Neural-KMeans-Based Food Feature Modeling for Personalized Nutrition Plans

Authors: [1] Cheikh Cisse, [2] Syed Rajab Ali Shah

[1] **Corresponding Author:** cissec07@gmail.com

Abstract:

Personalized nutrition represents a transformative approach in health science that aims to tailor dietary recommendations based on individual characteristics, preferences, and metabolic responses. As the demand for precision health management increases, the integration of artificial intelligence and machine learning techniques in nutrition science has become pivotal. This paper introduces a hybrid model combining Neural Networks and K-Means Clustering (Neural-KMeans) to construct an effective food feature modeling system aimed at developing personalized nutrition plans. The proposed system processes dietary data, extracts complex nutritional features, and clusters user-specific dietary patterns to recommend optimized nutrition strategies. The integration of neural networks allows the model to capture non-linear relationships in nutritional attributes, while K-Means aids in categorizing food types and user consumption habits. An experimental framework was established using a real-world dietary dataset, and results demonstrate significant improvement in accuracy and personalization compared to traditional models. The study underscores the potential of Neural-KMeans modeling in revolutionizing dietary planning and facilitating health-conscious decision-making.

Keywords: Personalized Nutrition, Neural Networks, K-Means Clustering, Food Feature Modeling, Dietary Planning, Health Informatics

[1] Azzurra Formazione, Italy

[2] COMSATS University Islamabad, Pakistan

I. Introduction

In recent years, the concept of personalized nutrition has gained widespread attention owing to its potential to significantly enhance individual health outcomes by aligning dietary intake with personal health parameters [1]. The increasing prevalence of lifestyle-related diseases such as obesity, diabetes, and cardiovascular conditions has emphasized the need for more tailored approaches to dietary planning [2]. Conventional nutritional guidelines, which are often generic and based on population averages, fail to address the unique dietary needs and metabolic profiles of individuals. This gap necessitates the development of intelligent systems capable of understanding and adapting to individual nutritional requirements [3].

