Disruptive Human Resource Management
S. Patnaik (Ed.)

doi:10.3233/ATDE240434

Market Segmentation and Personalized Marketing Strategy Optimization Driven by Big Data Analysis

Zifeng Zhang[1]
Shijiazhuang Vocational and Technical College, Shijiazhuang, Hebei 050081, China

Abstract. Market segmentation and personalized marketing strategies have become crucial for businesses in today's highly competitive landscape. With the increasing availability of customer data and advancements in technology, it is essential to explore the integration of big data analysis and neural networks to optimize market segmentation and evaluate the effectiveness of personalized marketing strategies. This work combines neural network to construct a big data-driven model to evaluate market segmentation and personalized marketing optimization strategies. To begin, this study enhances the classic whale optimization algorithm (WOA) by designing the Improved WOA algorithm with three key techniques for enhancement. Second, to address the shortcomings of deep belief network (DBN), this study employs the IWOA method to fine-tune the network's founding parameters before construction. Lastly, this work uses large data collection techniques to gather relevant information for training IWOA-DBN. IWOA-DBN is used to build a model for assessing approaches to market segmentation and individualized marketing optimization. In this study, the developed approach is experimentally tested in a systematic manner, and the findings demonstrate the method's superiority.

Keywords. Market segmentation, personalized marketing, WOA, DBN

1. Introduction

In today's dynamic and competitive business landscape, companies are continuously seeking innovative strategies to gain a competitive edge and drive sustainable growth [1]. One key aspect that has emerged as a critical success factor is market segmentation and personalized marketing. Market segmentation involves dividing a broad market into smaller, more manageable segments based on shared characteristics, needs, and preferences [2]. By understanding and targeting specific segments, businesses can tailor their marketing efforts to meet the unique requirements and desires of each group [3]. This targeted approach allows companies to develop more persuasive and compelling marketing messages, leading to higher customer engagement, conversion rates, and overall business performance.

The advent of big data has revolutionized the way businesses operate and make strategic decisions. With the proliferation of digital technologies and channels, companies now have access to an unprecedented amount of customer data [4]. This data encompasses various touchpoints, ranging from online interactions, social media

[1] Corresponding Author: Zifeng Zhang, zzf_123452023@163.com.

engagement, to transactional histories and customer feedback [5]. Such a wealth of information offers valuable insights into customer behavior, preferences, and patterns, driving the need for advanced analytics techniques to make sense of this vast dataset [6]. One such technique that has gained significant attention is the integration of big data analysis with neural networks. Neural networks, a subfield of artificial intelligence (AI), are powerful computational models inspired by the workings of the human brain [7]. These networks can analyze and process complex patterns and relationships within data, enabling businesses to uncover hidden insights and make accurate predictions [8].

By combining the capabilities of big data analysis with neural networks, companies can unlock valuable insights about their customer segments [9]. Sophisticated algorithms analyze diverse data sets, identifying patterns, behaviors, and preferences within each segment. This enables businesses to develop a comprehensive understanding of their target audience and customize their marketing strategies accordingly [10-11]. Personalized marketing, driven by market segmentation and data-driven insights, has proven to be highly effective in achieving customer satisfaction and fostering brand loyalty [12]. When customers receive targeted and relevant marketing messages, they perceive a deeper connection with the brand, leading to increased engagement, conversion rates, and ultimately, revenue growth [13]. Furthermore, the integration of big data analysis and neural networks enables businesses to continuously evaluate and optimize their marketing efforts. By leveraging machine learning algorithms, companies can analyze campaign performance metrics, customer feedback, and market trends, identifying areas for improvement and driving more impactful marketing strategies [14].

In this paper, we aim to delve deeper into the integration of big data analysis and neural networks for market segmentation and personalized marketing. We will explore the methodologies, techniques, and real-world applications of these approaches, showcasing their potential to transform marketing strategies and drive business success. By providing practical insights and highlighting industry best practices, we intend to enable businesses to harness the power of big data and AI-driven techniques to enhance their market segmentation approaches and deliver personalized marketing experiences. This work combines neural network to construct a big data-driven model to evaluate market segmentation and personalized marketing optimization strategies. To begin, this study enhances the classic whale optimization algorithm (WOA) by designing the Improved WOA algorithm with three key techniques for enhancement. Second, to address the shortcomings of deep belief network (DBN), this study employs the IWOA method to fine-tune the network's founding parameters before construction. Lastly, this work uses large data collection techniques to gather relevant information for training IWOA-DBN. IWOA-DBN is used to build a model for assessing approaches to market segmentation and individualized marketing optimization.

