TECHNIUM
SOCIAL SCIENCES JOURNAL

www.techniumscience.com

Vol. 68/2025

A New Decade for Social Changes

PLUS COMMUNICATION

International Communication & PR

Technium Social Sciences Journal
Vol. 68, 350-372, February, 2025
ISSN: 2668-7798
www.techniumscience.com

The Impact of Data Mining in Marketing: Transforming Consumer Insights into Business Success

Constantinos Halkiopoulos

Department of Management Science and Technology, University of Patras, Greece

halkion@upatras.gr

Abstract. This review paper discusses how data mining can revolutionize the way marketing strategies are drawn and business is conducted by offering comprehensive consumer insights. Data mining, through techniques such as association rule mining, clustering, and classification, helps businesses understand consumer behavior, forecast sales, optimize campaigns, and improve customer retention. It points out the emphasis on customer segmentation, predictive analytics, and the ethical considerations in data collection and usage, such as private information and AI or machine learning algorithms that bear the imprint of bias. The paper also highlights future trends and advancements in the technology of data mining, with a great emphasis on personalized marketing using a data-driven approach. The paper further stresses the need for ethical applications and careful handling of consumer data to increase trust and build brand loyalty. Overall, this research positions data mining as an essential tool in developing actionable business strategies, enhancing profitability, and reducing costs in the digital marketing landscape.

Keywords. Data Mining, Marketing Strategies, Consumer Insights, Association Rule Mining, Customer Segmentation, Predictive Analytics, Campaign Optimization, Customer Retention, Artificial Intelligence (AI), Machine Learning (ML), Ethical Considerations, Data Privacy

1. Introduction

The advent of big data in the twenty-first century has made data mining a central focus for marketing strategies across a broad range of industries. The insights that data mining provides are directly relevant to our understanding of consumers. Both in the physical and digital domains, data mining is a critical tool for understanding consumer behavior. Using data mining, businesses can not only identify who is purchasing which products and services, but also understand why they are doing so. The use of data mining in marketing can unlock vital information about what consumers require, enabling businesses to more accurately forecast their sales revenues and the resources required to meet customer needs effectively [1-7]. The evolution of data mining technologies over the past half century has seen the creation of major advances from developments in string search algorithms to enable the processing of complex data through to the automated capture of immense data sets through consumer activity. However, the underlying aim of data mining has remained evidently connected to the necessity of businesses to forecast the future using past events and changing conditions in both

Technium Social Sciences Journal
Vol. 68, 350-372, February, 2025
ISSN: 2668-7798
www.techniumscience.com

environments and consumer trends. Changes in market dynamics frequently create a chasm that can limit the success of both startups and experienced organizations. Whether arising from changing customer demand, technological advances, economic trends, or regulatory alterations, being able to make rapid decisions in this context can separate successful organizations from unsuccessful ones [8-15].

Data mining has incredible transformative potential based on the ability to make decisions and drive processes with such immediacy. Some of the data collected can be mundane or unuseful, but with the right market knowledge, data on consumer movements, what they choose to watch on television, and how many times they select a certain brand of tea can be considered an important cog in broader market strategies. Examples of data mining in marketing promotional campaigns for a new film or fashion brand can collect information from a range of sources, such as social media searches for target groups, and advertising investments and promotions from the leaders in that sub-sector. However, the growing integration of data mining only raises the question of privacy and the management of the public's personal data for commercial purposes [16-20].

Definition and Importance of Data Mining

Data mining is an advanced concept of marketing in which consumer insights are unveiled segment-wise. Furthermore, discovering interesting patterns is sought using various techniques within raw data by practicing data mining modeling. These interesting patterns put the decision makers who are directly concerned with marketing into thinking about new strategies for predicting or foreseeing the outcomes in current or future occurrences. Moreover, data mining indirectly becomes an exhaustive tool for crafting various information technology strategies, including but not limited to marketing. More deeply, data mining is expected to be employed as a means of generating creative definitions of consumer needs and their behavior by focusing on the four segment stages of micro, appropriate, personalized, and one-to-one marketing concepts. In sum, it can be recapitulated that data mining handles marketing views in unique issues, solutions, and decisions supported by various techniques that would not be available before a wide variety of data collection is made [21-28].

Advanced statistics computer software is needed to execute the techniques of data mining. Techniques such as decision trees, neural networks, market basket analysis, link analysis, and sequence or shopping cart analysis might be conducted to uncover valuable buried information. Another technique, known as cluster analysis, also unveils relevant information concerning the market segment or market basket. By using those tools, a long list of data can be uncovered and transformed into important knowledge of both problems and decisions. Indeed, this knowledge or information may represent interesting solution proposals to consider because they are framed based upon the situation in which consumers behave. Convincing marketers who are alert to employing those strategies have a larger number of loyal customers than their competitors who are unaware of those strategies. Ethically, the aspects of "where from" and "how" the data originates, as well as the "where" and "when" information is used, are concerns in data collection and data usage technology that are typically resistant to data mining and micro-marketing practices. Ethical convictions within data mining have become major debate issues and one of the research themes. Thus, given the unlocked potential of data mining, it is necessary to learn more advanced techniques and gain a better understanding of the data mining process and its advancements [29-35].

Technium Social Sciences Journal
Vol. 68, 350-372, February, 2025
ISSN: 2668-7798
www.techniumscience.com

2. Key Concepts and Techniques in Data Mining

Knowledge discovery in databases (KDD), or data mining, approaches large datasets in search of some regularity, some pattern, some form of insight that might not have been apparent otherwise. Data mining is, therefore, about knowledge generation, seeking to transform raw and often impenetrable data into useful management and marketing information. While data mining is often viewed as multi-disciplinary, marketing is probably the single area, from an applications standpoint, where data mining is most applied. At the heart of the KDD process are a series of techniques, which can be categorized into either descriptive or predictive modeling methodologies: 1. Descriptive techniques, such as segmentation, association analysis, and profiling, are designed to uncover patterns and relationships within data for the explicit purposes of exploring, understanding, and explaining why events and decisions occurred so that tangible or feasible actions can be taken in response. 2. Predictive modeling techniques, including artificial neural networks, decision trees, and regression analysis, allow us to make inferences about the future by testing and predicting the potential outcomes of different strategies, decisions, and events. These techniques have become increasingly important for identifying at-risk customers, planning targeted marketing strategies, and reducing advertising waste. The selection of an appropriate data mining technique, based on the marketing question being addressed, is paramount to the effectiveness of data mining. To assist practitioners who are often struggling with choosing the most appropriate technique for a given problem, the strengths and weaknesses of the techniques can be categorized and compared [36-40].

Association Rule Mining

Association rule mining is a widely used technique in data mining and has become a foundation of consumer purchase pattern analysis. The technique extracts meaningful patterns in a single scan of large datasets. These relationships are based on co-occurrence, and in the context of market basket analysis, can be used to determine the relative likelihood of one or more products being purchased. Association rule mining assumes a collection of tuples and creates rules using this data. For example, a large retailer could have a database of transactions, and a record of all products purchased. Association rule mining can analyze these records and generate a set of rules that indicate likely purchase behavior. A retailer could use items bought during one purchase cycle to identify items that may be bought together. This would enable the retailer to more accurately plan product assortments, plan and target promotions, and improve cross-sell and upsell items [41-45]. The set of association rules extracted could be large. However, filtering techniques can be applied, further reducing the list based on rule thresholds. Association rules use metrics to differentiate between significant rules. The three major metrics used are: Support is the frequency of item set occurrence in the transactions over all transactions - the fraction of the transactions that contain the specified item set. Confidence indicates how often a rule has been found to be true. It is the probability of occurrence of the consequent in the transactions when the antecedent occurs in the transactions already. Lift measures the significance of a rule for the items involved. It is used to positively or negatively assess the quality of rules. An increase in lift indicates that the antecedent and consequent have a strong relationship. Importantly for strategy, the use of the rule can strongly influence behavior. Analysis could also be made from the use of imported data from business use already. There could be association rules developed from initial insight data. This data could be resources reported on at different times to give insights into the production process. Ideas for variables and rules could be brought up or confirmed by the exploration of these two kinds of analyses [46-54].

