Management Letters / Cuadernos de Gestión 25/1 (2025) 57-70

FESIDE
Fundación Emilio Soldevilla para la Investigación y Desarrollo de la Economía de la Empresa

enpresa institutua
Instituto de Economía Aplicada a la Empresa

Management Letters / Cuadernos de Gestión

journal homepage: https://ojs.ehu.eus/index.php/CG

ISSN: 1131-6837 / e-ISSN: 1988-2157

Exploring the Frontier: Generative AI Applications in Online Consumer Behavior Analytics

Explorando la frontera: Aplicaciones generativas de la IA en el análisis del comportamiento de los consumidores en línea

Takuma Kimura*

* ***Corresponding author:*** *Hosei University, Japan. 2-17-1, Fujimi, Chiyoda-Ku, Tokyo, 102-8160– ktakuma@hosei.ac.jp – Showa Women's University, Japan, 1-7-57, Taishido, Setagaya-ku, Tokyo, 154-8533 – 221055@swu.ac.jp – https://orcid.org/0000-0001-7126-188X*

ARTICLE INFO

Received 23 November 2023,
Accepted 14 March 2024

Available online 11 September 2024

DOI: 10.5295/cdg.232121tk

JEL: M31

ABSTRACT

This paper presents a systematic review of the application of generative artificial intelligence (AI) in online consumer behavior analytics (OCBA). With the advent of e-commerce and social media, consumer behavior increasingly occurs online, generating vast amounts of data. This shift necessitates advanced analytical tools, and generative AI emerges as a pivotal technology. Generative AI, distinct from traditional AI, can autonomously generate new content based on learned data patterns, offering innovative approaches to OCBA. Based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology and data synthesis method proposed by Webster and Watson (2002), this study analyzes 28 peer-reviewed papers, focusing on how generative AI is applied in OCBA and how it can enhance OCBA performance. The findings show that generative adversarial networks (GANs) are the most used, followed by variational autoencoders (VAEs) and autoregressive models. This review categorizes the application areas of generative AI in OCBA and examines how these technologies enhance OCBA's effectiveness and efficiency. Furthermore, the paper discusses the challenges associated with generative AI, emphasizing the need to consider ethical issues, such as bias and data privacy. This comprehensive review contributes to a deeper understanding of generative AI's role in OCBA, outlining its applications and functionalities from a technical perspective. It guides future research and practice, highlighting areas for further exploration and improvement in leveraging generative AI for consumer behavior analytics.

Keywords: Generative artificial intelligence, Generative adversarial network, Variational autoencoders, Autoregressive model, Generative pre-trained transformer, Online consumer behavior analytics.

RESUMEN

Este artículo presenta una revisión sistemática de la aplicación de la IA generativa en el Análisis del Comportamiento del Consumidor Online (OCBA). Con la llegada del comercio electrónico y las redes sociales, el comportamiento de los consumidores se produce cada vez más en línea, lo que genera enormes cantidades de datos. Este cambio requiere herramientas analíticas avanzadas, y la IA generativa emerge como una tecnología fundamental. La IA generativa, distinta de la IA tradicional, puede generar de forma autónoma nuevos contenidos basados en patrones de datos aprendidos, ofreciendo enfoques innovadores a la OCBA. Basado en la metodología PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) y el método de síntesis de datos propuesto por Webster y Watson (2002), este estudio analiza 28 artículos revisados por pares, centrándose en cómo se aplica la IA generativa en OCBA y cómo puede mejorar el rendimiento de OCBA. Los resultados muestran que las redes generativas adversariales (GAN) son las más utilizadas, seguidas de los autocodificadores variacionales (VAE) y los modelos autorregresivos. La revisión clasifica las áreas de aplicación de la IA generativa en OCBA y examina cómo estas tecnologías mejoran la eficacia y la eficiencia de OCBA. Además, el artículo analiza los retos asociados a la IA generativa, haciendo hincapié en la necesidad de tener en cuenta cuestiones éticas como la parcialidad y la privacidad de los datos. Esta revisión contribuye a una comprensión más profunda del papel de la IA generativa en la OCBA, esbozando sus aplicaciones y funcionalidades desde una perspectiva técnica.

Palabras clave: Inteligencia artificial generativa, Redes generativas adversariales, Autocodificadores variacionales, Modelo autorregresivo, Transformador generativo preentrenado, Análisis del comportamiento del consumidor online..

1. INTRODUCTION

With the rise of e-commerce and the proliferation of social media, consumer behavior is increasingly taking place online. This trend was accelerated by the COVID-19 pandemic (Sajid *et al.*, 2022). Online purchasing decreases consumers' travel time, and offers wider product choice and easier price comparisons (Jiang *et al.*, 2013). As online consumption becomes increasingly common, businesses need to discover patterns in online consumer behavior and accurately predict future behavior to make optimal offers at the right time to maximize profits.