2. Method

2.1. Deep Belief Network Algorithm

DBN is a classic deep learning model since its RBM can be decomposed into a probabilistic generation model. In the DBN, the input layer employs RBM to convert the data from the feature vectors of the nonlinear process into the language of the hidden layers. Data features are extracted while the network constantly maps from the hidden layer's lower to upper layers. RBM is the cornerstone of the DBN model architecture; it is an energy-based probabilistic model with a Boltzmann-type sample distribution. It is feasible to represent every statistical probability distribution in terms of energy. This is why RBM is so helpful; it can generate a learning model for the data of interest without our knowing the underlying distribution of the relevant statistical variables. Connections between the visible and invisible layers are only given bidirectional weights in the RBM model. An RBM has two layers, one that is exposed and one that is hidden, with no link weights between units in the same layer. The RBM framework is shown in Figure 1.

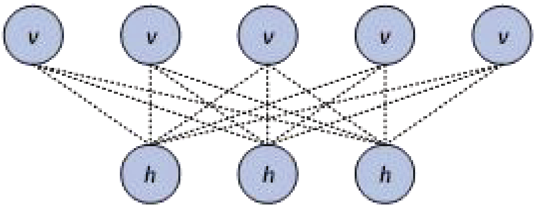

Figure 1. RBM structure.

The joint state energy function between each pair of visible layer neurons and hidden layer neurons may be computed given the RBM's visible layer, hidden layer vector, weight, and bias as Equation (1).

$$E =- \sum_{i=1}^{m} a_i v_i - \sum_{j=1}^{n} b_j h_j - \sum_{i=1}^{m} \sum_{j=1}^{n} v_i w_{ij} h_j \tag{1}$$

where w is weight, a and b are bias.

When the energy function is at its minimum, the RBM network structure is at its most efficient configuration for a given distribution of probabilities. The RBM model needs to be trained over and over again until it reaches its ideal state, which is done by decreasing the energy function. Both the hidden and the revealed layers' combined probability distributions may be calculated as Equation (2) and Equation (3).

$$P(v, h) = \frac{e^{-E(v,h)}}{Z} \tag{2}$$

$$Z = \sum_v \sum_h e^{-E(v,h)} \tag{3}$$

where E is energy.

The single-layer network topology of DBN can only learn simple characteristics since it is considered to be a superposition of several RBMs. Therefore, deep features may be extracted from complicated data by stacking numerous RBMs. Classification and regression analyses benefit from these derived deep features. However, data categorization and regression cannot be performed directly by stacking RBMs. Adding a regression layer or a classification layer on top of the stacked RBM is required to obtain the entire DBN model. Figure 2 depicts the DBN's graphical design.

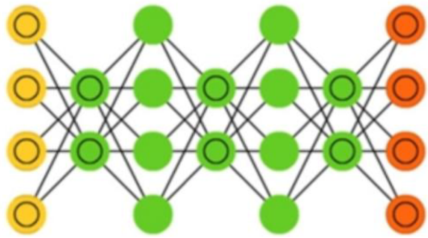

Figure 2. DBN structure.

For this purpose, we employ a multi-RBM and multi-classification-layer DBN. The first RBM is supplied a vector of raw sample data via its visible layer. It is then discovered that the original RBM's settings for the visible and hidden layers should remain constant. A new layer of information is learned by both RBMs while their parameters are locked in place; the output of the first RBM's hidden layer is fed into the output of the second RBM's visible layer. The item layer follows the continuous RBM as the classification layer, where the softmax function is used to translate the features produced by the last RBM to the category labels that best describe the sample data. Stacked RBMs mirror DBN, and more RBMs allow for more efficient extraction of abstract characteristics from sample data. Layer-by-layer training and learning can also guarantee roughly optimum settings for each layer. At this point, the DBN as a whole has model parameters that are suboptimal, necessitating more training.

2.2. WOA pipeline

DBN's random initialization of weights and thresholds raises the possibility that the network may not converge to the global optimum. Therefore, the WOA method is used to set the starting values in this study. When hunting for food, humpback whales employ a unique bubble net in the shape of a spiral. From a depth of roughly 15 meters, humpback whales swim up to the surface in a spiraling motion. In this way, a network of bubbles powerful enough to encircle the prey can rise to the surface at almost the same time as the first bubbles. By using this tactic, humpback whales may herd their food into the center of the bubble net, where they can quickly and easily consume it by diving under the school and racing towards the center. The humpback whale will take over the task of releasing bubbles throughout the predation process, allowing the other members of the team to employ the same tactic.