Clustering

Clustering refers to the practice of identifying natural groupings of consumers, where individuals within the same group share some degree of similarity, while those in different groups are different with respect to the criteria within the clustering analysis. The grouping could be in terms of customer demographics, preferences, or more. The role of cluster analysis, typically found in marketing, is used for segmenting the consumer population into distinct and meaningful subgroups or segments. Once these segments are identified, firms are better able to match their marketing efforts directed at these segments. There are many clustering techniques available, such as K-means, Two-step clustering, Classification and Regression Trees, Hierarchical clustering, Model-based clustering, and Kohonen SOM [55-60]. Clustering is classified as unsupervised data mining methods. There are several clustering algorithms; K-means and Hierarchical clustering are the most used in decision-making in management and marketing studies. The K-means algorithm is a partitioning method. It is simple and mostly uses clustering technique with the shortest computation speed. Furthermore, Hierarchical clustering methods are the most popular types, describing an agglomerative or divisive form. Cluster analysis can be seen as a tool for understanding consumers, as well as the behavior consumers have about issues, products, or brands. To identify the importance of personalization, the clustering process is conducted to segment target consumers to have more potential. The clustering process can be conducted to demonstrate the behavior of consumers, so that it is possible to treat them differently and more personally than before. By clustering consumers based on their characteristics, marketing can be evaluated using relevant management, for example, better services, information service improvements for different segments, and assessing whether to launch a new product. With clustering, capacity can be accessed to increase the possibility of success in marketing campaigns. Marketers can analyze the heterogeneity of the samples both in the questionnaire and non-questionnaire base. Besides, the aggregation level of samples in the dataset can be determined. We can aggregate the characteristics or behavior of customers. On a lower level of aggregation, companies have a plethora of data to draw from clusters to draw a more reliable conclusion. Thus, a product can be more personalized to better match the personal needs and wants of consumers. Marketing should enable companies to engage and bring value to the customers [61-67].

Classification

Classification concerns categorizing data into predefined, discrete classes or labels based on features. The distinguishing characteristic of classification is that these classes or labels must already be known from historical data. Once a model has been built, the input data can then be separated into these classes. Classification has a wide variety of applications across various fields, and marketing is where this technique is perhaps the most heavily implemented. Many algorithms have been put forward over time to automate classification prediction and decision-making, with common examples including decision trees, neural networks, support vector machines, k-nearest neighbors, and random forests among many others. Whereas decisions were previously made based on some intuition, these automated algorithms allow companies to predict a specific outcome with a high degree of accuracy. In terms of marketing, if used properly, classification allows companies to take full advantage of their trend and effect discovery to separate the "lookers" from the "buyers" and the "retail bleeders" that are wasting financial resources, while also finding "high-value hidden niches." In a full marketing context, classification could be used to help with lead scoring and a CRM to assist with cross-selling and upselling for customer retention predictions [68-74]. Classifiers are built by applying algorithmic models to historical data, training the classifier on past marketing examples to predict an outcome. Once the classifier has been trained, identifying or targeting the desired audience can attract the necessary attention that the audience deserves, providing a high level of service for a potential long-term relationship. The built classifier can be used to sort out customers with probabilities of responses within the database and later score new prospect database lists based on a high prospect's reply to a very high prospect. Parameters might include targeting variables, comparable response history, probability-based scores, and more. Classifier construction and use, however, present substantial challenges. Classifier accuracy may be highly sensitive to the data used in training the algorithm. Therefore, ensuring high-quality training data is critical, with invalid data or poor data management causing inaccuracies and the potential for target audience separation as to the potential consumer who may take over. Classifiers are also amenable, so significant and ongoing big data knowledge within marketing is required to capture a database that accurately represents current consumer environments, not just past ones. This is a challenging development for companies engaging in digital markets and anyone outside retailing [75-81].

3. Applications of Data Mining in Marketing

In addition to identifying what data mining is, it is important to consider the actual applications of data mining in marketing and how these can be turned into business benefits. One such application is customer segmentation or market basket analysis, which uses data mining techniques to carry out market research by dividing customer bases into groups, so that it is easier for the company to identify their customers' needs and personalize its marketing strategy. In this framework, predictive analytics is used to uncover consumer trends by examining demographic and transaction-related data and, in general, to create statistical models capable of predicting consumers' future behavior. Other applications include campaign optimization techniques such as position-aware methods, which find the "best contact" customers in a database and maximize the profit of a multi-wave marketing campaign, and next-best-offer systems, which choose the best product to promote to a given customer, considering the cost of the promotion and the probability of acceptance [82-85]. Ultimately, similar requirements can be found in loyalty and churn management, which focuses on analyzing consumer behavior to develop approaches to keep them from leaving a service provider. Furthermore, churn prediction techniques foresee customer abandonment, allowing the company to intervene with a retention strategy in advance. Overall, a wide variety of processes have been found in specialized literature. These techniques also raise ethical implications, especially concerning people's privacy, as outlined by the negative sentiment that many customers show towards the collection and use of their data. It is also important to remark that the actual ethical use of data is a necessary assumption for the successful deployment of all these marketing campaigns. In conclusion, today's trend towards data-driven marketing is not surprising if one considers the capacity of data mining to transform simple consumer insights into business strategies for boosting profits and reducing costs [86-90].

Customer Segmentation

In today's digital landscape, the customer's journey is more varied and complex than ever. With consumers accessing multiple devices across various channels, marketers are met with the task of personalizing their campaigns to rise above the noise and connect with the right people. Data mining bridges the gap between consumer insights and effective business strategies. This section explores its application in customer segmentation [91-96]. Segmentation is one of the strategic practices that has benefited the most from data mining. This practice allows companies to organize their audiences into smaller groups that share certain characteristics. Segmentation can be based on different categories such as:

- Geography; - Demographics (age, sex, schooling, income, profession, social class, etc.); - Behavior (occasional buyers/users, loyalty, scope of use, etc.); - Purchasing patterns and consumption (products to win others, to consume together, to satisfy the same secondary needs, etc.); - Satisfaction (happy, disappointed customers, etc.).

Strategies for targeting either one of these can be used in conjunction with the others, which significantly increases the effectiveness of advertising or promotional messages directed to certain segments. Customer segmentation is an integral part of building brand loyalty because customers are more likely to be interested in the products of the company in this way. In line with the goals of the company, methods should also be employed to take customers from the target group to the next segment. Preferences cater to consumers based on segmentation, ensuring that satisfaction with the product is high [97-100]. Given the importance of customer segmentation, there has been a wide discussion on this topic by several scholars and practitioners in management science, sales, and marketing. Although there are many characteristics used for segmentation, adapting to the behavior of the user population is one of the essential criteria for segmenting. There are many methods and tools developed for customer segmentation using methods of data mining. Clustering techniques enable the mining of data and can be incorporated into segmentation analysis as a stage of data mining. Other procedures developed under this heading are classification and decision trees, focusing on the recognition of patterns that can justify the segmentation of the user base. Being a domain-driven approach, data mining can be used either to confirm the goodness of such criteria or to define entirely new segmentation approaches. Ethical considerations are considered during the development of such approaches [101-104].

Predictive Analytics

Predictive analytics is designed to forecast future events or consumer behavior. It can be used to predict future customer needs, as well as potential opportunities and threats, and to influence them. Predictive analytics utilizes data mining, including statistical techniques and machine learning models, to determine the likelihood of future outcomes based on historical data. Consequently, it relies not only on the accuracy of data collected but also on the ability to identify patterns. As a result, predictive analytics cannot be used in an organization where the use of new data proposing innovation from other units is looked down upon [105-108]. The use of predictive analytics offers evident benefits to businesses in many promotional, pricing, and operational areas, allowing for the anticipation of consumer behavior and future trends. The healthcare industry can make technology announcements based on ratings, estimates, projections, and customer feedback. Predictive analytics assists banks and credit card providers in detecting abnormal behavior that could indicate fraudulent activities. Further, various reports indicate significant growth in data-driven decision-making strategies following the implementation of predictive analytics. In real marketing practices, companies widely apply predictive analytics, such as lead scoring to develop the right application of direct marketing, customer credit forecasting, and personalized product or movie recommendations, among others. Nevertheless, the application of predictive models can be difficult when existing customer data collection is inaccurate, incomplete, or weak, or when the model is overfitted, meaning it does not generalize well to new data [109-111]. Challenges Analytical models forecast a future condition primarily based on historical knowledge. As a result, the reliability

Technium Social Sciences Journal
Vol. 68, 350-372, February, 2025
ISSN: 2668-7798
www.techniumscience.com

of the estimates is predicated on the validity of past occurrences to future arbitrarily similar ones. More substantial challenges to the usage of predictive analytics lie in reliability and accuracy, inappropriate data collection and reporting capability, business domain knowledge, and privacy and ethical implications regarding data, which will be discussed in the following sections [112-114].

