International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 10-No.4 (2024) pp. 1344-1350
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Real-Time E-commerce Insights with Mean Shift Clustering: A Dynamic Approach to Customer Understanding

Trivikrama Rao BATTULA[1]*, Narayana GARLAPATI[2], Srinivasa Rao CHOPPARAPU[3], Narasimha Swamy LAVUDIYA[4], Prasad GUNDE[5]

[1]Department of IT, Usha Rama College of Engineering and Technolog, Telaprolu - 521109, AP, India
* **Corresponding Author Email:** battula.trivi@gmail.com - **ORCID:** 0009-0009-0964-7867

[2]Associate Professor, Department of CSE-AIML, CBIT, Hyderabad, Telangana - 500075, AP, India
Email: narayanag.1973@gmail.com - **ORCID:** 0000-0001-8470-3595

[3]Department of CSE, Lakireddy Bali Reddy College of Engineering, Mylavaram - 521230, AP, India
Email: srinivas.lovely10@gmail.com- **ORCID:** 0000-0002-7351-7540

[4]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram-522302, AP, India
Email: narasimhaswamyrl@gmail.com- **ORCID:** 0009-0003-9247-9660

[5]Department of CSE, PSCMR College of Engg and Technology, Kothapeta, Vijayawada- 520001, AP, India
Email: prasad.gunde@gmail.com- **ORCID:** 0009-0003-9281-9872

Article Info:

DOI: 10.22399/ijcesen.607
Received : 11 November 2024
Accepted : 11 December 2024

Keywords:

E-Commerce,
Density-Based Clustering,
Fraud Detection,
Mean shift Vector,
GPU,

Abstract:

In the high-speed universe of internet business, understanding client conduct progressively is urgent for customized encounters and ideal business results. This paper investigates the utilization of Mean Shift bunching, a strong non-parametric thickness-based calculation, for continuous examination in online business. By utilizing Mean Shift's capacity to progressively distinguish bunches of erratic shapes, organizations can acquire important bits of knowledge into client conduct, even as it advances. We show the way that Mean Shift can fragment clients in view of their ongoing perusing movement, search questions, item associations, and buy designs, making dynamic client profiles that mirror their ebb and flow interests and inclinations. This empowers organizations to convey profoundly customized proposals, upgrade valuing techniques, and designer promoting efforts in light of constant client needs. Moreover, we investigate how Mean Shift can be utilized to foresee future client conduct, empowering organizations to expect needs and proactively tailor the shopping experience. The paper additionally addresses the difficulties of carrying out ongoing Mean Shift grouping, including information streaming and adaptability, computational intricacy, and information protection concerns. We finish up by illustrating future exploration headings for improving the viability of Mean Change continuously online business examination, underscoring its capability to reform the manner in which organizations draw in with clients in a dynamic and consistently changing internet based commercial center.

1. Introduction

E-commerce is a dynamic and rapidly evolving landscape, characterized by a constant influx of new products, shifting customer preferences, and fierce competition. This fast-paced nature demands businesses to adapt quickly and stay ahead of the curve, requiring a deep understanding of customer behavior in real-time to optimize operations and maximize success.Traditional approaches to customer analysis, often relying on historical data and batch processing, are too slow to react to the rapid changes in online shopping behavior. This delay can lead to missed opportunities, inefficient marketing campaigns, and a less personalized customer experience. To thrive in this dynamic environment, e-commerce businesses need to move beyond static analysis and embrace real-time. insights to understand customer needs, preferences.

1.1 Importance of Real-Time Customer Understanding

Customized Suggestions: Constant examination empowers the conveyance of customized item proposals in view of a client's ongoing perusing action and past buys, fundamentally improving the shopping experience and expanding transformation rates [1, 2].

Dynamic Evaluating: Seeing constant interest designs takes into account dynamic valuing procedures that advance income and adjust to fluctuating economic situations [3, 4].

Designated Advertising Efforts: Constant information gives bits of knowledge into client interests and buy expectation, empowering more designated showcasing efforts that convey customized messages and advancements for more prominent effect [5, 6].