The proposal of WOA was motivated by the hunting and foraging habits of humpback whales. The three primary phases of WOA are the "surround" phase, "bubble net" phase, and the "global search" phase. To model whale swarms' foraging behavior, the algorithm includes an exploration phase in which whales follow the leader on a random, worldwide hunt for food. The whale swarm can take one of two approaches during the mining phase of the algorithm. The whales swarm closer to the prey as the gap between them and the leader decreases. Another method is spiral bubble net hunting, in which the predator swims in a spiral pattern toward the target while blowing bubbles. The standard whale optimization technique may now achieve global optimization performance thanks to the power of global search. The whale pod is in a condition of random walking during the global search phase since no one knows where the prey is. Right now, the whale community utilizes the location data of randomly selected whales to navigate to new feeding grounds. The detailed calculation are from Equation (4) to Equation (7).

$$X(t+1) = X_{rand} - AD_{rand} \tag{4}$$

$$D_{rand} = |cX_{rand} - X(t)| \tag{5}$$

$$A = 2ar - a \tag{6}$$

$$c = 2r \tag{7}$$

During the encirclement and predation phase, the prey's location is unclear at the outset of the hunt, therefore the whale pod adjusts its position in relation to the leader. As the quantity of prey increases, their location becomes more apparent, and the whale pod adjusts its position accordingly. The detailed calculation are from Equation (8) to Equation (9).

$$X(t+1) = X^{*}(t) + AB \tag{8}$$

$$B = |cX^{*}(t) - X(t)| \tag{9}$$

where $X^{*}(t)$ is the optimal solution.

When using bubble nets to catch prey, humpback whales have essentially figured out how far apart individual whales are. In the spiral bubble net hunting technique, individual whales approach their prey from below by swimming in a spiraling upward manner, spitting bubbles as they go. Both encircling predation and spiral bubble net predation are typical components of the hunting mechanism during the actual hunting activity. To model the hunting behavior of whales that perform encircling predation and spiral bubble net predation simultaneously, a random variable is constructed to decide between the two location update procedures at predetermined intervals. When the likelihood is high enough, the pod adjusts its hunting location in accordance with its spiral bubble net predation method. When the chance of finding food is low enough, the whale group adjusts its hunting stance to take advantage of its surroundings.

2.3. IWOA-DBN

WOA has difficulties, such as a poor convergence time and imprecise solution accuracy, when applied to challenging global optimization problems. This effort has resulted in several improvements to the IWOA design. Swarm intelligence employs a population-based iterative search method, the efficacy of which is proportional to the quality of the initial population. The algorithm's global convergence time and the quality of the solution can both benefit from a more diversified starting population. The WOA approach is used to find the optimal solution to an optimization issue without any prior knowledge of that solution. As a result of the above, it is customary to employ a random generation strategy when constructing the initial population. However, the variety and even distribution of the starting population created by stochastic approaches are also low. This reduces the search efficiency since it is unable to collect relevant data from the solution space. To sum up, chaos is a nonlinear phenomena that may be both chaotic and predictable. In order to establish an initial population, this article uses a chaotic sequence. A wide variety of chaotic models can produce chaotic sequences. In this work, we choose the Skew Tent mapping approach to provide chaotic sequences for population startup. The detailed calculation are from Equation (10) to Equation (11).

$$x_{k+1} = \frac{x_k}{\varphi}, 0 < x_k < \varphi \tag{10}$$

$$x_{k+1} = (1 - x_k)/(1 - \varphi), \varphi < x_k < 1 \tag{11}$$

where $\varphi \in (0,1)$.

Intelligent optimization methods that use iterative population evolution carry out both global and local searches. The convergence rate of the evolutionary phase of the algorithm might be slowed down by a lack of coordination between these two types of operations, or it can converge too fast. Finding anything will need organizations to broaden their searches to include more of the globe. By doing so, we maintain a genetically diversified population and prevent the algorithm from getting stuck at a local optimum. To speed up the convergence of the process, the swarm can do a local search, where it zeroes in on a relatively small region of the solution space. If we want to make our unique swarm intelligence optimization method even more efficient, we need to find the optimal balance between global and local search. When it comes to both global and local search, the WOA algorithm relies heavily on the convergence factor. Because of this, it is essential to fine-tune the value of the convergence factor so that the WOA algorithm's global search capacity is well-balanced with its local search ability. The convergence factor in WOA, on the other hand, decreases linearly with time. An initial greater value allows the algorithm to discover a bigger search region and eventually leads to a global search since the algorithm is satisfied. Over time, the algorithm will improve to the point where local accurate search may be performed with a relatively small value for the convergence factor. Due to nonlinear changes that occur throughout the evolutionary search phase of the WOA technique, a linearly dropping convergence factor strategy is not a good representation of the optimal search process. Based on the inertia weight configuration of the PSO algorithm, this study proposes a nonlinear convergence factor approach. The detailed calculation is Equation (12).

$$a = (a_i - a_f)(t_{max} - t)/t_{max} \tag{12}$$

where a_i and a_f are initial and end value.