Campaign Optimization

In traditional marketing, the marketer designs a marketing campaign, executes the campaign, and measures campaign performance metrics: how much money is made from executing the campaign versus if the campaign were not executed. If the marketing campaign results in a positive ROI, then the process is iterated. Data mining allows discovering all sorts of patterns to refine the strategies outlined above—specifically, optimizing campaigns. Analysts look closely at the navigation of website visitors, and based on these habits, visitors could be classified and analyzed. The resulting profiles will enable the marketing department to tailor landing pages to the specific visitor groups. Because the data mining outputs depend directly on the current campaigns, data mining is very useful for real-time decision making [115-117]. Several techniques can be used for campaign optimization, including cluster analysis, decision trees, and neuro-fuzzy technologies. Two simple campaign optimization techniques are A/B testing and multivariate testing. With A/B testing, one can alternate the page that the visitor sees when he or she requests a webpage. By doing this randomly, a full population of visitors can be used for the tests. If it is desired to incorporate the results in a data mining tool, the selection of which page to show and the result ought to be saved. The result could typically be the click-through rate, the actual order, or other performance measures of the website. The success of a webpage is typically measured as the conversion of a webpage click: the visitor clicks on the back, next, or exit button, or provides other explicit indications that he or she has initiated a purchasing procedure. One can then calculate the conversion rate and assess how successful the different webpages are. This technique enables the webpage owner to determine which webpage is the most successful. The disadvantage of A/B testing is that a single webpage at a time can be tested. Running a campaign where alternatives in content cannot be tested is hard to optimize. In these cases, one describes the possible alternatives and tries to determine the mix that best corresponds to the aimed result. These mixes can be determined using fractional versions of designs that vary each of the canisters at different levels. When the tests are over, a model makes the conversion between the content of the website and the conversion rate. The set of canisters that is predicted from the datasets will typically have a different mix of content than the original mix. The advantage of using this compared to A/B testing is that the whole web content can be tested and that the combined interaction effects between contents can be assessed [118-123]. Campaign optimization is not always easy. Once the proper audience has been determined, another challenge is to provide the proper canister generators with feedback on which criteria to use. To evaluate if this campaign has been successful, the same score must be computed based on the same profile. This forms an evaluation score that can be used to validate the conversion from webpage outputs to canister responses. If a prototype marketing campaign is close to achieving the company's goal, the campaign can go into production. The campaign will then be orchestrated based on the marketing data scenarios determined and tested by the business analysts. Once the campaign is executed in the real market, extensive data is generated. Some of this data is not used for data preparation for the campaign but can be used to improve the description of the customer [124-128].

4. Challenges and Ethical Considerations in Data Mining

There are several challenges associated with data mining, many of which are shared with machine learning applications in general. Common challenges include but are not limited to data quality, time and computational complexity, and heterogeneity and integration across data sources. In addition, data mining also raises substantial practical problems in personal data protection, as well as ethical concerns. As such, it is important that data governance protocols be put in place to ensure that the data and insights derived from data mining applications are reliable and trustworthy [129-133]. Ethical issues surrounding data mining, especially in the marketing context, have become increasingly central for current debates on the digital transformation of the business world. Many consumers are skeptical about data collection practices, particularly when the usage of the data was not made transparent, and they have real and legitimate concerns about the protection of their private lives and data. In the marketing and digital business context, concerns of data security are part of the equation, but there is also a need to consider the legal implications of potential misuse of consumer data, primarily related to data protection and privacy. Moreover, few researchers are exploring the potential of ethics in marketing analytics to guide strategic and marketing decisions and actions. However, as digital marketing develops, the ethical implications of data analysis will become increasingly important. This paper discusses some of the main ethical issues associated with data mining, which have been recognized by scholars in the field. The key considerations include obtaining informed consent from data subjects to collect and use data, respecting individuals' privacy, and the potential side effects of the deployment of data mining algorithms [134-139].

Data Privacy and Security

Concerns about the privacy and security of consumer data hold significant relevance in the context of data mining. The risks associated with data-driven insights and predictions have led to widespread discussions and regulations to help safeguard against the more harmful effects. On the regulations front, there has been considerable movement in recent years, including new provisions aimed at improving control for individuals and increasing obligations. While the advancements are significant, security and privacy appear to be issues that continue to evolve and require adaptation. A data breach, which resulted in hundreds of thousands of voyage records, is an example of the impacts that can affect customer trust. Policies of fair information and data collection are ethical, yet companies have somehow exhibited a higher incidence of inappropriate behavior. Big data has shifted the focus away from consumer data divulgence and has ignored the principles of notice and consent. These concerns led to the development of tools that can be added through a mobile phone's settings. It is of utmost importance to ensure that organizations are following the data regulations to protect the data shared by individuals. Trust is a vital component when viewing from the aspect of establishing a customer relationship. If an individual lacks trust, that will eventually hinder any future from consummating. Techniques to build consumer confidence in privacy protection should be considered among organizations and policies to combat the abuse of private information. If organizations are utilizing personal data, they can ensure secure practices and offer fair opportunities for consumers to make well-informed decisions. Supervisory and regulatory agencies must be diligent and hold organizations accountable for such actions, as this is an avoidable result. In conclusion, there is a need to balance rigorous use with respecting privacy requirements in ethical data mining [140-143].

Bias and Fairness

A common idea in data mining is that "the data never lies." But this is a harmful myth unsupported by work that demonstrates the reality of biased data and the impact of that bias not only on the results of data mining techniques, but on the insights and ultimately actions that flow from those results. Because data mining techniques often use historical data to learn what the patterns of success are for targeting consumers, the demographics and interactions of the people in the training set will shape outcomes. Many times, these training sets contain significant bias. This bias can come from a variety of sources; in marketing training sets, bias is often a result of the way interactions have historically been recorded or completed. In some online systems, women may be offered lower credit limits than men, leading to skewed data. Some companies may sell disproportionately more to certain consumers, leading to marketing and sales training sets that do not accurately reflect the purchasing preferences of other groups. At times, a consumer's race may be inferred, leading to inferred bias in the results of focused models. Other times, model training and test datasets become misaligned due to changes in consumer behavior, distribution methods, or the products or responses being modeled. Additionally, the algorithms chosen for data mining processes can introduce their own bias, compounding or counteracting the existing bias in the training data. This occurs when user interactions are not equally detectable—for example, in a test where expensive software is given to a user to test and cheaper software is not given, even if the user prefers the cheaper product, expensive software will be preferred. In these cases, the advent of commercial data mining vendors and their unwillingness to widely share the processes that they use makes it hard to document. Fairness should be assessed regularly, in case of previously unbiased interactions or models becoming biased. When consumers of marginalized groups are underserved or adversely impacted, unfairness represents significant harm. Even though it is just bad business sense to disproportionately push products on a consumer who is not interested, the default solution is not enough. The best way to guard against this unfairness is to remove or limit the use of sensitive features during data mining and the model operations, when targeting, but also in any decision making. Additionally, accountability and transparency are important ethical concepts that will be violated if marketers cannot identify the logic of their actions or haven't taken actions to address bias. This high-stakes marketing makes it even more important for the fair literature in computer science to be extended and built upon by experts in the data mining community. Given the harm that possibly biased targeting can cause, we propose an extension of the Code of Ethics and Professional Conduct to include language describing the necessity for fairness in targeting, to make data miners and technology professionals who work with them aware of the ethics and industry norms [144-148].

5. Future Trends in Data Mining and Marketing

One of the future trends in data mining, marketing, and customer relationship management is the growing influence of AI and machine learning, set to advance the capabilities of data mining. Artificial intelligence applications and machine learning technology enable data mining to make more accurate predictions. This will enhance the predictive analytics capabilities of marketing and enable a better understanding of customer purchase intentions, behavior, loyalty, and the factors that shape them. Automation in data mining is progressing to the extent that strategic marketing will be driven by live data and the direct action of these automation systems. High-frequency trading data stream analytics in B2B relationships are also increasing and are expected to extend to B2C relationships in social commerce and social media marketing [149-153]. Many new laws have been considered to regulate the mining of big data; however, a new report suggests that there are not going to be any new government data mining and AI laws in 2020. This is because data mining greatly benefits governments and companies. Alternatively, initiatives to increase transparency, regulate data sales, and ensure that the data-driven decision-making process is fair are expected to be introduced. Adaptability is an important skill for future marketing managers as they navigate the potential impacts of advances in data mining technology. Customer insights can be derived from a growing number of consumer data sources, such as smartphone apps and mobile phones, financial credit data, social media, new forms of online communication, and virtual and augmented reality. A marketing leader and an innovative predictive mining organization, whose focus in 2020 is to implement AI-enhanced real-time results marketing solutions for attracting initial investment [154-156].

Artificial Intelligence and Machine Learning

The evolution of the data-driven marketing space has also seen advancements in data mining techniques. Increased processing power and computer memory facilitated the application of artificial intelligence and machine learning in data mining, allowing for more complex analyses than ever. These techniques are based on learning from past data, using feedback as a tool for further learning. These capabilities facilitate more precise predictive analyses, going beyond the mere observation of coinciding scenarios. Today, these techniques have a wide range of applications in marketing, such as consolidation of data from different data sources, the unification and transformation of the data into customer journeys, automated customer clustering, prediction and a posteriori rejection of convenient future courses of action, and/or real-time decision-making. One of their most critical aspects is a kind of prediction that can explain why consumers deviate from predicted behavior – hence an AI and ML engine forms the basis for an online system that provides an on-the-fly answer, interacting dynamically with the user. Such technologies can also be used in real time to personalize the offer and content shown to the user. Key issues in the fruitful employment of AI and ML for marketing are their accuracy, regarding the integrity of the data, which can be corrupted by human biases, the fact that the results coming from them can be machine-driven/interpretability, and the ethical implications of using those technologies. Marketers will have to deal carefully with this, because consumer and customer reactions could be unpredictable, especially when issues like privacy are involved [157-159].

