The proposed algorithm has two stages. In the first stage, there is a clustering algorithm which groups users by behaviour and a neural network, estimating demand considering the grouping made. This neural network receives a time series of the sales history of each product in each channel and returns the forecast of demand for each time series. In the second stage, the estimated time series are received along with inventory information, which is fed to a simulation algorithm that uses genetic algorithms and hybrid simulation methods. The genetic algorithm will return the supply distribution plan and the simulation will use this planning and other information to calculate the cost considering all channels. This value goes back to the genetic algorithm and the process is repeated until the stopping criterion is reached. The final output will be supplying distribution planning, as well as demand estimation for each product.

In (He et al., 2020), technological advances that have been producing profound changes in supply chains are addressed, using IoT devices, as these allow the creation of ways of functioning, partially based on the analysis of a large volume of data. In a supply chain environment, IoT devices can be used in a way that allows the consumer to play a fundamental role in decision-making by those responsible. This process is achieved through the use of SCPs (smart connected products), which are situated between the aforementioned entities. A product may, in fact, serve this function. Compared to traditional products, SCPs can provide new features, greater reliability, greater product usage, and overcome the limits of traditional products. In many SCPs, software replaces physical accessories or allows a single physical device to operate under different conditions. The connection between products can have two functions: allowing the exchange of information between products, manufacturers, operating systems, and users; and creating digital functions independent of physical products.

3.3 Recommendation and Stock Management Systems

In (Ding et al., 2023), topics such as IoT, promise to revolutionize supply chain operations, enabling real-time monitoring, data-based decision making, and greater efficiency. However, realizing the full potential of IoT in promoting sustainability in supply chains requires a comprehensive understanding of its impacts and interactions. Simulations were carried out within the scope of ABM (Agent-Based Modelling), where different scenarios were evaluated to assess the impact of IoT implementation on sustainability KPIs (Key Performance Indicators), such as carbon emissions, resource use, and operational efficiency. By manipulating parameters such as IoT penetration rate, data exchange protocols, and decision-making algorithms, we elucidate the complex dynamics that influence supply chain sustainability outcomes. The authors demonstrated the significant potential of IoT to improve sustainable supply chain performance across multiple dimensions. By monitoring and optimizing processes in real-time, IoT-based supply chains demonstrate a reduced environmental footprint, greater resource efficiency, and greater resilience to disruption. Additionally, the integration of IoT data streams enables greater visibility and coordination between supply chain partners, fostering collaboration and innovation towards sustainability goals.

In (Dutta et al., 2023), it is demonstrated that the use of algorithms such as K-Means, logistic regression, and collaborative filters can optimize price estimation and personalized recommendation systems, this being a customer perspective. Techniques such as the EOQ model are employed to optimize inventory management. This involves calculating the optimal quantity of products, from the perspective of the seller.

4 Discussion

After reviewing the articles, clustering algorithms such as K-NN and K- Means are used to group similar customers and/or products to recommend products, considering all the information. Regressions also proved to be important for product recommendations. They were used in (Bag et al., 2019) revealing that several attributes have a significant impact on consumers' intentions to purchase durable goods. Among these, price emerged as a critical factor, indicating that consumers are sensitive to the cost of durable goods when making purchasing decisions. Furthermore, product quality and brand reputation also strongly influence purchase intentions, highlighting the importance of perceived value and trust in consumer decision-making. Furthermore, certain product characteristics, such as durability and functionality, were identified as determining factors of purchase intentions, highlighting the importance of product attributes in shaping consumer preferences. Another use of regressions was presented in (Marchand & Marx, 2020), but they observed that for certain users, the regression estimates were completely wrong, which is why they also used collaborative filters, making the algorithm hybrid. The results show that the hybrid model outperformed all methods used, including the Netflix Prize-winning algorithm. Therefore, it was concluded that there should not be a simple recommendation algorithm, but rather a set of them that adapt to different groups of users. Considering these abrupt differences between user tastes that can overturn models based on simple regressions or even clustering algorithms, (Fu et al., 2017) built a model based on uncertainties, concluding that FS only works well in low dimensions, the SFS-SISOM can deal with greater dimensionality, but not in high non-linear dimensions, with HFS being more appropriate for this, even though it tends to fall into relative minima. HFS also has lower space complexity. With these algorithms it is possible to estimate prices and discounts, based on some inputs, considering that these may have uncertain values. Relying more on active product search (Tsagkias et al., 2021) suggests the use of NLP-based algorithms, but a more complete and distributed approach was presented (Malhotra & Rishi, 2021) using Hadoop. the authors also carried out a comparison between the various deployment systems, because creating a tool of this type requires consideration of several aspects such as partial failure support capabilities, fault tolerance, scaling, real-time processing, and execution efficiency. The systems compared were: HDFS, Spark, Peer-to-Peer, HPC Cluster, Multicore, GPU, and FPGA. They compared them, using a range of values, between 1 and 5, with 1 being the best option and 5 being the worst. After this comparison, the authors chose the HDFS platform due to its high scalability and fault tolerance, which are the two most important requirements in an E-Commerce environment. Therefore, for recommendation systems, it is necessary to use algorithms that consider the uncertainties and diversity of users, in addition to ensuring efficiency and easy scaling. Thus, using Hadoop, fuzzy algorithms, and hybrid algorithms can achieve good results.