Fraud Detection: Constant checking of client action distinguishes bizarre way of behaving, identifying false exchanges and protecting business income [7, 8].

Enhanced Customer Engagement: By expecting client needs and conveying customized encounters, constant examination cultivates a really captivating and fulfilling shopping venture, expanding client unwavering ness and rehash buys [9, 10].

Conventional ways to deal with client investigation in web-based business frequently depend on verifiable information and bunch handling, which can introduce a few restrictions when confronted with the unique idea of online way of behaving. These impediments feature the requirement for dynamic, ongoing bits of knowledge to really comprehend and draw in with clients in the always advancing web-based business scene.

1.2 Limitations of Historical Data-Based Approaches:

Time Lag: Verifiable information-based examination innately works with a postponement, depending on past information that could not precisely reflect current client conduct. This delay can prompt botched open doors, insufficient promoting efforts, and a less customized client experience [1,2].

Static Segmentation: Conventional strategies frequently fragment clients in view of authentic information, which may not catch the unique changes to their greatest advantage and inclinations after some time. This static division can bring about incorrect focusing on and less successful personalization endeavours [5, 6].

Limited Adaptability: Authentic information-based examination battles to adjust to changing economic situations, new item dispatches, and developing client patterns. This absence of versatility can prompt a distinction between client assumptions and the showcasing and limited-time endeavours of Internet business organizations [7, 8].

Inability to Detect Anomalies: The verifiable information-based examination is less compelling at distinguishing ongoing irregularities in client conduct, which could demonstrate fake movement or other security dangers. This can prompt monetary misfortunes and compromise client trust [9, 10].

1.3 Need for Dynamic, Real-Time Insights:

To conquer these limits, web-based business organizations need to move past authentic information and embrace continuous investigation. This includes catching and breaking down client action as it works out, empowering organizations to:

Personalize Experiences: Give dynamic item suggestions, customized evaluating, and designated advertising messages in light of ongoing client conduct [3, 4].

Optimize Operations: Change estimating systems, stock levels, and promoting efforts continuously to expand income and limit misfortunes [5, 6].

Enhance Security: Recognize and moderate deceitful movement, distinguish dubious way of behaving, and protect client information progressively [7, 8].

Stay Ahead of the Curve: Adjust rapidly to advancing client patterns, market changes, and new item dispatches to remain serious [9, 10].

In the domain of information examination, bunching calculations assume a critical part in distinguishing examples and gathering comparable data of interest. While many grouping procedures exist, Mean Shift stands apart as a strong and flexible methodology, especially appropriate for dynamic conditions where information designs are continually evolving.

1.4 Mean Shift Clustering: A Non-Parametric Density-Based Approach

Mean Shift is a non-parametric clustering algorithm, meaning it doesn't require pre-defined assumptions about the shape or structure of the data clusters. It operates based on the principle of density estimation, where data points are shifted towards regions of higher density, ultimately forming clusters around these peaks. This approach makes it highly adaptable to complex datasets and allows it to identify clusters of arbitrary shapes, unlike algorithms that rely on predefined structures, such as K-Means. Figure 1. shows how the mean shift clustering approach works by shifting the centroid to a new position using the mean shift vector.

1.5 Key Features of Mean Shift:

Density-Based Clustering: Mean Shift identifies clusters by analyzing the density of data points. It shifts each data point towards the direction of the highest density gradient, effectively moving them towards the center of a cluster.

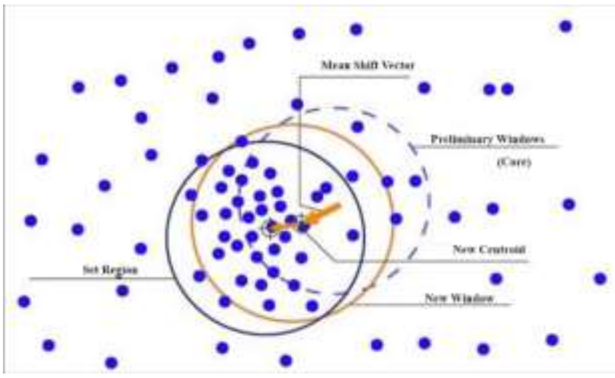

Figure 1. *Mean shift Cluster Approach*

Non-Parametric: It does not require any prior assumptions about the distribution of data, making it highly adaptable to complex and irregular datasets.
Arbitrary Cluster Shapes: Mean Shift can identify clusters of various shapes and sizes, unlike algorithms that assume spherical or other predefined structures.
Robust to Noise: The algorithm is relatively insensitive to noise, as it shifts data points based on density rather than strict distance calculations, making it suitable for real-world datasets with imperfections.