6. Conclusion

Data mining decodes consumer behavior and helps predict sales revenues, thereby giving insight for a data-driven marketing strategy. It helps techniques of association rule mining, clustering, and classification that uncover meaningful insights on customer segmentation and personalization. These are some factors that revolutionize marketing effort. The power of data mining is quite well-translational from retail to healthcare sectors. However, there is a need to harness the capability for data mining in an ethical manner, addressing privacy concerns and ensuring fairness. Advanced AI and machine learning integrated into data mining hold promise for more precise predictive analytics, real-time decision-making, and solidify its importance in the ever-evolving landscape of digital marketing.

7. References

[1] Vanhala, M., Lu, C., Peltonen, J., Sundqvist, S., Nummenmaa, J., & Järvelin, K. (2020). The usage of large data sets in online consumer behaviour: A bibliometric and

computational text-mining-driven analysis of previous research. Journal of Business Research, 106, 46–59. https://doi.org/10.1016/j.jbusres.2019.09.009

[2] Van Nguyen, T., Zhou, L., Chong, A. Y. L., Li, B., & Pu, X. (2020). Predicting customer demand for remanufactured products: A data-mining approach. European Journal of Operational Research, 281(3), 543–558. https://doi.org/10.1016/j.ejor.2019.08.015

[3] Ebrahimi, P., Basirat, M., Yousefi, A., Nekmahmud, M., Gholampour, A., & Fekete-Farkas, M. (2022). Social networks marketing and consumer purchase behavior: The combination of SEM and unsupervised machine learning approaches. Big Data and Cognitive Computing, 6(2), 35. https://doi.org/10.3390/bdcc6020035

[4] Javaid, H. A. (2024). Improving fraud detection and risk assessment in financial service using predictive analytics and data mining. Integrated Journal of Science and Technology.

[5] Sundararaj, V., & Rejeesh, M. R. (2021). A detailed behavioral analysis on consumer and customer changing behavior with respect to social networking sites. Journal of Retailing and Consumer Services. https://doi.org/10.1016/j.jretconser.2020.102190

[6] Nagaraj, P., Muneeswaran, V., Dharanidharan, A., Aakash, M., Balananthanan, K., & Rajkumar, C. (2023, January). E-Commerce customer churn prediction scheme based on customer behavior using machine learning. In 2023 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1–6). IEEE. https://doi.org/10.1109/ICCCI56745.2023.10128498

[7] Arefin, S., Parvez, R., Ahmed, T., Ahsan, M., Sumaiya, F., Jahin, F., & Hasan, M. (2024, May). Retail industry analytics: Unraveling consumer behavior through RFM segmentation and machine learning. In 2024 IEEE International Conference on Electro Information Technology (eIT) (pp. 545–551). IEEE. https://doi.org/10.1109/eIT60633.2024.10609927

[8] Pan, H., & Zhou, H. (2020). Study on convolutional neural network and its application in data mining and sales forecasting for e-commerce. Electronic Commerce Research. https://doi.org/10.1007/s10660-020-09409-0

[9] Sheng, J., Amankwah-Amoah, J., Khan, Z., & Wang, X. (2021). COVID-19 pandemic in the new era of big data analytics: Methodological innovations and future research directions. British Journal of Management, 32(4), 1164–1183. https://doi.org/10.1111/1467-8551.12441

[10] Halkiopoulos, C., & Gkintoni, E. (2024). Leveraging AI in E-Learning: Personalized Learning and Adaptive Assessment through Cognitive Neuropsychology—A Systematic Analysis. Electronics, 13(18), 3762. https://doi.org/10.3390/electronics13183762

[11] Hewamalage, H., Ackermann, K., & Bergmeir, C. (2023). Forecast evaluation for data scientists: Common pitfalls and best practices. Data Mining and Knowledge Discovery, 37(2), 788–832. https://doi.org/10.1007/s10618-022-00894-5

[12] Ashtiani, M. N., & Raahemi, B. (2023). News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2023.119509

[13] Nti, I. K., Quarcoo, J. A., Aning, J., & Fosu, G. K. (2022). A mini-review of machine learning in big data analytics: Applications, challenges, and prospects. Big Data Mining and Analytics, 5(2), 81–97. https://doi.org/10.26599/BDMA.2021.9020028

[14] Giakoumi, A., & Halkiopoulos, C. (2024). Tourism Digital Transformation and Innovative Travel Facilitation Solutions Based on COVID-19: A Case Study of “Fit2Fly” Testing. Recent Advancements in Tourism Business, Technology and Social Sciences, 189–207. https://doi.org/10.1007/978-3-031-54338-8_12

[15] Zhong, Y., Chen, L., Dan, C., & Rezaeipanah, A. (2022). A systematic survey of data mining and big data analysis in the Internet of Things. The Journal of Supercomputing, 78(17), 18405–18453. https://doi.org/10.1007/s11227-022-04594-1

[16] Varadarajan, R. (2020). Customer information resources advantage, marketing strategy and business performance: A market resources-based view. Industrial Marketing Management. https://doi.org/10.1016/j.indmarman.2020.03.003

[17] Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. Journal of the Academy of Marketing Science. https://doi.org/10.1007/s11747-020-00749-9

[18] Gupta, M. K., & Chandra, P. (2020). A comprehensive survey of data mining. International Journal of Information Technology. https://doi.org/10.1007/s41870-020-00427-7

[19] Bharadiya, J. P. (2023). A comparative study of business intelligence and artificial intelligence with big data analytics. American Journal of Artificial Intelligence.

[20] Suma, V., & Hills, S. M. (2020). Data mining-based prediction of demand in the Indian market for refurbished electronics. Journal of Soft Computing Paradigm (JSCP).

[21] Kim, S., Lee, S., & McCulloch, R. (2024). EXPRESS: A topic-based segmentation model for identifying segment-level drivers of star ratings from unstructured text reviews. Journal of Marketing Research. https://doi.org/10.1177/00222437241246752

[22] Khatri, P., Duggal, H. K., Dutta, S., Kumari, P., Thomas, A., Brod, T., & Colimoro, L. (2023). Unveiling heterogeneous knowledge-oriented leadership and knowledge acquisition-based hybrid work agility of knowledge workers. Journal of Knowledge Management, 27(11), 253–278. https://doi.org/10.1108/JKM-10-2022-0793

[23] Juita, V., Pujani, V., Rahim, R., & Rahayu, R. (2024). Dataset on online impulsive buying behavior of "buy now pay later" users and non-buy now pay later users in Indonesia using the stimulus-organism-response model. Data in Brief. https://doi.org/10.1016/j.dib.2024.110500

[24] De Caigny, A., De Bock, K. W., & Verboven, S. (2024). Hybrid black-box classification for customer churn prediction with segmented interpretability analysis. Decision Support Systems. https://doi.org/10.1016/j.dss.2024.114217

[25] Huang, L., Shan, M., Weng, L., & Meng, L. (2024). Graph convolutional spectral clustering for electricity market data clustering. Applied Sciences. https://doi.org/10.3390/app14125263

[26] Chen, W., Liang, Y., Zhu, Y., Chang, Y., Luo, K., Wen, H., ... & Zheng, Y. (2024). Deep learning for trajectory data management and mining: A survey and beyond. arXiv preprint arXiv:2403.14151.

[27] Ryu, S., Kwak, K., & Lee, S. (2024). The growth patterns of firms in the industrial value chain using the hidden Markov model: Evidence from the Korean semiconductor industry. Computers & Industrial Engineering. https://doi.org/10.1016/j.cie.2024.110187

[28] San, S. (2023). Optimizing sales performance in creative-as-a-service (CaaS) companies: A machine learning approach to opportunity time-series forecasting.