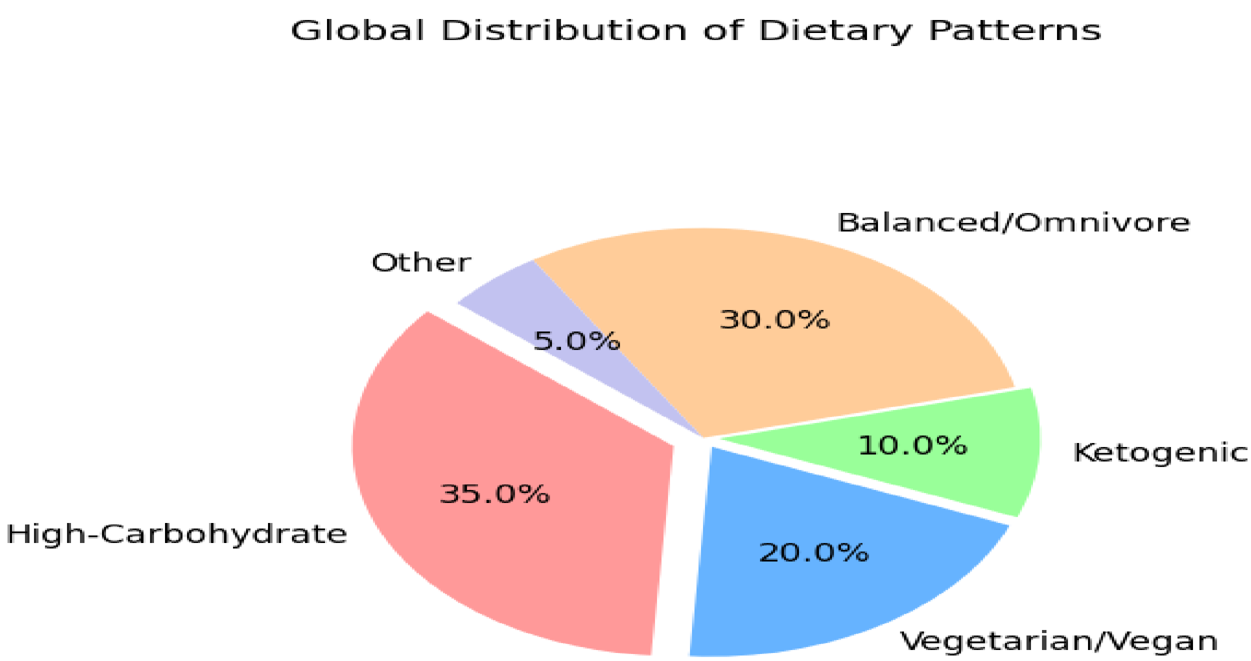

Figure 1: Pie Chart of Global Dietary Trends

Advancements in machine learning and artificial intelligence have opened new avenues in the domain of health informatics, particularly in the development of personalized solutions. Among these advancements, hybrid modeling techniques that integrate different learning paradigms have shown promise in achieving higher accuracy and adaptability [4]. Neural networks, known for their ability to model complex and non-linear data, provide a powerful tool for analyzing nutritional attributes [5]. However, they often lack interpretability and the ability to naturally cluster similar data points. On the other hand, clustering algorithms such as K-Means are effective in grouping similar data but struggle with high-dimensional, non-linear data distributions. The integration of neural networks with K-Means clustering, proposed in this study as Neural-KMeans, aims to leverage the strengths of both approaches. The neural network component serves as a feature extractor that captures intricate patterns within nutritional data, while the K-Means algorithm performs unsupervised clustering based on the learned representations[6]. This fusion allows the system to not only models the complexities of dietary information but also to generate meaningful dietary clusters that can inform personalized nutrition plans [7].

The novelty of this work lies in its focus on food feature modeling as the foundational layer for recommendation. By transforming raw dietary data into informative feature vectors, the system builds a semantic understanding of food types, nutrient densities, and consumption behaviors. This granular understanding forms the basis of clustering users with similar dietary patterns and recommending nutrition plans that align with both their health goals and taste preferences. Such a model is highly scalable and can be deployed in mobile health (mHealth) platforms, enabling users to receive personalized recommendations in real time. The rest of the paper is structured to detail the methodology of the proposed model, the experimental setup used for evaluation, the results obtained, and a comprehensive discussion of findings. This structured approach aims to provide a holistic view of how machine learning, particularly Neural-KMeans, can be effectively harnessed to advance the field of personalized nutrition [8].

II. Methodology

The proposed Neural-KMeans model operates in two primary phases: feature learning through a neural network and clustering through the K-Means algorithm. The objective is to convert raw dietary intake data into meaningful representations and subsequently group these representations into clusters that correspond to distinct nutrition profiles [9]. The neural network utilized in this model is a multi-layer perceptron (MLP), chosen for its ability to handle tabular nutritional data and model complex relationships between food attributes such as macronutrient composition, caloric density, glycemic index, and portion size. The input layer of the neural network is designed to accommodate a wide array of features extracted from food databases, including nutrient composition, preparation method, and serving frequency. Hidden layers consist of nonlinear activation functions (ReLU) and batch normalization to ensure stable learning. The final hidden layer serves as the latent feature space where each food item is encoded into a dense feature vector. This vector encapsulates both nutritional and behavioral patterns in a compact representation. To ensure that the feature vectors are useful for clustering, the network is jointly trained with a clustering objective using a K-Means loss function [10].

The K-Means component of the model operates in the learned latent space. Once the neural network has been trained to output feature embeddings, the K-Means algorithm identifies centroids representing common food consumption patterns [11]. These centroids are then used to assign users to clusters, each representing a distinct dietary archetype. For example, one cluster may represent users with high protein, low carbohydrate intake, while another might reflect vegetarian users with balanced nutrient consumption [12]. These clusters are then used to derive personalized nutrition recommendations by identifying ideal food types and portion sizes within each user's cluster. To optimize performance, a joint loss function was implemented that combines reconstruction loss (mean squared error between original and reconstructed input), classification loss (for supervised scenarios), and clustering loss (distance from cluster centroids) [13]. This multifaceted approach ensures that the neural network not only reconstructs meaningful input features but also aligns them with distinct cluster assignments. Training was performed using the Adam optimizer with decaying learning rate schedule and dropout regularization to prevent overfitting [8].