In terms of inventory management, in (Hameed et al., 2023) the use of association rules allowed to show that some items in the store have associations, for example, whoever buys a set of tea glasses buys different colours, the same for lunch boxes. So, this algorithm, applied to small retail stores, can help understand customer behaviour and thus adjust inventory and prices based on associations between products. However, you must have a large data set from different years to increase the strength of the associations and find more relationships. The temporal analysis performed also allows planning the times at which promotions, previously orchestrated by association analysis, should be launched. In addition to these rules, other articles show that the use of RFID and sensors can help organize products in the warehouse and minimize shipping routes, saving time and costs. Already in (Sulova et al., 2022) proposed to study the risk of delay in delivery, estimating whether a delivery will arrive on time or late, using a binomial problem. The H20 AutoML method allows you to run multiple algorithms at the same time, choosing the ones that obtain the best results or using them in a large ensemble model. A SEAM-type ensemble (Stacked Ensemble All Models), powered by predictive algorithms from different families (Deep Learning, XGBoost, Random Forests, etc …), showed the best result, with an accuracy of 90%. A more complex method was presented in (Pereira & Frazzon, 2021), where the algorithm was applied to data obtained from a Brazilian retailer that operates 139 physical stores and sells home supplies and electronics online. The case study uses a four-stage omni-channel supply chain with physical stores, distribution centres, suppliers, and physical and online customers. The results showed that using information about estimated demand and sharing information allowed costs to be minimized. The use of the neural network also showed better results when it came to allocating products to the correct channel at the correct time. Thus, it was found that the use of this algorithm showed positive results for more efficient supply chain management, with fewer costs and fewer errors in estimating demand and supply. Already in (Dutta et al., 2023) the use of clustering algorithms, collaborative filters, and EOQ proved to be fast both in recommendation systems and in inventory management. Therefore, considering the ensembles and the complex model that uses neural networks and genetic algorithms, it can also be seen that hybrid models tend to have better results, as they adapt to the uncertainty of the enormous amount of data from most different users. However, it must be considered that faster algorithms can sometimes obtain satisfactory results.

5 Conclusion

After the systematic review, it can be concluded that the first approach used both in customizing recommendation algorithms and inventory management is the use of simple models that make some measurements in several dimensions and return estimates. However, to make the experience more robust and personalized, it is necessary to consider hybrid models that take uncertainties and diversity into account. In addition, there is a need for systems that allow scaling and efficiency in the search for information in the era of Big Data. Specifically for inventory management, it was found that RFID is inevitable for controlling and improving the management and location of products. In this way, it was possible to answer both questions that supported this systematic review.