[29] Wu, W. T., Li, Y. J., Feng, A. Z., Li, L., Huang, T., Xu, A. D., & Lyu, J. (2021). Data mining in clinical big data: The frequently used databases, steps, and methodological models. Military Medical Research, 8(1), 1–12. https://doi.org/10.1186/s40779-021-00338-z

[30] Pan, Y., & Zhang, L. (2021). A BIM-data mining integrated digital twin framework for advanced project management. Automation in Construction. https://doi.org/10.1016/j.autcon.2021.103564

[31] Gordan, M., Sabbagh-Yazdi, S. R., Ismail, Z., Ghaedi, K., Carroll, P., McCrum, D., & Samali, B. (2022). State-of-the-art review on advancements of data mining in structural health monitoring. Measurement, 193, 110939. https://doi.org/10.1016/j.measurement.2022.110939

[32] Naeem, M., Jamal, T., Diaz-Martinez, J., Butt, S. A., Montesano, N., Tariq, M. I., ... & De-La-Hoz-Valdiris, E. (2022). Trends and future perspective challenges in big data. In Advances in Intelligent Data Analysis and Applications: Proceedings of the Sixth Euro-China Conference on Intelligent Data Analysis and Applications, 15–18 October 2019, Arad, Romania (pp. 309–325). Springer Singapore. https://doi.org/10.1007/978-981-16-5036-9_30

[33] Krassowski, M., Das, V., Sahu, S. K., & Misra, B. B. (2020). State of the field in multi-omics research: From computational needs to data mining and sharing. Frontiers in Genetics, 11, 610798. https://doi.org/10.3389/fgene.2020.610798

[34] Shafiq, M., Tian, Z., Bashir, A. K., Jolfaei, A., & Yu, X. (2020). Data mining and machine learning methods for sustainable smart cities traffic classification: A survey. Sustainable Cities and Society, 60, 102177. https://doi.org/10.1016/j.scs.2020.102177

[35] Alkashami, M., Taamneh, A., Khadragy, S., Shwedeh, F., Aburayya, A., & Salloum, S. (2023). AI different approaches and ANFIS data mining: A novel approach to predicting early employment readiness in Middle Eastern nations. International Journal of Data and Network Science, 7(3), 1267–1282. https://doi.org/10.5267/j.ijdns.2023.4.011

[36] Sunhare, P., Chowdhary, R. R., & Chattopadhyay, M. K. (2022). Internet of things and data mining: An application-oriented survey. Journal of King Saud University-Computer and Information Sciences, 34(6), 3569–3590. https://doi.org/10.1016/j.jksuci.2020.07.002

[37] Sarker, I. H. (2021). Data science and analytics: An overview from data-driven smart computing, decision-making, and applications perspective. SN Computer Science. https://doi.org/10.1007/s42979-021-00765-8

[38] Yang, A., Zhang, W., Wang, J., Yang, K., Han, Y., & Zhang, L. (2020). Review on the application of machine learning algorithms in the sequence data mining of DNA. Frontiers in Bioengineering and Biotechnology, 8, 1032. https://doi.org/10.3389/fbioe.2020.01032

[39] Kong, Y., Feng, C., & Yang, J. (2020). How does China manage its energy market? A perspective of policy evolution. Energy Policy. https://doi.org/10.1016/j.enpol.2020.111898

[40] Ageed, Z. S., Zeebaree, S. R., Sadeeq, M. M., Kak, S. F., Rashid, Z. N., Salih, A. A., & Abdullah, W. M. (2021). A survey of data mining implementation in smart city applications. Qubahan Academic Journal, 1(2), 91–99. https://doi.org/10.48161/qaj.v1n2a52

[41] Telikani, A., Gandomi, A. H., & Shahbahrami, A. (2020). A survey of evolutionary computation for association rule mining. Information Sciences. https://doi.org/10.1016/j.ins.2020.02.073

[42] Santoso, M. H. (2021). Application of association rule method using the Apriori algorithm to find sales patterns: A case study of Indomaret Tanjung Anom. Brilliance: Research of Artificial Intelligence. https://doi.org/10.47709/brilliance.v1i2.1228

[43] Khan, A., & Ghosh, S. K. (2021). Student performance analysis and prediction in classroom learning: A review of educational data mining studies. Education and Information Technologies. https://doi.org/10.1007/s10639-020-10230-3

[44] Salloum, S. A., Alshurideh, M., Elnagar, A., & Shaalan, K. (2020). Mining in educational data: Review and future directions. In Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020) (pp. 92–102). Springer International Publishing. https://doi.org/10.1007/978-3-030-44289-7_9

[45] Han, J., Pei, J., & Tong, H. (2022). Data mining: Concepts and techniques.

[46] Lu, P. H., Keng, J. L., Kuo, K. L., Wang, Y. F., Tai, Y. C., & Kuo, C. Y. (2020). An Apriori algorithm-based association rule analysis to identify herb combinations for treating uremic pruritus using Chinese herbal bath therapy. Evidence-Based Complementary and Alternative Medicine, 2020(1), 8854772. https://doi.org/10.1155/2020/8854772

[47] Lu, P. H., Keng, J. L., Tsai, F. M., Lu, P. H., & Kuo, C. Y. (2021). An Apriori algorithm-based association rule analysis to identify acupoint combinations for treating diabetic gastroparesis. Evidence-Based Complementary and Alternative Medicine, 2021(1), 6649331. https://doi.org/10.1155/2021/6649331

[48] Iqbal, S., Khan, R., Khan, H. U., Alarfaj, F. K., Alomair, A. M., & Ahmed, M. (2022). Association rule analysis-based identification of influential users in social media. Computers, Materials & Continua, 73(3). https://doi.org/10.32604/cmc.2022.030881

[49] de la Cruz-Ruiz, F., Canul-Reich, J., Rivera-López, R., & de la Cruz-Hernández, E. (2024). Impact of data balancing a multiclass dataset before the creation of association rules to study bacterial vaginosis. Intelligent Medicine, 4(03), 188–199. https://doi.org/10.1016/j.imed.2023.02.001

[50] Altay, E. V., & Alatas, B. (2021). Differential evolution and sine cosine algorithm-based novel hybrid multi-objective approaches for numerical association rule mining. Information Sciences. https://doi.org/10.1016/j.ins.2020.12.055

[51] Ampornphan, P., & Tongngam, S. (2020). Exploring technology influencers from patent data using association rule mining and social network analysis. Information. https://doi.org/10.3390/info11060333

[52] Yao, Q., Yang, H., Bao, B., Yu, A., Zhang, J., & Cheriet, M. (2021). Core and spectrum allocation based on association rules mining in spectrally and spatially elastic optical networks. IEEE Transactions on Communications, 69(8), 5299–5311. https://doi.org/10.1109/TCOMM.2021.3082768

[53] Ospina-Mateus, H., Quintana Jiménez, L. A., Lopez-Valdes, F. J., Berrio Garcia, S., Barrero, L. H., & Sana, S. S. (2021). Extraction of decision rules using genetic algorithms and simulated annealing for prediction of severity of traffic accidents by motorcyclists. Journal of Ambient Intelligence and Humanized Computing, 12(11), 10051–10072. https://doi.org/10.1007/s12652-020-02759-5

[54] Chowdhury, S., & Zhu, J. (2023). Investigation of critical factors for future-proofed transportation infrastructure planning using topic modeling and association rule

mining. Journal of Computing in Civil Engineering. https://doi.org/10.1061/(ASCE)CP.1943-5487.0001059
[55] Nnoaham, K. E., & Cann, K. F. (2020). Can cluster analyses of linked healthcare data identify unique population segments in a general practice-registered population? BMC Public Health. https://doi.org/10.1186/s12889-020-08930-z
[56] Hernani-Merino, M., Lazo, J. G. L., López, A. T., Mazzon, J. A., & López-Tafur, G. (2020). An international market segmentation model based on susceptibility to global consumer culture. Cross Cultural & Strategic Management, 28(1), 108–128. https://doi.org/10.1108/CCSM-04-2019-0081
[57] Abdul-Rahman, S., Arifin, N. F. K., Hanafiah, M., & Mutalib, S. (2021). Customer segmentation and profiling for life insurance using k-modes clustering and decision tree classifier. International Journal of Advanced Computer Science and Applications, 12(9), 434–444. https://doi.org/10.14569/IJACSA.2021.0120950
[58] Ghosal, A., Nandy, A., Das, A. K., Goswami, S., & Panday, M. (2020). A short review on different clustering techniques and their applications. In Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018 (pp. 69–83). Springer. https://doi.org/10.1007/978-981-13-7403-6_9
[59] Lee, Z. J., Lee, C. Y., Chang, L. Y., & Sano, N. (2021). Clustering and classification based on distributed automatic feature engineering for customer segmentation. Symmetry. https://doi.org/10.3390/sym13091557
[60] Reutterer, T., & Dan, D. (2021). Cluster analysis in marketing research. In Handbook of Market Research. https://doi.org/10.1007/978-3-319-57413-4_11
[61] Neeraj, K. N., & Maurya, V. (2020). A review on machine learning (feature selection, classification, and clustering) approaches of big data mining in different areas of research. Journal of Critical Reviews.
[62] Benslama, T., & Jallouli, R. (2020). Clustering of social media data and marketing decisions. In Digital Economy. Emerging Technologies and Business Innovation: 5th International Conference on Digital Economy, ICDEc 2020, Bucharest, Romania, June 11-13, 2020, Proceedings 5 (pp. 53–65). Springer International Publishing. https://doi.org/10.1007/978-3-030-64642-4_5
[63] Das, S., & Nayak, J. (2022). Customer segmentation via data mining techniques: State-of-the-art review. In Computational Intelligence in Data Mining: Proceedings of ICCIDM 2021 (pp. 489–507). Springer. https://doi.org/10.1007/978-981-16-9447-9_38
[64] Yoseph, F., Ahamed Hassain Malim, N. H., Heikkilä, M., Brezulianu, A., Geman, O., & Paskhal Rostam, N. A. (2020). The impact of big data market segmentation using data mining and clustering techniques. Journal of Intelligent & Fuzzy Systems, 38(5), 6159–6173. https://doi.org/10.3233/JIFS-179698
[65] Hicham, N., & Karim, S. (2022). Analysis of unsupervised machine learning techniques for an efficient customer segmentation using clustering ensemble and spectral clustering. International Journal of Advanced Computer Science and Applications, 13(10). https://doi.org/10.14569/IJACSA.2022.0131016
[66] Faizan, M., Zuhairi, M. F., Ismail, S., & Sultan, S. (2020). Applications of clustering techniques in data mining: A comparative study. International Journal of Advanced Computer Science and Applications, 11(12). https://doi.org/10.14569/IJACSA.2020.0111218