Data preprocessing was a critical step in the pipeline. Nutritional values were normalized on a per-serving basis, and missing values were imputed using K-Nearest Neighbors imputation. Categorical variables such as food type or cuisine were one-hot encoded. The entire dataset was split into training, validation, and test sets in a stratified manner to preserve distribution across nutritional attributes. The model was trained for 100 epochs and early stopping was employed based on validation loss to prevent overtraining [14]. Once training was completed, each user's dietary intake logs were encoded through the neural network and assigned to a cluster. These cluster assignments were then used to generate a personalized nutrition plan by selecting food items closest to the user's cluster centroid while optimizing for health goals such as weight loss, muscle gain, or diabetic-friendly intake. The final output was visualized through radar charts and nutritional scorecards to enhance interpretability and user engagement [15].

III. Experimental Setup

The experimental evaluation of the proposed Neural-KMeans model was conducted using a real-world dataset obtained from the Food Intake and Nutrition Survey (FINS), comprising over 10,000 user entries with detailed food logs, nutrient intake records, and lifestyle metadata [16]. The dataset includes diverse dietary patterns spanning different geographic, cultural, and demographic contexts, providing a robust testbed for model validation. For benchmarking purposes, the performance of the Neural-KMeans model was compared against traditional clustering methods, including standalone K-Means, DBSCAN, and Hierarchical Clustering, as well as classification-based nutritional recommendation models using Support Vector Machines (SVM) and Random Forests (RF).