References

Bag, S., Tiwari, M. K., & Chan, F. T. S. (2019). Predicting the consumer's purchase intention of durable goods: An attribute-level analysis. Journal of Business Research, 94, 408-419. Retrieved from https://www.sciencedirect.com/science/article/pii/S0148296317304770 doi : https://doi.org/https://doi.org/10.1016/j.jbusres.2017.11.031

Campbell, C., Sands, S., Ferraro, C., Tsao, H.-YJ, & Mavrommatis, A. (2020). From data to action: How marketers can leverage AI. Business Horizons, 63, 227-243. Retrieved from https://www.sciencedirect.com/science/article/pii/S0007681319301624 doi : https://doi.org/https://doi.org/10.1016/j.bushor.2019.12.002

Chew , L.-J., Haw , S.-C., & Subramaniam , S. (2020, 6). Recommender system for retail domain. In (pp. 9-13). ACM. Retrieved from https://dl.acm.org/doi/10.1145/3408066.3408101 doi : https://doi.org/10.1145/3408066.3408101

Corsten, D., & Gruen, T. (2003, 12). Desperately seeking shelf availability: an examination of the extent, the causes, and the efforts to address retail out-of-stocks. International Journal of Retail & Distribution Management, 31, 605-617. Retrieved from https://www.emerald.com/insight/content/doi/10.1108/09590550310507731/full/html doi : https://doi.org/10.1108/09590550310507731

Corsten, D., & Gruen, T. (2005, 12). On shelf availability: An examination of the extent, the causes, and the efforts to address retail out-of-stocks. In (pp. 131-149). Springer-Verlag. Retrieved from http://link.springer.com/10.1007/3-540-27059-09 doi : https://doi.org/10.1007/3-540-27059-09

Ding, S., Ward, H., Cucurachi , S., & Tukker, A. (2023). Revealing the hidden potentials of internet of things (iot) - an integrated approach using agent-based modeling and system dynamics to assess sustainable supply chain performance. Journal of Cleaner Production, 421, 138558. Retrieved from https://www.sciencedirect.com/science/article/pii/S0959652623027166 doi : https://doi.org/https://doi.org/10.1016/j. jclepro.2023.138558

Dutta, A., Yadav, R. K., & G, M. (2023, 11). The convergence of modern digital solutions with online retail offers a comprehensive examination from seller and buyer perspectives. In (p. 1-6). IEEE. Retrieved from https://ieeexplore.ieee.org/document/10419945/ doi : https://doi.org/10.1109/ICRASET59632.2023.10419945

Fildes , R., Ma, S., & Kolassa , S. (2022, 10). Retail forecasting: Research and practice. International Journal of Forecasting, 38, 1283-1318. doi : https://doi.org/10.1016/J.IJFORECAST.2019.06.004

Fu, X., Zeng, X.-J., Luo, X., Wang, D., Xu, D., & Fan, Q.-L. (2017). Designing an intelligent decision support system for effective negotiation pricing: A systematic and learning approach. Decision Support Systems, 96, 49-66. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167923617300192 doi : https://doi.org/https://doi.org/10.1016/j.dss.2017.02.003

Govindan, K., Kannan, D., Jørgensen, T. B., & Nielsen, T. S. (2022, 8). Supply chain 4.0 performance measurement: A systematic literature review, framework development, and empirical evidence. Transportation Research Part E: Logistics and Transportation Review, 164, 102725. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S1366554522001168 doi : https://doi.org/10.1016/j.tre.2022.102725

Hameed, VA, Rana, ME, & Enn, LH (2023, 7). Apriori algorithm based association rule mining to enhance small-scale retailer sales. In (pp. 187-191). IEEE. Retrieved from https://ieeexplore.ieee.org/document/10256952/ doi : https://doi.org/10.1109/BDAI59165.2023.10256952

He , L., Xue , M., & Gu , B. (2020). Internet-of-things enabled supply chain planning and coordination with big data services: Certain theoretical implications. Journal of Management Science and Engineering, 5 , 1-22. Retrieved from https://www.sciencedirect.com/science/article/pii/S2096232020300172 doi : https://doi.org/https://doi.org/10.1016/j.jmse.2020.03.002

Isinkaye, FO, Folajimi, YO, & Ojokoh , BA (2015, 11). Recommendation systems: Principles, methods and evaluation. Egyptian Informatics Journal, 16, 261-273. doi : https://doi.org/10.1016/J.EIJ.2015.06.005

Ko, H., Lee, S., Park, Y., & Choi, A. (2022, 1). A survey of recommendation systems: Recommendation models, techniques, and application fields. Electronics 2022 , Vol . /11/1/141 doi : https://doi.org/10.3390/ELECTRONICS11010141