[67] Ramasubbareddy, S., Srinivas, T. A. S., Govinda, K., & Manivannan, S. S. (2020). Comparative study of clustering techniques in market segmentation. In Innovations in Computer Science and Engineering: Proceedings of 7th ICICSE (pp. 117–125). Springer. https://doi.org/10.1007/978-981-15-2043-3_15

[68] Khan, Y., Sadia, H., Ali Shah, S. Z., Khan, M. N., Shah, A. A., Ullah, N., ... & Khan, M. I. (2022). Classification, synthetic, and characterization approaches to nanoparticles, and their applications in various fields of nanotechnology: A review. Catalysts, 12(11), 1386. https://doi.org/10.3390/catal12111386

[69] Omer, S., Forgách, L., Zelkó, R., & Sebe, I. (2021). Scale-up of electrospinning: Market overview of products and devices for pharmaceutical and biomedical purposes. Pharmaceutics. https://doi.org/10.3390/pharmaceutics13020286

[70] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications, and research directions. SN Computer Science. https://doi.org/10.1007/s42979-021-00592-x

[71] Barhoum, A., García-Betancourt, M. L., Jeevanandam, J., Hussien, E. A., Mekkawy, S. A., Mostafa, M., ... & Bechelany, M. (2022). Review on natural, incidental, bioinspired, and engineered nanomaterials: History, definitions, classifications, synthesis, properties, market, toxicities, risks, and regulations. Nanomaterials, 12(2), 177. https://doi.org/10.3390/nano12020177

[72] Hassani, H., & Silva, E. S. (2023). The role of ChatGPT in data science: How AI-assisted conversational interfaces are revolutionizing the field. Big Data and Cognitive Computing. https://doi.org/10.3390/bdcc7020062

[73] Huan, Y., Kong, Q., Mou, H., & Yi, H. (2020). Antimicrobial peptides: Classification, design, application, and research progress in multiple fields. Frontiers in Microbiology. https://doi.org/10.3389/fmicb.2020.582779

[74] Wedel, M., Bigné, E., & Zhang, J. (2020). Virtual and augmented reality: Advancing research in consumer marketing. International Journal of Research in Marketing, 37(3), 443–465. https://doi.org/10.1016/j.ijresmar.2020.04.004

[75] Almustafa, K. M. (2020). Prediction of heart disease and classifiers' sensitivity analysis. BMC Bioinformatics. https://doi.org/10.1186/s12859-020-03626-y

[76] Yoo, S. H., Geng, H., Chiu, T. L., Yu, S. K., Cho, D. C., Heo, J., ... & Lee, H. (2020). Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging. Frontiers in Medicine, 7, 427. https://doi.org/10.3389/fmed.2020.00427

[77] Dai, E., & Wang, S. (2021). Say no to the discrimination: Learning fair graph neural networks with limited sensitive attribute information. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining (pp. 680–688). https://doi.org/10.1145/3437963.3441752

[78] Kim, Y., Cho, D., Han, K., Panda, P., & Hong, S. (2021). Domain adaptation without source data. IEEE Transactions on Artificial Intelligence, 2(6), 508–518. https://doi.org/10.1109/TAI.2021.3110179

[79] Althnian, A., AlSaeed, D., Al-Baity, H., Samha, A., Dris, A. B., Alzakari, N., ... & Kurdi, H. (2021). Impact of dataset size on classification performance: An empirical evaluation in the medical domain. Applied Sciences, 11(2), 796. https://doi.org/10.3390/app11020796

[80] Tang, Y. X., Tang, Y. B., Peng, Y., Yan, K., Bagheri, M., Redd, B. A., ... & Summers, R. M. (2020). Automated abnormality classification of chest radiographs using deep

convolutional neural networks. NPJ Digital Medicine, 3(1), 70. https://doi.org/10.1038/s41746-020-0273-z
[81] Rankin, D., Black, M., Bond, R., Wallace, J., Mulvenna, M., & Epelde, G. (2020). Reliability of supervised machine learning using synthetic data in healthcare: Model to preserve privacy for data sharing. JMIR Medical Informatics, 8(7), e18910. https://doi.org/10.2196/18910
[82] Wu, J., Shi, L., Lin, W. P., Tsai, S. B., Li, Y., Yang, L., & Xu, G. (2020). [Retracted] An empirical study on customer segmentation by purchase behaviors using a RFM model and K-Means algorithm. Mathematical Problems in Engineering, 2020(1), 8884227. https://doi.org/10.1155/2020/8884227
[83] Brahmana, R. S., Mohammed, F. A., & Chairuang, K. (2020). Customer segmentation based on RFM model using K-means, K-medoids, and DBSCAN methods. Lontar Komput. J. Ilm. Teknol. Inf, 11(1), 32. https://doi.org/10.24843/LKJITI.2020.v11.i01.p04
[84] Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM ranking—An effective approach to customer segmentation. Journal of King Saud University-Computer and Information Sciences, 33(10), 1251–1257. https://doi.org/10.1016/j.jksuci.2018.09.004
[85] Wang, C. (2022). Efficient customer segmentation in digital marketing using deep learning with swarm intelligence approach. Information Processing & Management. https://doi.org/10.1016/j.ipm.2022.103085
[86] Geetha, B. T., Yenugula, M., Randhawa, N., Purohit, P., Maney, K. L., & Venkteshwar, A. (2024). Advancement improving the acquisition of customer insights in digital marketing by utilising advanced artificial intelligence algorithms. In 2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies (pp. 1–7). https://doi.org/10.1109/TQCEBT59414.2024.10545055
[87] Panas, G., Thrasidi, N., Halkiopoulos, C., Gkintoni, E. (2021). Consumer Behavior and Cognitive Factors in Relation to Gastronomic Tourism and Destination Marketing in Greece. In: Katsoni, V., Şerban, A.C. (eds) Transcending Borders in Tourism Through Innovation and Cultural Heritage. Springer Proceedings in Business and Economics. Springer, Cham. https://doi.org/10.1007/978-3-030-92491-1_40
[88] Gupta, S., Leszkiewicz, A., Kumar, V., Bijmolt, T., & Potapov, D. (2020). Digital analytics: Modeling for insights and new methods. Journal of Interactive Marketing, 51(1), 26–43. https://doi.org/10.1016/j.intmar.2020.04.003
[89] Bharadiya, J. P. (2023). Machine learning and AI in business intelligence: Trends and opportunities. International Journal of Computer (IJC).
[90] Patil, D., Rane, N. L., & Rane, J. (2024). Applications of ChatGPT and generative artificial intelligence in transforming the future of various business sectors. In The Future Impact of ChatGPT on Several Business Sectors (pp. 1–47). https://doi.org/10.70593/978-81-981367-8-7_1
[91] Mariani, M. M., & Wamba, S. F. (2020). Exploring how consumer goods companies innovate in the digital age: The role of big data analytics companies. Journal of Business Research. https://doi.org/10.1016/j.jbusres.2020.09.012
[92] Panas, G., Thrasidi, N., Halkiopoulos, C., & Gkintoni, E. (2022). Consumer Behavior and Cognitive Factors in Relation to Gastronomic Tourism and Destination Marketing in Greece. Springer Proceedings in Business and Economics, 655–677. https://doi.org/10.1007/978-3-030-92491-1_40