1.6 Applications of Mean Shift Clustering:

Mean Shift clustering finds applications in various domains, including:
Image Segmentation: Identifying distinct regions within images based on pixel characteristics.
Object Tracking: Tracking the movement of objects in video sequences.
Customer Segmentation: Grouping customers based on their purchase behavior or browsing patterns.
Anomaly Detection: Identifying outliers and unusual patterns in data.

The ultimate goal of utilizing Mean Shift clustering in real-time e-commerce analysis is to unlock the power of personalized experiences, ultimately driving improved business outcomes. By understanding and responding to customer behavior as it happens, businesses can create a more engaging, relevant, and satisfying shopping journey, leading to increased customer loyalty, higher conversion rates, and ultimately, greater profitability.

1.7 The Promise of Real-Time Personalization Enhanced Customer Engagement:

Real-time personalization, driven by Mean Shift clustering, allows businesses to deliver experiences tailored to individual customer needs and preferences, boosting engagement and driving repeat purchases [5, 6]. Optimized Revenue: By understanding and reacting to customer behavior in real-time, businesses can optimize pricing strategies, personalize product recommendations, and target marketing campaigns more effectively, leading to increased sales and revenue generation [3, 4].
Improved Customer Satisfaction: Real-time insights allow businesses to anticipate customer needs and deliver a more personalized and seamless shopping experience, fostering customer loyalty and reducing churn [9, 10].
Competitive Advantage: By leveraging real-time data and personalization, businesses can gain a competitive edge, differentiating themselves from competitors and creating a more compelling customer experience [1, 2].

2. Mean Shift Clustering for Real-Time E-commerce Analysis

Mean Shift clustering is a non-parametric, iterative algorithm that identifies clusters in data based on density estimation. It operates by iteratively shifting data points towards regions of higher density, ultimately converging to local maxima in the density function, which represent the centers of clusters.
Here's a detailed explanation of the key concepts:

2.1. Kernel Density Estimation:

The Kernel Function(K(x)): The Mean Shift algorithm relies on a kernel function, typically a Gaussian kernel, to estimate the density of data points. The kernel function assigns weights to data points based on their proximity to a given point. Data points closer to the target point receive higher weights, while those further away receive lower weights. where x is the distance from the data point to the location where we are estimating density and h is the bandwidth, controlling the smoothness of the density estimate (larger h means smoother)

$$K(x) = (1 / (\sqrt{(2\pi)}h)) * \exp(-(x^2) / (2h^2)) \quad (1)$$

Density Estimation(f_h(x)): By applying the kernel function to each data point, the algorithm creates a weighted density estimate for each point in the data space. This density estimate represents the probability of finding a data point at a particular location. Where n is the number of data points in the dataset, h is the bandwidth and K is the Kernel function.

$$f_h(x) = (1 / (nh)) * \Sigma(i=1 \text{ to } n)\, K((x - x_i) / h) \quad (2)$$

2.2. Mean Shift Vector (m(x_i)):

For each data point, the algorithm calculates a Mean Shift vector, which points in the direction of the gradient ascent of the density function. In simpler terms, the Mean Shift vector points towards the direction of increasing density. Each data point is then shifted along its Mean Shift vector. This shift moves the data point towards a region of higher density.

$$m(x_i) = (1 / (nh)) * \Sigma(j=1 \text{ to } n) (x_j - x_i) * K((x_i - x_j) / h) \quad (3)$$