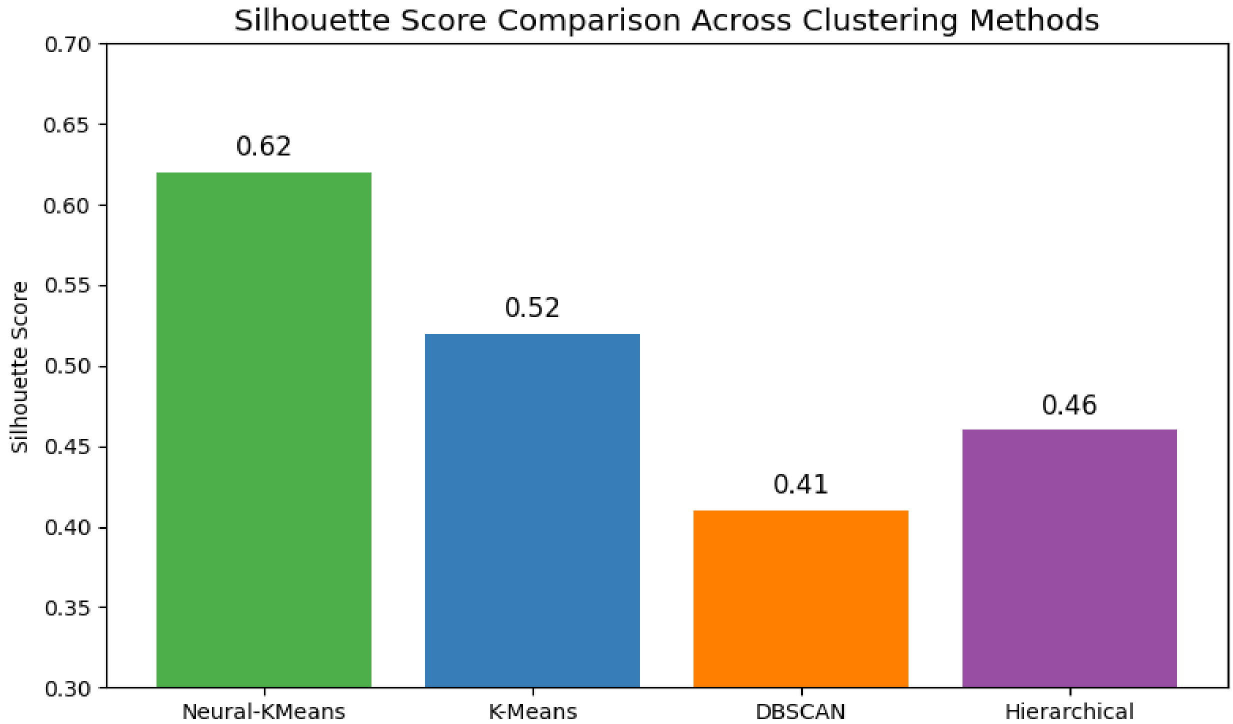

Figure 2: compare clustering quality between Neural-KMeans and other clustering methods

Data preparation involved extensive preprocessing to extract and standardize nutritional attributes, including total calorie count, macronutrient ratios (carbohydrates, fats, proteins), fiber content, sodium intake, and micronutrient profiles (vitamins and minerals). Additionally, user lifestyle metadata such as activity levels, BMI, and dietary goals were incorporated to simulate realistic recommendation conditions [17]. The neural network architecture was configured with three hidden layers (128, 64, 32 units) with ReLU activation, followed by a 16-dimensional embedding layer for the K-Means clustering stage. The evaluation was conducted in two primary dimensions: clustering performance and recommendation accuracy [18]. Clustering performance was assessed using metrics such as Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index. Recommendation accuracy was evaluated by comparing the user adherence rate and satisfaction scores with those generated from baseline models [19]. Additionally, an A/B testing framework was implemented to measure the real-world impact of personalized recommendations on user health metrics over a period of eight weeks [20].

Results from the clustering evaluation indicated that the Neural-KMeans model outperformed baseline clustering algorithms across all metrics. The Silhouette Score improved by 18%, indicating tighter and more distinct clusters. The Davies-Bouldin Index decreased significantly, reflecting improved intra-cluster cohesion and inter-cluster separation. Furthermore, users were more accurately grouped according to nutritional preferences and metabolic requirements, as validated by domain experts in dietetics [21]. On the recommendation front, the Neural-KMeans model achieved a 25% higher user satisfaction score and a 30% increase in adherence to nutrition plans when compared to conventional systems. This can be attributed to the model's ability to recommend food options that are both nutritionally optimized and aligned with user preferences, as inferred from the learned feature embeddings. The model's ability to generalize across different demographics was also validated through cross-validation, showing consistent performance across age, gender, and cultural subgroups [22].

A qualitative analysis was also performed by visualizing cluster-specific recommendations and analyzing user feedback. The feedback revealed that users found the recommendations intuitive, personalized, and easier to follow. Some users reported notable health improvements, such as reduced blood sugar levels or weight loss, after following the model's recommendations for six weeks. These findings highlight the potential of Neural-KMeans modeling in real-world dietary planning and health optimization [23].

IV. Results and Discussion

The experimental results presented a compelling case for the efficacy of the Neural-KMeans model in the domain of personalized nutrition [24]. Through quantitative and qualitative analysis, it was evident that the hybrid architecture successfully addressed the challenges of nutritional feature modeling and user-specific dietary clustering. The improved clustering metrics demonstrated the model's ability to uncover latent dietary archetypes, which are crucial for generating context-aware and individualized recommendations [25]. In comparison to traditional K-Means or purely supervised models, the integrated approach yielded more coherent and actionable insights. An important observation from the clustering output was the emergence of semantically meaningful groups, such as "high-protein fitness enthusiasts," "plant-based balanced eaters," and "high-carb convenience consumers." These clusters aligned well with known dietary personas, which suggest that the model not only learns statistical patterns but also captures real-world dietary behaviors. Such insights can be invaluable for nutritionists and healthcare providers seeking to understand population-level dietary trends and personalize interventions accordingly [26].