Similar to other swarm intelligence optimization techniques, the fundamental WOA algorithm has a shortcoming. In particular, late-stage iterations provide a high risk of being mired in a local optimum. This article addresses this limitation by applying chaotic disturbance to the present optimum individual, which combines the ergodicity, regularity, and unpredictability of the chaotic sequence. This can cause the WOA algorithm to escape the local optimum, improving its search efficiency and precision. The detailed calculation is Equation (13).

$$y_k^{'} = (1 - \tau)y^* + \tau y_k \tag{13}$$

where y^* and y_k are chaotic variable.

The method begins by setting up the DBN and the necessary IWOA settings. After that, the relevant IWOA operation is carried out based on the results of the fitness calculation. Parameter values are then assigned to the DBN's initial weights and thresholds when the termination criteria have been satisfied. Finally, the DBN is trained until the goal is achieved.

3. Result

This job begins with the data collection stage. Table 1 shows the unique data distribution and how the dataset's training samples and test samples differ in terms of data amount.

Table 1. Data distribution of dataset

Item	Quantity
Training set	30225
Test set	16521
Total set	46746

In this study, we compare IWOA-DBN against other competing approaches to demonstrate IWOA-DBN's superiority. Experimental data is depicted in Figure 3 to make the parameter settings as uniform as feasible and preserve the validity of the comparison.

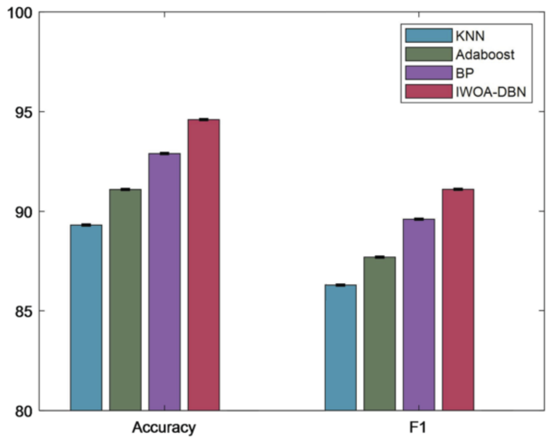

Figure 3. Comparison with different methods.

IWOA-DBN has superior performance than other conventional approaches. Skew Tent mapping is used to produce the prototype population in this study. In Figure 4, we can see how this mapping approach compares to the random mapping method in terms of IWOA-DBN performance.

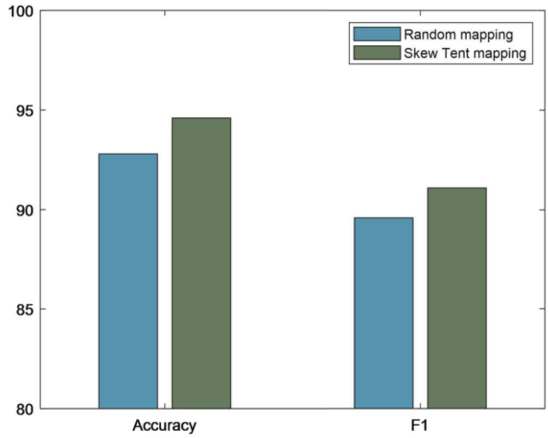

Figure 4. Comparison of mapping methods.

Skew Tent mapping improves model accuracy and F1 score compared to random mapping. The reliability of Skew Tent mapping has been confirmed. The nonlinear convergence factor is used as a controller in this study. In Figure 5, we see a comparison between the convergence factor used in this study and the linear convergence factor used in previous work to evaluate the performance of IWOA-DBN.

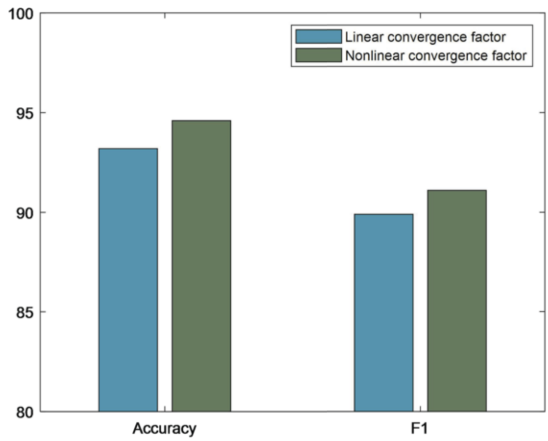

Figure 5. Comparison of convergence factor.