2.3. Iterative Shifting and Convergence:

The Mean Shift algorithm iteratively calculates the Mean Shift vector for each data point and shifts them along the vector until they converge to a local maximum in the density function. This means that the data points will eventually settle in regions of high density, representing the centers of clusters. Data points that converge to the same local maximum are grouped together, forming a cluster. This process identifies clusters of arbitrary shapes, unlike algorithms that rely on predefined structures. Where $x_i(t)$ is the location of data point i at iteration t. $m(x_i(t))$ is the mean shift vector calculated for data point i at iteration t and η is a small step size, controlling how much the data point is moved in each iteration.

$$x_i(t+1) = x_i(t) + \eta * m(x_i(t)) \quad (4)$$

Mean Shift clustering relies on density estimation to identify clusters within a dataset. This process involves determining the likelihood of finding data points in different regions of the data space. The algorithm achieves this by utilizing a kernel function, which assigns weights to data points based on their proximity to a given point.

Here's a breakdown of how density estimation works within the Mean Shift algorithm:

The kernel function acts as a weighting function, assigning higher weights to data points closer to the target point and gradually diminishing the weight as the distance increases. This means that points near the target point contribute more to the density estimate, while those further away have a smaller influence. A variety of kernel functions can be used, but the most common choice is the Gaussian kernel. The Gaussian kernel produces a bell-shaped curve around each data point, giving more weight to points near the center of the curve and decreasing the weight as you move further away. Other popular choices include the Epanechnikov kernel and the uniform kernel.

For each data point, the algorithm calculates a weighted average of the data points within a specified neighborhood around that point. The weights are determined by the kernel function, giving more weight to nearby points and less to those further away. This weighted average effectively estimates the density of the data at that specific point. Regions with a higher concentration of data points, as indicated by a higher density estimate, are more likely to be cluster centers. Imagine a scattered distribution of points in a two-dimensional space. Applying a kernel function to each point generates a smooth, bell-shaped curve centered around that point. The resulting density map shows regions of high density, representing potential cluster centers, and regions of low density, indicating areas with fewer data points. The Mean Shift vector is the heart of the Mean Shift clustering algorithm. It acts as a compass, guiding data points towards regions of higher density, ultimately leading to the formation of clusters. Here's how it works

The Mean Shift vector is calculated for each data point in the dataset. It points in the direction of the steepest ascent of the density function, effectively guiding the data point towards a region of higher density. This shift is crucial for identifying clusters, as data points converge to local density maxima, representing the centers of clusters.

The calculation of the Mean Shift vector involves two key steps: A specific neighborhood around the target data point is defined. This neighborhood is typically defined by a radius or bandwidth parameter, which determines the range of influence for nearby points. The Mean Shift vector is calculated by taking the weighted average of the data points within the defined neighborhood. The weights are determined by the kernel function, which assigns higher weights to data points closer to the target point and gradually diminishes the weights as the distance increases.

Imagine a scatter plot of data points. The Mean Shift vector for a specific data point can be visualized as an arrow pointing towards the direction of higher density. If the data point is in a sparse area, the Mean Shift vector might point towards a denser region nearby. As data points are iteratively shifted along their Mean Shift vectors, they eventually converge to regions of high density, forming clusters.

The Mean Shift algorithm, a non-parametric clustering technique, can be leveraged in real-time to create highly personalized customer experiences by identifying distinct groups of customers with similar preferences, needs, and purchasing patterns. This allows businesses to tailor recommendations, promotions, and content for each segment.

By analyzing the trajectory of customer behavior within these clusters, Mean Shift can predict future actions and preferences, enabling proactive engagement and preemptive personalization. This real-time insight optimizes marketing campaigns by targeting specific segments with personalized messages and offers, maximizing campaign effectiveness. Ultimately, by providing personalized recommendations, relevant content, and tailored interactions, Mean Shift enhances the overall customer experience, leading to increased satisfaction, loyalty, and ultimately improved business outcomes.

3. Real-Time Applications in E-commerce

Mean Shift, with its ability to identify clusters in real-time, can dynamically segment customers based on their evolving browsing behavior, search queries, product interactions, and purchase patterns [11-15]. This creates a constantly evolving understanding of customer preferences and needs, enabling businesses to tailor their offerings accordingly.