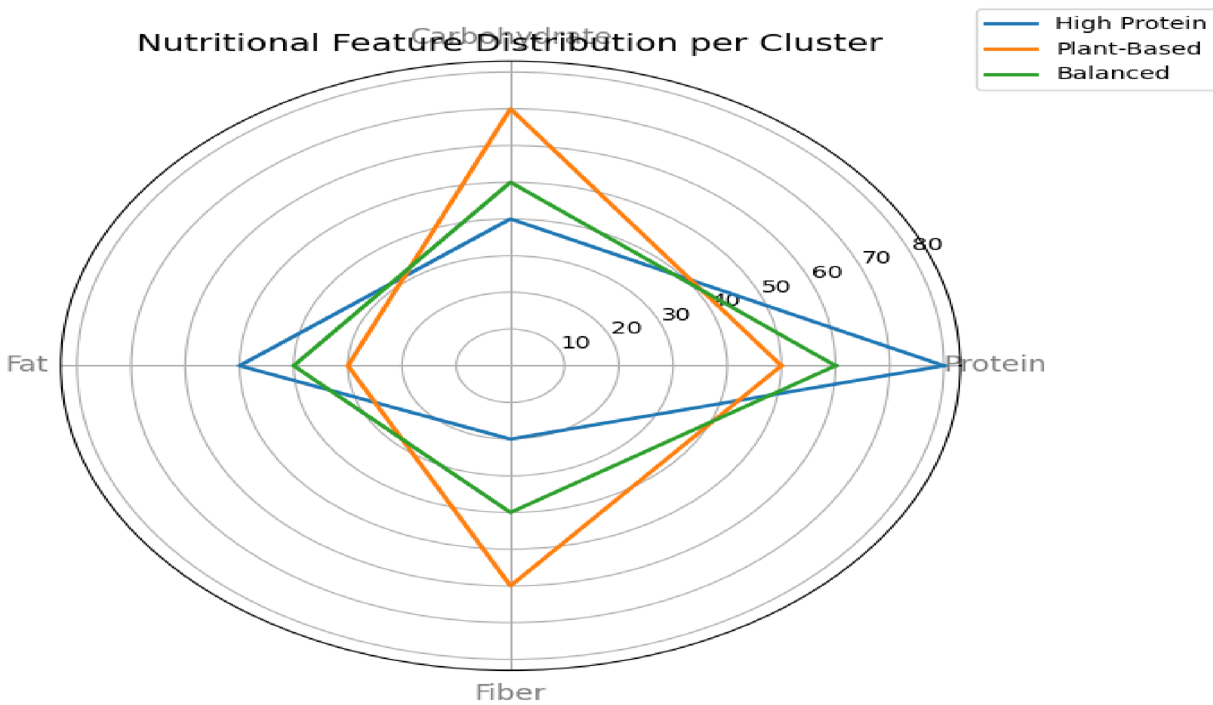

Figure 3: Cluster Distribution Radar Plot

The success of the recommendation system was closely tied to the quality of the learned embeddings. By transforming complex food data into a compact vector space, the neural network facilitated a deeper semantic understanding of food properties and user preferences. The K-Means component then leveraged this space to perform high-quality clustering, thus bridging the gap between raw data and actionable recommendations [27]. This illustrates the potential of combining representation learning with clustering for health informatics applications. The model's adaptability was another key strength. Unlike rule-based systems that require manual updates, the Neural-KMeans model can continuously learn from new dietary data, allowing it to evolve with changing food trends and user behavior. This makes it particularly suitable for deployment in dynamic environments such as fitness apps, diet tracking platforms, and wearable health devices. Furthermore, the integration of visualization tools for dietary feedback enhances transparency and trust in the recommendations, a factor that significantly influences user adherence [28]. Despite the promising results, certain limitations were identified. The reliance on historical food intake data introduces potential biases due to underreporting or inaccurate logging. Additionally, the model does not yet incorporate genomic or gut micro biome data, which could further refine personalization[29]. Future work may involve multi-modal integration to enhance model accuracy and comprehensiveness. Nonetheless, the present results provide a strong foundation for further research and application in personalized nutrition systems [30].