Nonlinear convergence factors enhance model accuracy and F1 score beyond what is possible with linear convergence factors. Nonlinear convergence resilience is proven by this test.

4. Conclusion

In order to analyze market segmentation and tailored marketing optimization tactics, this work uses neural networks to build a large data-driven model. To begin, the original WOA is improved in this paper by incorporating three new strategies into a new algorithm called the Improved WOA. Second, this research uses the IWOA technique to fine-tune the network's initial parameters before formation, therefore mitigating the drawbacks of DBNs. The information used to train IWOA-DBN is gathered through the use of huge data gathering techniques, which is the final contribution of this study. A model for evaluating strategies for market segmentation and personalized marketing optimization is constructed using IWOA-DBN. In order to analyze market segmentation and tailored marketing optimization tactics, this work uses neural networks to build a large data-driven model. To begin, the original whale optimization algorithm (WOA) is improved in this paper by incorporating three new strategies into a new algorithm called the Improved WOA. Second, this research uses the IWOA technique to fine-tune the network's initial parameters before formation, therefore mitigating the drawbacks of deep belief networks (DBNs). The information used to train IWOA-DBN is gathered through the use of huge data gathering techniques, which is the final contribution of this study. A model for evaluating strategies for market segmentation and personalized marketing optimization is constructed using IWOA-DBN. With the continuous growth of big data and AI technologies, these approaches are expected to play an increasingly significant role in driving successful marketing strategies and achieving business success.

References

[1] Kumar A, Shankar R, Aljohani NR. A big data driven framework for demand-driven forecasting with effects of marketing-mix variables. Industrial marketing management. 2020; 90: 493-507.
[2] Li L, Zhang J. Research and analysis of an enterprise E-commerce marketing system under the big data environment. Journal of Organizational and End User Computing (JOEUC). 2021; 33(6): 1-19.
[3] Nti IK, Adekoya AF, Weyori BA. A novel multi-source information-fusion predictive framework based on deep neural networks for accuracy enhancement in stock market prediction. Journal of Big data. 2021; 8(1): 1-28.
[4] Ghimire A, Thapa S, Jha A K. Accelerating business growth with big data and artificial intelligence. 2020 Fourth International Conference on I-SMAC; 2020: p. 441-448.
[5] Mumtaz S., Huq KMS, Radwan A, Rodriguez J, Aguiar R L. Energy efficient interference-aware resource allocation in LTE-D2D communication. In 2014 IEEE International Conference on Communications; 2014; p. 282-287.
[6] Kumar D T S. Data mining based marketing decision support system using hybrid machine learning algorithm. Journal of Artificial Intelligence and Capsule Networks. 2020; 2(3): 185-193.
[7] Yoseph F, Ahamed Hassain Malim NH, Heikkilä M, Brezulianu A, Geman O, Paskhal Rostam NA. The impact of big data market segmentation using data mining and clustering techniques. Journal of Intelligent & Fuzzy Systems. 2020; 38(5): 6159-6173.
[8] Li J, Li S, Cheng L, Liu Q, Pei J, Wang S. BSAS: A Blockchain-Based Trustworthy and Privacy-Preserving Speed Advisory System. IEEE Transactions on Vehicular Technology. 2022; 71(11): 11421-11430.
[9] Chang MS, Kim HJ. A customer segmentation scheme base on big data in a bank. Journal of Digital Contents Society. 2018; 19(1): 85-91.
[10] Alghamdi A. A hybrid method for customer segmentation in Saudi Arabia restaurants using clustering, neural networks and optimization learning techniques. Arabian Journal for Science and Engineering. 2023; 48(2): 2021-2039.
[11] Mumtaz S, Lundqvist H, Huq KMS, Rodriguez J, Radwan, A. Smart Direct-LTE communication: An energy saving perspective. Ad Hoc Networks. 2014; 13: 296-311.

[12] Serrano W. Neural networks in big data and Web search. Data. 2018; 4(1): 7.
[13] Kauko TOM, Hooimeijer P, Hakfoort J. Capturing housing market segmentation: An alternative approach based on neural network modelling. Housing Studies. 2002; 17(6): 875-894.
[14] Pei J, Zhong K, Li J, Yu Z. PAC: partial area clustering for re-adjusting the layout of traffic stations in city's public transport. IEEE Transactions on Intelligent Transportation Systems. 2022; 24(1): 1251-1260.