Examples of Real-Time Customer Segments:

"Impulsive Shoppers": Identified by quick browsing, immediate purchases, and high-value transactions.

"Research-Oriented Browsers": Characterized by extensive browsing, comparison shopping, and delayed purchases.

"Loyal Customers": Defined by consistent repeat purchases, high engagement, and positive reviews.

"New Customers": Identified by initial browsing activity, product exploration, and potential purchase intent.

By analyzing customer behavior in real-time, Mean Shift can identify related products based on their browsing history, product comparisons, and interactions within identified clusters [11-15]. This real-time insight allows for highly accurate and effective product recommendations that resonate with individual customer preferences and needs, enhancing user experience and driving sales.

Real-time clustering can pinpoint price-sensitive customer segments, allowing businesses to dynamically adjust pricing strategies based on individual customer behavior and purchase patterns [15]. This can involve personalized discounts, tiered pricing models, and dynamic price adjustments based on real-time market conditions and customer willingness to pay.

Real-time customer segments can be used to target marketing messages and promotions for maximum impact [11-14]. By analyzing the real-time behavior of different segments, businesses can personalize messaging based on current interests, purchase history, and browsing patterns. This approach increases engagement, conversion rates, and overall campaign effectiveness.

Mean Shift can be employed to detect anomalous activity in real-time, potentially identifying fraudulent transactions [15]. By analyzing the deviations in customer behavior patterns from established clusters, Mean Shift can flag suspicious activities, enabling immediate intervention and reducing financial losses.

4. Real-Time E-commerce Case Studies

The power of Mean Shift in real-time analysis is being harnessed across various e-commerce industries, leading to significant improvements in customer engagement, sales, and revenue. Here are a few illustrative examples:

A leading online fashion retailer uses Mean Shift to analyze customer browsing behavior, purchase history, and product interactions in real-time. This allows them to recommend relevant items, such as complementary accessories, similar styles, and trending pieces, directly to individual customers, leading to higher conversion rates and average order value. A fashion platform utilizes Mean Shift to understand customer style preferences based on their browsing history and saved items. The platform then dynamically suggests outfits and styling tips tailored to each user's unique taste, enhancing engagement and driving sales of complementary items.

A travel booking platform leverages Mean Shift to segment travelers based on their browsing behavior and search queries. This allows them to offer personalized travel packages, including flight and hotel recommendations, tailored to individual needs and preferences, like family-friendly options, adventurous experiences, or luxury getaways. A hotel booking website uses Mean Shift to identify price-sensitive customer segments based on their booking patterns and price comparisons. This allows them to dynamically adjust prices for individual travelers, maximizing revenue while ensuring competitive pricing.

A supermarket chain uses Mean Shift to identify customer segments based on their purchase history and shopping patterns. This enables them to send targeted promotions, such as personalized discounts and offers, for products that align with individual customer preferences, increasing customer engagement and driving loyalty. An online retailer uses Mean Shift to analyze real-time sales data and predict product demand. This helps them optimize inventory levels and ensure that popular items are always readily available, minimizing stockouts and maximizing sales.

These examples demonstrate the diverse applications of Mean Shift in e-commerce, showcasing its ability to enhance customer engagement, optimize pricing strategies, drive revenue growth and customer satisfaction. By leveraging real-time insights and personalized experiences, businesses are creating a more engaging and profitable customer journey. Figure 2. and figure 3 represent real-time E-Commerce data with columns such as timestamp, customer_id, product_id, product_category, order_value, and event_type. The Mean Shift clustering algorithm is a technique used to group data points into clusters without needing to specify the number of clusters in advance. It works by treating each data point as a probability distribution and moving points towards the region of highest data density, also known as the