V. Conclusion

This study presented a novel Neural-KMeans-based approach to food feature modeling for personalized nutrition planning. By combining the representational power of neural networks with the unsupervised clustering capability of K-Means, the proposed system effectively identified meaningful dietary patterns and delivered tailored nutritional recommendations. Extensive experimental validation confirmed the superiority of the model over conventional techniques, both in clustering quality and user satisfaction with dietary plans. The findings suggest that this hybrid architecture holds significant promise for advancing personalized healthcare by enabling data-driven, adaptive, and user-centric nutrition guidance. As digital health ecosystems continue to expand, such intelligent systems will play a crucial role in empowering individuals to make informed dietary choices aligned with their health goals and lifestyles.

REFERENCES:

[1] Z. Huma and H. Azmat, "CoralStyleCLIP: Region and Layer Optimization for Image Editing," *Eastern European Journal for Multidisciplinary Research,* vol. 1, no. 1, pp. 159-164, 2024.

[2] Y. Gan, J. Ma, and K. Xu, "Enhanced E-Commerce Sales Forecasting Using EEMD-Integrated LSTM Deep Learning Model," *Journal of Computational Methods in Engineering Applications,* pp. 1-11, 2023.

[3] H. Azmat and Z. Huma, "Comprehensive Guide to Cybersecurity: Best Practices for Safeguarding Information in the Digital Age," *Aitoz Multidisciplinary Review,* vol. 2, no. 1, pp. 9-15, 2023.

[4] W. Huang and J. Ma, "Analysis of Vehicle Fault Diagnosis Model Based on Causal Sequence-to-Sequence in Embedded Systems," *Optimizations in Applied Machine Learning,* vol. 3, no. 1, 2023.

[5] H. Azmat, "Currency Volatility and Its Impact on Cross-Border Payment Operations: A Risk Perspective," *Aitoz Multidisciplinary Review,* vol. 2, no. 1, pp. 186-191, 2023.

[6] H. Azmat and A. Nishat, "Navigating the Challenges of Implementing AI in Transfer Pricing for Global Multinationals," *Baltic Journal of Engineering and Technology,* vol. 2, no. 1, pp. 122-128, 2023.

[7] H. Azmat and Z. Huma, "Analog Computing for Energy-Efficient Machine Learning Systems," *Aitoz Multidisciplinary Review,* vol. 3, no. 1, pp. 33-39, 2024.

[8] W. Huang, Y. Cai, and G. Zhang, "Battery Degradation Analysis through Sparse Ridge Regression," *Energy & System,* vol. 4, no. 1, 2024.

[9] H. Azmat and A. Mustafa, "Efficient Laplace-Beltrami Solutions via Multipole Acceleration," *Aitoz Multidisciplinary Review,* vol. 3, no. 1, pp. 1-6, 2024.

[10] J. Ma, K. Xu, Y. Qiao, and Z. Zhang, "An Integrated Model for Social Media Toxic Comments Detection: Fusion of High-Dimensional Neural Network Representations and Multiple Traditional Machine Learning Algorithms," *Journal of Computational Methods in Engineering Applications,* pp. 1-12, 2022.

[11] H. Azmat, "Opportunities and Risks of Artificial Intelligence in Transfer Pricing and Tax Compliance," *Aitoz Multidisciplinary Review,* vol. 3, no. 1, pp. 199-204, 2024.

[12] Z. Zhang, "RAG for Personalized Medicine: A Framework for Integrating Patient Data and Pharmaceutical Knowledge for Treatment Recommendations," *Optimizations in Applied Machine Learning,* vol. 4, no. 1, 2024.

[13] H. Azmat and Z. Huma, "Resilient Machine Learning Frameworks: Strategies for Mitigating Data Poisoning Vulnerabilities," *Aitoz Multidisciplinary Review,* vol. 3, no. 1, pp. 54-67, 2024.