[93] Sahu, M. K. (2022). Machine learning for personalized marketing and customer engagement in retail: Techniques, models, and real-world applications. Journal of Artificial Intelligence Research and Applications, 2(1), 219–254.
[94] Igoumenakis, G., Antonopoulou, H., & Halkiopoulos, C. (2024). Digital Marketing and the Contribution to Sustainable Tourism: Case Study of Zakynthos Island. Recent Advancements in Tourism Business, Technology and Social Sciences, 293–321. https://doi.org/10.1007/978-3-031-54338-8_17
[95] Ranfagni, S., Faraoni, M., Zollo, L., & Vannucci, V. (2021). Combining online market research methods for investigating brand alignment: The case of Nespresso. British Food Journal. https://doi.org/10.1108/BFJ-06-2020-0462
[96] Perez-Vega, R., Hopkinson, P., Singhal, A., & Mariani, M. M. (2022). From CRM to social CRM: A bibliometric review and research agenda for consumer research. Journal of Business Research, 151, 1–16. https://doi.org/10.1016/j.jbusres.2022.06.028
[97] Kimura, M. (2022). Customer segment transition through the customer loyalty program. Asia Pacific Journal of Marketing and Logistics. https://doi.org/10.1108/APJML-09-2020-0630
[98] Shetty, K., & Fitzsimmons, J. R. (2022). The effect of brand personality congruence, brand attachment, and brand love on loyalty among HENRY's in the luxury branding sector. Journal of Fashion Marketing and Management: An International Journal, 26(1), 21–35. https://doi.org/10.1108/JFMM-09-2020-0208
[99] Huangfu, Z., Ruan, Y., Zhao, J., Wang, Q., & Zhou, L. (2022). Accessing the influence of community experience on brand loyalty toward virtual brand community: Developing country perspective. Frontiers in Psychology, 13, 865646. https://doi.org/10.3389/fpsyg.2022.865646
[100] Halkiopoulos, C., Antonopoulou, H., Gkintoni, E., & Aroutzidis, A. (2022). Neuromarketing as an Indicator of Cognitive Consumer Behavior in Decision-Making Process of Tourism destination—An Overview. Springer Proceedings in Business and Economics, 679–697. https://doi.org/10.1007/978-3-030-92491-1_41
[101] Li, Y., Chu, X., Tian, D., Feng, J., & Mu, W. (2021). Customer segmentation using K-means clustering and the adaptive particle swarm optimization algorithm. Applied Soft Computing. https://doi.org/10.1016/j.asoc.2021.107924
[102] Yadegaridehkordi, E., Nilashi, M., Nasir, M. H. N. B. M., Momtazi, S., Samad, S., Supriyanto, E., & Ghabban, F. (2021). Customer segmentation in eco-friendly hotels using multi-criteria and machine learning techniques. Technology in Society, 65, 101528. https://doi.org/10.1016/j.techsoc.2021.101528
[103] Theodorakopoulos, L., Karras, A., Theodoropoulou, A., & Kampiotis, G. (2024). Benchmarking Big Data Systems: Performance and Decision-Making Implications in Emerging Technologies. Technologies, 12(11), 217. https://doi.org/10.3390/technologies12110217
[104] Sahu, M. K. (2020). Machine learning algorithms for personalized financial services and customer engagement: Techniques, models, and real-world case studies. Distributed Learning and Broad Applications in Scientific Research, 6, 272–313.
[105] Feng, G., Fan, M., & Chen, Y. (2022). Analysis and prediction of students' academic performance based on educational data mining. IEEE Access. https://doi.org/10.1109/ACCESS.2022.3151652

[106] Strielkowski, W., Vlasov, A., Selivanov, K., Muraviev, K., & Shakhnov, V. (2023). Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review. Energies, 16(10), 4025. https://doi.org/10.3390/en16104025
[107] Kumar, D., Haque, A., Mishra, K., Islam, F., Mishra, B. K., & Ahmad, S. (2023). Exploring the transformative role of artificial intelligence and metaverse in education: A comprehensive review. Metaverse Basic and Applied Research, 2, 55–55. https://doi.org/10.56294/mr202355
[108] Hasan, R., Palaniappan, S., Mahmood, S., Abbas, A., Sarker, K. U., & Sattar, M. U. (2020). Predicting student performance in higher educational institutions using video learning analytics and data mining techniques. Applied Sciences, 10(11), 3894. https://doi.org/10.3390/app10113894
[109] Cherif, A., Badhib, A., Ammar, H., Alshehri, S., Kalkatawi, M., & Imine, A. (2023). Credit card fraud detection in the era of disruptive technologies: A systematic review. Journal of King Saud University-Computer and Information Sciences, 35(1), 145–174. https://doi.org/10.1016/j.jksuci.2022.11.008
[110] Kumar, S., Ahmed, R., Bharany, S., Shuaib, M., Ahmad, T., Tag Eldin, E., ... & Shafiq, M. (2022). Exploitation of machine learning algorithms for detecting financial crimes based on customers' behavior. Sustainability, 14(21), 13875. https://doi.org/10.3390/su142113875
[111] Theodorakopoulos, L., Thanasas, G., & Halkiopoulos, C. (2024). Implications of Big Data in Accounting: Challenges and Opportunities. Emerging Science Journal, 8(3), 1201–1214. https://doi.org/10.28991/esj-2024-08-03-024
[112] Tellman, B., Sullivan, J. A., Kuhn, C., Kettner, A. J., Doyle, C. S., Brakenridge, G. R., ... & Slayback, D. A. (2021). Satellite imaging reveals increased proportion of population exposed to floods. Nature, 596(7870), 80–86. https://doi.org/10.1038/s41586-021-03695-w
[113] Yang, D., Alessandrini, S., Antonanzas, J., Antonanzas-Torres, F., Badescu, V., Beyer, H. G., ... & Zhang, J. (2020). Verification of deterministic solar forecasts. Solar Energy, 210, 20–37. https://doi.org/10.1016/j.solener.2020.04.019
[114] Deeks, J. J., Bossuyt, P. M., Leeflang, M. M., & Takwoingi, Y. (2023). Cochrane handbook for systematic reviews of diagnostic test accuracy. https://doi.org/10.1002/9781119756194
[115] Rosário, A. T., & Dias, J. C. (2023). How has data-driven marketing evolved: Challenges and opportunities with emerging technologies. International Journal of Information Management Data Insights, 3(2), 100203. https://doi.org/10.1016/j.jjimei.2023.100203
[116] Adeniran, I. A., Efunniyi, C. P., Osundare, O. S., & Abhulimen, A. O. (2024). Transforming marketing strategies with data analytics: A study on customer behavior and personalization. International Journal of Management & Entrepreneurship Research, 6(8).
[117] Palle, R. R. (2021). Discuss the role of data analytics in extracting meaningful insights from social media data, influencing marketing strategies and user engagement. Journal of Artificial Intelligence and Machine Learning in Management, 5(1), 64–69.
[118] Baier, D., & Rese, A. (2020). Increasing conversion rates through eye tracking, TAM, A/B tests: A case study. Advanced Studies in Behaviormetrics and Data Science:

Essays in Honor of Akinori Okada, 341–353. https://doi.org/10.1007/978-981-15-2700-5_21
[119] Quin, F., Weyns, D., Galster, M., & Silva, C. C. (2024). A/B testing: A systematic literature review. Journal of Systems and Software. https://doi.org/10.1016/j.jss.2024.112011
[120] Nandy, P., Venugopalan, D., Lo, C., & Chatterjee, S. (2021). A/B testing for recommender systems in a two-sided marketplace. Advances in Neural Information Processing Systems, 34, 6466–6477.
[121] Westland, J. C. (2022). A comparative study of frequentist vs Bayesian A/B testing in the detection of E-commerce fraud. Journal of Electronic Business & Digital Economics. https://doi.org/10.1108/JEBDE-07-2022-0020
[122] Larsen, N., Stallrich, J., Sengupta, S., Deng, A., Kohavi, R., & Stevens, N. T. (2024). Statistical challenges in online controlled experiments: A review of A/B testing methodology. The American Statistician, 78(2), 135–149. https://doi.org/10.1080/00031305.2023.2257237
[123] Croxen-John, D., & Van Tonder, J. (2020). E-commerce website optimization: Why 95% of your website visitors don't buy, and what you can do about it.
[124] Kedi, W. E., Ejimuda, C., Idemudia, C., & Ijomah, T. I. (2024). Machine learning software for optimizing SME social media marketing campaigns. Computer Science & IT Research Journal, 5(7), 1634–1647. https://doi.org/10.51594/csitrj.v5i7.1349
[125] Du, R. Y., Netzer, O., Schweidel, D. A., & Mitra, D. (2021). Capturing marketing information to fuel growth. Journal of Marketing, 85(1), 163–183. https://doi.org/10.1177/0022242920969198
[126] Haleem, A., Javaid, M., Qadri, M. A., Singh, R. P., & Suman, R. (2022). Artificial intelligence (AI) applications for marketing: A literature-based study. International Journal of Intelligent Networks, 3, 119–132. https://doi.org/10.1016/j.ijin.2022.08.005
[127] Callejo, P., Gramaglia, M., Cuevas, R., & Cuevas, A. (2022). A deep dive into the accuracy of IP geolocation databases and its impact on online advertising. IEEE Transactions on Mobile Computing, 22(8), 4359–4373. https://doi.org/10.1109/TMC.2022.3166785
[128] Breza, E., Stanford, F. C., Alsan, M., Alsan, B., Banerjee, A., Chandrasekhar, A. G., ... & Duflo, E. (2021). Effects of a large-scale social media advertising campaign on holiday travel and COVID-19 infections: A cluster randomized controlled trial. Nature Medicine, 27(9), 1622–1628. https://doi.org/10.1038/s41591-021-01487-3
[129] Wu, X., Duan, R., & Ni, J. (2024). Unveiling security, privacy, and ethical concerns of ChatGPT. Journal of Information and Intelligence. https://doi.org/10.1016/j.jiixd.2023.10.007
[130] Thapa, C., & Camtepe, S. (2021). Precision health data: Requirements, challenges, and existing techniques for data security and privacy. Computers in Biology and Medicine. https://doi.org/10.1016/j.compbiomed.2020.104130
[131] Ogbuke, N. J., Yusuf, Y. Y., Dharma, K., & Mercangoz, B. A. (2022). Big data supply chain analytics: Ethical, privacy and security challenges posed to business, industries, and society. Production Planning & Control, 33(2-3), 123–137. https://doi.org/10.1080/09537287.2020.1810764
[132] Roshanaei, M., Khan, M. R., & Sylvester, N. N. (2024). Enhancing cybersecurity through AI and ML: Strategies, challenges, and future directions. Journal of Information Security. https://doi.org/10.4236/jis.2024.153019