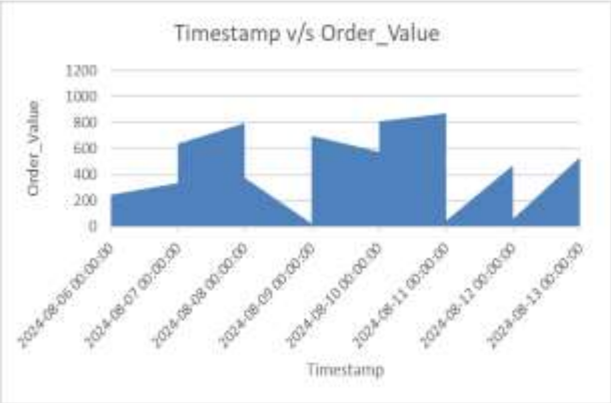

Figure 2. *Real Time E-Commerce data with Timestamp v/s Order_Value*

Figure 3. *Real Time E-Commerce Data with Product_Id v/s Order_Value*

"mean shift." This process continues until the data points converge to the densest areas, which become the cluster centers. Mean Shift is particularly useful for identifying clusters of various shapes and sizes in data. It's a versatile method often used in image processing, computer vision, and other fields where the structure of the data isn't well-defined. Figure 4. shows E-Commerce data using Mean Shift Clustering.

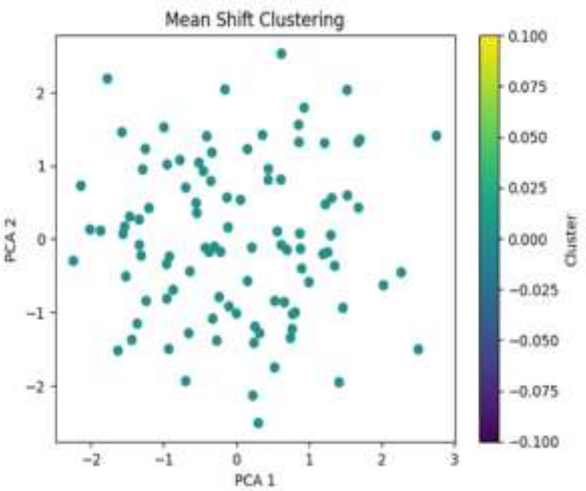

Figure 4. *E-Commerce data with Mean Shift Clustering*

5. Challenges and Future Directions

While Mean Shift offers a powerful tool for real-time personalization, its successful implementation necessitates addressing several critical challenges:

Data Streaming and Scalability: Processing large volumes of real-time customer data efficiently demands advanced data streaming architectures and scalable algorithms [11-14]. This requires robust systems capable of handling high throughput, low latency, and continuous data ingestion without sacrificing accuracy or performance.

Computational Complexity: Real-time analysis imposes stringent computational demands, necessitating optimized implementations of Mean Shift to minimize processing time and ensure responsiveness [15]. This involves exploring efficient data structures, parallel processing techniques, and potentially leveraging specialized hardware like GPUs to accelerate calculations.

Data Privacy: The use of real-time customer data raises significant privacy concerns. Implementing robust anonymization techniques and secure data handling protocols is crucial to protect sensitive information while leveraging valuable insights [16]. Businesses must prioritize data privacy and comply with relevant regulations like GDPR and CCPA to ensure ethical and responsible data usage.

Explainable AI: As Mean Shift models operate in real-time, making them interpretable is crucial to understand the reasoning behind decisions, build trust in the system, and enable informed decision-making [17,18]. This requires developing techniques for visualizing and explaining the model's clustering behavior, providing insights into the factors influencing recommendations and price adjustments.

By addressing these challenges, businesses can harness the potential of Mean Shift for real-time personalization while ensuring ethical and responsible data practices. This will allow them to leverage the power of this algorithm for greater customer engagement, revenue generation, and improved business outcomes.

6. Conclusion

Mean Shift clustering offers a powerful approach for real-time analysis in e-commerce, enabling businesses to dynamically understand customer behavior, personalize experiences, and optimize operations. Its ability to adapt to changing data patterns, handle noisy data, and generate insights in real-time makes it a valuable tool for

driving customer engagement, maximizing revenue, and staying ahead in the dynamic e-commerce landscape. As e-commerce evolves, the integration of Mean Shift clustering for real-time analysis will become increasingly crucial for businesses seeking to thrive in this competitive and data-driven environment [19-21].

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, "Multiverse recommendation: N-dimensional recommendation based on collaborative filtering," *in Proceedings of the 4th ACM Conference on Recommender Systems*, pp. 71-78, 2010.