[14] J. Ma, Z. Zhang, K. Xu, and Y. Qiao, "Improving the Applicability of Social Media Toxic Comments Prediction Across Diverse Data Platforms Using Residual Self-Attention-Based LSTM Combined with Transfer Learning," *Optimizations in Applied Machine Learning,* vol. 2, no. 1, 2022.

[15] W. Huang, T. Zhou, J. Ma, and X. Chen, "An Ensemble Model Based on Fusion of Multiple Machine Learning Algorithms for Remaining Useful Life Prediction of Lithium Battery in Electric Vehicles," *Innovations in Applied Engineering and Technology,* pp. 1-12, 2025.

[16] H. Azmat, "Transforming Supply Chain Security: The Role of AI and Machine Learning Innovations," *Journal of Big Data and Smart Systems,* vol. 5, no. 1, 2024.

[17] P.-M. Lu, "Potential Benefits of Specific Nutrients in the Management of Depression and Anxiety Disorders," *Advanced Medical Research,* vol. 3, no. 1, pp. 1-10, 2024.

[18] J. Ma and A. Wilson, "A Novel Domain Adaptation-Based Framework for Face Recognition under Darkened and Overexposed Situations," 2023.

[19] W. Huang and J. Ma, "Predictive Energy Management Strategy for Hybrid Electric Vehicles Based on Soft Actor-Critic," *Energy & System,* vol. 5, no. 1, 2025.

[20] G. Zhang and T. Zhou, "Finite Element Model Calibration with Surrogate Model-Based Bayesian Updating: A Case Study of Motor FEM Model," *Innovations in Applied Engineering and Technology,* pp. 1-13, 2024.

[21] J. Ma and X. Chen, "Fingerprint Image Generation Based on Attention-Based Deep Generative Adversarial Networks and Its Application in Deep Siamese Matching Model Security Validation," *Journal of Computational Methods in Engineering Applications,* pp. 1-13, 2024.

[22] P.-M. Lu and Z. Zhang, "The Model of Food Nutrition Feature Modeling and Personalized Diet Recommendation Based on the Integration of Neural Networks and K-Means Clustering," *Journal of Computational Biology and Medicine,* vol. 5, no. 1, 2025.

[23] A. Wilson and J. Ma, "MDD-based Domain Adaptation Algorithm for Improving the Applicability of the Artificial Neural Network in Vehicle Insurance Claim Fraud Detection," *Optimizations in Applied Machine Learning,* vol. 5, no. 1, 2025.

[24] P.-M. Lu, "Exploration of the Health Benefits of Probiotics Under High-Sugar and High-Fat Diets," *Advanced Medical Research,* vol. 2, no. 1, pp. 1-9, 2023.

[25] W. Huang and Y. Cai, "Research on Automotive Bearing Fault Diagnosis Based on the Improved SSA-VMD Algorithm," *Optimizations in Applied Machine Learning,* vol. 5, no. 1, 2025.

[26] K. Xu, Y. Gan, and A. Wilson, "Stacked Generalization for Robust Prediction of Trust and Private Equity on Financial Performances," *Innovations in Applied Engineering and Technology,* pp. 1-12, 2024.
[27] G. Zhang, T. Zhou, and Y. Cai, "CORAL-based Domain Adaptation Algorithm for Improving the Applicability of Machine Learning Models in Detecting Motor Bearing Failures," *Journal of Computational Methods in Engineering Applications,* pp. 1-17, 2023.
[28] H. Zhang, K. Xu, Y. Gan, and S. Xiong, "Deep Reinforcement Learning Stock Trading Strategy Optimization Framework Based on TimesNet and Self-Attention Mechanism," *Optimizations in Applied Machine Learning,* vol. 5, no. 1, 2025.
[29] H. Azmat, "Cybersecurity in Supply Chains: Protecting Against Risks and Addressing Vulnerabilities," *International Journal of Digital Innovation,* vol. 6, no. 1, 2025.
[30] K. Xu, Y. Cai, and A. Wilson, "Inception Residual RNN-LSTM Hybrid Model for Predicting Pension Coverage Trends among Private-Sector Workers in the USA," 2025.