[133] Andrew, J., & Baker, M. (2021). The General Data Protection Regulation in the age of surveillance capitalism. Journal of Business Ethics. https://doi.org/10.1007/s10551-019-04239-z
[134] Hemker, S., Herrando, C., & Constantinides, E. (2021). The transformation of data marketing: How an ethical lens on consumer data collection shapes the future of marketing. Sustainability. https://doi.org/10.3390/su132011208
[135] Florea, D., & Florea, S. (2020). Big data and the ethical implications of data privacy in higher education research. Sustainability. https://doi.org/10.3390/su12208744
[136] Mullins, M., Holland, C. P., & Cunneen, M. (2021). Creating ethics guidelines for artificial intelligence and big data analytics customers: The case of the consumer European insurance market. Patterns. https://doi.org/10.1016/j.patter.2021.100362
[137] Halkiopoulos, C., Antonopoulou, H., Gkintoni, E., Aroutzidis, A. (2021). Neuromarketing as an Indicator of Cognitive Consumer Behavior in Decision Making Process of Tourism Destination. In: Katsoni, V., Şerban, A.C. (eds) Transcending Borders in Tourism Through Innovation and Cultural Heritage. Springer Proceedings in Business and Economics. Springer, Cham. https://doi.org/10.1007/978-3-030-92491-1_41
[138] Illia, L., Colleoni, E., & Zyglidopoulos, S. (2023). Ethical implications of text generation in the age of artificial intelligence. Business Ethics, the Environment & Responsibility, 32(1), 201–210. https://doi.org/10.1111/beer.12479
[139] Graafland, J., & Wells, T. R. (2021). In Adam Smith's own words: The role of virtues in the relationship between free market economies and societal flourishing, a semantic network data-mining approach. Journal of Business Ethics. https://doi.org/10.1007/s10551-020-04521-5
[140] Okazaki, S., Eisend, M., Plangger, K., de Ruyter, K., & Grewal, D. (2020). Understanding the strategic consequences of customer privacy concerns: A meta-analytic review. Journal of Retailing, 96(4), 458–473. https://doi.org/10.1016/j.jretai.2020.05.007
[141] Quach, S., Thaichon, P., Martin, K. D., Weaven, S., & Palmatier, R. W. (2022). Digital technologies: Tensions in privacy and data. Journal of the Academy of Marketing Science, 50(6), 1299–1323. https://doi.org/10.1007/s11747-022-00845-y
[142] Akour, I., Alnazzawi, N., Alshurideh, M., Almaiah, M. A., Al Kurdi, B., Alfaisal, R. M., & Salloum, S. (2022). A conceptual model for investigating the effect of privacy concerns on E-commerce adoption: A study on United Arab Emirates consumers. Electronics, 11(22), 3648. https://doi.org/10.3390/electronics11223648
[143] Sun, Z., Strang, K. D., & Pambel, F. (2020). Privacy and security in the big data paradigm. Journal of Computer Information Systems. https://doi.org/10.1080/08874417.2017.1418631
[144] Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. Journal of Business Research, 144, 201–216. https://doi.org/10.1016/j.jbusres.2022.01.083
[145] Hovy, D., & Prabhumoye, S. (2021). Five sources of bias in natural language processing. Language and Linguistics Compass. https://doi.org/10.1111/lnc3.12432
[146] Susarla, A., Gopal, R., Thatcher, J. B., & Sarker, S. (2023). The Janus effect of generative AI: Charting the path for responsible conduct of scholarly activities in information systems. Information Systems Research, 34(2), 399–408. https://doi.org/10.1287/isre.2023.ed.v34.n2

[147] Yu, Y., Zhuang, Y., Zhang, J., Meng, Y., Ratner, A. J., Krishna, R., ... & Zhang, C. (2024). Large language model as attributed training data generator: A tale of diversity and bias. Advances in Neural Information Processing Systems, 36.

[148] Borji, A. (2023). A categorical archive of ChatGPT failures. arXiv preprint arXiv:2302.03494. https://doi.org/10.21203/rs.3.rs-2895792/v1

[149] Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: Machine intelligence approach for drug discovery. Molecular Diversity, 25, 1315–1360. https://doi.org/10.1007/s11030-021-10217-3

[150] Albahri, A. S., Hamid, R. A., Alwan, J. K., Al-Qays, Z. T., Zaidan, A. A., Zaidan, B. B., ... & Madhloom, H. T. (2020). Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): A systematic review. Journal of Medical Systems, 44, 1–11. https://doi.org/10.1007/s10916-020-01582-x

[151] Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A. H., Alomari, S., Gök, M., Alaabdin, A. M. Z., & Abdulrhman, S. H. (2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. Iraqi Journal for Computer Science and Mathematics, 3(1), 115–123. https://doi.org/10.52866/ijcsm.2022.01.01.013

[152] Theodorakopoulos, L., Theodoropoulou, A., & Halkiopoulos, C. (2024). Enhancing Decentralized Decision-Making with Big Data and Blockchain Technology: A Comprehensive Review. Applied Sciences, 14(16), 7007. https://doi.org/10.3390/app14167007

[153] Gladju, J., Kamalam, B. S., & Kanagaraj, A. (2022). Applications of data mining and machine learning framework in aquaculture and fisheries: A review. Smart Agricultural Technology. https://doi.org/10.1016/j.atech.2022.100061

[154] Krafft, M., Kumar, V., Harmeling, C., Singh, S., Zhu, T., Chen, J., ... & Rosa, E. (2021). Insight is power: Understanding the terms of the consumer-firm data exchange. Journal of Retailing, 97(1), 133–149. https://doi.org/10.1016/j.jretai.2020.11.001

[155] Holmlund, M., Van Vaerenbergh, Y., Ciuchita, R., Ravald, A., Sarantopoulos, P., Ordenes, F. V., & Zaki, M. (2020). Customer experience management in the age of big data analytics: A strategic framework. Journal of Business Research, 116, 356–365. https://doi.org/10.1016/j.jbusres.2020.01.022

[156] Perez-Vega, R., Kaartemo, V., Lages, C. R., Razavi, N. B., & Männistö, J. (2021). Reshaping the contexts of online customer engagement behavior via artificial intelligence: A conceptual framework. Journal of Business Research, 129, 902–910. https://doi.org/10.1016/j.jbusres.2020.11.002

[157] Papadopoulos, D., Gkintoni, E., Halkiopoulos, C., & Antonopoulou, H. (2023). e-Tourism and Online Reservation Systems in Electronic Business. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4319582

[158] Perera, A., & Iqbal, K. (2021). Big data and emerging markets: Transforming economies through data-driven innovation and market dynamics. Journal of Computational Social Dynamics.

[159] Kozak, J., Kania, K., Juszczuk, P., & Mitręga, M. (2021). Swarm intelligence goal-oriented approach to data-driven innovation in customer churn management. International Journal of Information Management, 60, 102357. https://doi.org/10.1016/j.ijinfomgt.2021.102357

[160] Saura, J. R. (2021). Advanced digital marketing strategies in a data-driven era. https://doi.org/10.4018/978-1-7998-8003-5