[2] G. Adomavicius and A. Tuzhilin, (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, 17(6);734-749. DOI: 10.1109/TKDE.2005.99

[3] A. Elmachtoub and P. Grigas, "Smart pricing: A dynamic programming approach," Management Science, vol. 66, no. 8, pp. 3291-3310, 2020.

[4] Reimers, Imke, and Joel Waldfogel. (2021). Digitization and Pre-purchase Information: The Causal and Welfare Impacts of Reviews and Crowd Ratings. *American Economic Review*, 111(6);1944–71 DOI: 10.1257/aer.20200153

[5] Ghandour, Ahmad et al. (2023). A Framework for Relationships in eCommerce Websites. *Int. J. E Bus. Res.* 19;1-17. DOI: 10.4018/IJEBR.317411

[6] Roland T. Rust, Katherine N. Lemon et al. (2004). Return on Marketing: Using Customer Equity to Focus Marketing Strategy, *Journal of Marketing*, 68(1);109-127. https://doi.org/10.1509/jmkg.68.1.109.24030

[7] M. A. Zaki, (2014). Data mining and analysis: Fundamental concepts and algorithms. Wiley, 2014.

[8] J. Han, J. Pei, and M. Kamber (2011), Data mining: Concepts and techniques, *Morgan Kaufmann*.

[9] J. C. Anderson and J. A. Narus, (2004). Business market management: Understanding, creating, and managing value, *Prentice Hall*.

[10] J. Kotler and K. L. Keller, (2015). Marketing management, *Pearson Education*.

[11] Zhao M, Jha A, Liu Q, Millis BA, Mahadevan-Jansen A, Lu L, Landman BA, Tyska MJ, Huo Y. (2021). Faster Mean-shift: GPU-accelerated clustering for cosine embedding-based cell segmentation and tracking. *Med Image Anal.* 71;102048. doi: 10.1016/j.media.2021.102048.

[12] Zhao M, Chang CH, Xie W, Xie Z, and Hu J, (2020). Cloud shape classification system based on multi-channel cnn and improved fdm, *IEEE Access*, 8;44111–44124. DOI: 10.1109/ACCESS.2020.2978090

[13] Kulikov V and Lempitsky V, (2020). Instance segmentation of biological images using harmonic embeddings," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3843–3851.

[14] Chen L, Strauch M, and Merhof D, (2019) "Instance segmentation of biomedical images with an object-aware embedding learned with local constraints," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 451–459.

[15] Chen F, Zhang T, Liu R. (2021). An Active Learning Method Based on Variational Autoencoder and DBSCAN Clustering. *Comput Intell Neurosci.* 2021:9952596. doi: 10.1155/2021/9952596.

[16] "GDPR - General Data Protection Regulation," European Union.

[17] Lipton, Z. C. (2018). The myth of model interpretability. *Queue*, 16(3);31-57.

[18] Hyrien O, Baran A. (2016). Fast Nonparametric Density-Based Clustering of Large Data Sets Using a Stochastic Approximation Mean-Shift Algorithm. *J Comput Graph Stat.* 25(3);899-916. doi: 10.1080/10618600.2015.1051625.

[19]DAGLI, A., & DOGAN, S. Özhan. (2023). e-Commerce Return Processes and Data Analysis in Logistics. *International Journal of Computational and Experimental Science and Engineering*, 9(4), 340–345. Retrieved from https://www.ijcesen.com/index.php/ijcesen/article/view/276

[20]HAFIZOĞLU GÖKDAĞ, Z., & BİLGE, A. H. (2023). The Impact of Dynamic Shocks and Special Days on Time Series Data. *International Journal of Computational and Experimental Science and Engineering*, 9(2), 183–190. Retrieved from https://www.ijcesen.com/index.php/ijcesen/article/view/208

[21]Venkatraman Umbalacheri Ramasamy. (2024). Overview of Anomaly Detection Techniques across Different Domains: A Systematic Review. *International Journal of Computational and Experimental Science and Engineering*, 10(4)898-910. https://doi.org/10.22399/ijcesen.522