Malhotra, D., & Rishi, O. (2021). An intelligent approach to design of e-commerce metasearch and ranking system using next-generation big data analytics. Journal of King Saud University - Computer and Information Sciences, 33, 183-194. Retrieved from https://www.sciencedirect.com/science/article/pii/S1319157817303440 doi : https://doi.org/https://doi.org/10.1016/j.jksuci.2018.02.015

Marchand, A., & Marx, P. (2020). Automated product recommendations with preference-based explanations. Journal of Retailing, 96, 328-343. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022435920300014 (¡ b¿ PAGO ?¡ /b¿) doi : https://doi.org/https://doi.org/10.1016/ j.jretai.2020.01.001

Nabizadeh , A.H., Salim , N., Aghdam , A.R., Rafsanjani , AHN, & Fard , K.B. (2013, 5). Recommendation systems: a review. Retrieved from https://www.researchgate.net/publication/325074466

Page, MJ, McKenzie, JE, Bossuyt, PM, Boutron , I., Hoffmann, TC, Mulrow, CD, . . . Moher, D. (2021, 3). The prism 2020 statement: an updated guideline for reporting systematic reviews. BMJ, 372. Retrieved from https://www.bmj.com/content/372/bmj.n71 https://www.bmj.com/content/372/bmj.n71.abstract doi : https://doi.org/10.1136/ BMJ.N71

Pappas , A., Fumagalli , E., Rouziou , M., & Bolander , W. (2023). More than machines: The role of the future retail salesperson in enhancing the customer experience. Journal of Retailing, 99, 518-531. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022435923000490 doi : https://doi.org/https://doi.org/10.1016/j.jretai.2023.10.004

Pereira, MM, & Frazzon , EM (2021). A data-driven approach to adaptive synchronization of demand and supply in omni-channel retail supply chains. International Journal of Information Management, 57, 102165. Retrieved from https://www.sciencedirect.com/science/article/pii/S026840122030205X doi : https://doi.org/https://doi.org/10.1016/j .ijinfomgt.2020.102165

Raga, A., Nag, A., Reddy, BR, Shetty, DD, Prashanth, BS, Kumar, MVM, & Darshan, SLS (2021, 12). Goget - a digital shop recommendation system to empower retailers using machine learning. In (Vol. 1, p. 1-7). IEEE. Retrieved from https://ieeexplore.ieee.org/document/9702606/ doi : https://doi.org/10.1109/FABS52071.2021.9702606

Riley, J. (2023, 12). Ai powers the digital economy. Ubiquity, 2023 . Retrieved from https://doi.org/10.1145/3636478 doi : https://doi.org/10.1145/3636478

Sivapalan, S., Sadeghian, A., Rahnama, H., & Madni, A. M. (2014, 8). Recommending systems in e-commerce. In (pp. 179-184). IEEE. Retrieved from http://ieeexplore.ieee.org/document/6935763/ doi : https://doi.org/10.1109/WAC.2014.6935763

Sulova , S., Aleksandrova , Y., Stoyanova , M., & Radev , M. (2022, 6). A predictive analytics framework using machine learning for the logistics industry. In (p. 39-44). ACM. Retrieved from https://dl.acm.org/doi/10.1145/3546118.3546130 doi : https://doi.org/10.1145/3546118.3546130

Taj, S., Imran, A.S., Kastrati, Z., Daudpota, S.M., Memon, R.A., & Ahmed, J. (2023, 12). Iot -based supply chain management: A systematic literature review. Internet of Things, 24, 100982. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S2542660523003050 doi : https://doi.org/10.1016/j.iot.2023.100982

Tan, W. C., & Sidhu, M. S. (2022). Review of rfid and iot integration in supply chain management. Operations Research Perspectives, 9, 100229. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S2214716022000070 doi : https://doi.org/10.1016/j.orp.2022.100229

Tsagkias , M., King, T. H., Kallumadi , S., Murdock, V., & de Rijke, M. (2021, 2). Challenges and research opportunities in ecommerce search and recommendations. SIGIR Forum, 54 . Retrieved from https://doi.org/10.1145/3451964.3451966 doi : https://doi.org/10.1145/3451964.3451966