Online consumer behavior includes touchpoints such as web searches, social media interactions, and posting online reviews. These activities rapidly generate vast amounts of data on consumer behavior. The data can be used to discern consumer behavior patterns, predict future consumer behavior, and form strategies tailored to specific audiences (Anshari *et al.*, 2019). Thus, organizational competence for online consumer behavior analytics (OCBA), an analytic method to discover patterns in online consumer behavior, can determine an organization's competitiveness.

The emergence of generative Artificial Intelligence (AI) has been touted as the most thrilling innovation in AI in history. The release of ChatGPT in 2023 focused the public's attention on generative AI and dramatically raised the profile of AI in general. While traditional AI excels at analysis and prediction based on existing data, generative AI marks a revolutionary shift in AI's creative capabilities. It autonomously generates new content based on patterns found in the data it is trained on.

The use of generative AI in OCBA has previously been discussed in the literature and research on generative AI algorithms for OCBA is ongoing. However, prior research has yet to provide a systematic view of how generative AI algorithms and their core technologies are utilized in OCBA and how they contribute to improving the performance of these analytic methods. Therefore, this paper systematically reviews prior research on using generative AI in OCBA from a technical perspective. Specifically, we review the OCBA areas where generative AI has been developed and applied in previous studies. We also systematically identify the mechanisms by which generative AI enhances the effectiveness and efficiency of OCBA.

The subsequent chapters of this paper are organized as follows. Chapter 2 outlines the differences between generative AI and traditional AI, describes the main generative AI algorithms, and presents the research questions this review addresses. Chapter 3 describes the methodology of the systematic review. Chapter 4 answers the research questions using a systematic synthesis of previous studies. Chapter 5 summarizes the findings and discusses the limitations of this study and implications for future research and practice.

2. GENERATIVE AI FOR OCBA

2.1. Machine learning and its application to OCBA

Consumer behavior research has traditionally been conducted through observational studies, focus groups, and questionnaire-based surveys. A drawback of observational studies and focus groups is that they are time-consuming and expensive to conduct (Breen, 2006; Grove & Fisk, 1992). Furthermore, these methods often have small sample sizes, so the results may not represent populations. In contrast, questionnaire surveys are quicker and cheaper, allowing for large sample sizes. However, because questionnaires are generally self-reported, they are subject to social desirability bias that may prevent participants from giving honest and accurate responses (Kimura, 2023).

As more and more consumer behavior occurs online, new opportunities for consumer behavior analytics have emerged. Every click, purchase, and online interaction on online platforms generates enormous amounts of data. Additionally, companies can now collect continuously generated and updated data in real-time. Moreover, the collected data is not limited to numerical information but can include diverse modes such as language, images, audio, and video. This variety and volume of data have enabled and increased the importance of data-driven decision-making.

Machine learning, an artificial intelligence (AI) technique, enables machines to learn from data without detailed programming (Samuel, 1959). Machine learning technologies can learn from massive amounts of data, gain meaningful insights, and make predictions and decisions. The most notable strength of machine learning compared to traditional statistical analysis is its adaptable autonomous learning ability. In traditional statistical analysis, the analyst must formulate a mathematical model in advance based on a hypothesis or research question. In contrast, machine learning models can autonomously find patterns in the data and discover relationships among variables that were not apparent *a priori*. This strength of machine learning is critical when there are complex, nonlinear relationships among variables and when the data is unstructured.

Because of its dynamic nature, machine learning can obtain insights from continuously (sometimes real-time) updated data. Thus, machine learning has been applied to various consumer behavior analytics such as customer segmentation, customer churn prediction, sentiment analysis, and personalized recommendation (Ngai & Wu, 2022; Policarpo *et al.*, 2021).

2.2. Generative AI

The emergence of generative AI has brought about recent innovations in OCBA (Baek, 2023). Generative AI is a program that can create new content, such as text, images, music, and videos, after being trained on large datasets. Generative AI differs from traditional AI in that it can go beyond data analysis and prediction and generate novel creations. There are various algorithms for developing generative AI. Popular options include energy-based models (EBMs), generative adversarial networks (GANs), variational autoencoders (VAEs), autoregressive models, and normalizing flows (Bond-Taylor *et al.*, 2022).

A. EBMs

An EBM is a generative model that learns underlying data distribution from a sample dataset and generates similar datasets. The critical components of EBMs are an energy function and a partition function (Sun *et al.*, 2021). The energy function is a scalar function that assigns an energy value to each possible configuration of the input data. The energy value represents the compatibility between the observed and latent variables. If the energy value is large, the configuration of the observed and latent variables is less compatible. The energy function is trained to assign low energy values to the actual data and high energy values to the generated data, whereas, the partition function is trained to normalize the energy values.

B. GANs

GANs are neural networks that can generate new data by learning from existing data (Goodfellow *et al.*, 2020). GANs consist of two networks: a generator and a discriminator (Pan *et al.*, 2019). The generator network is trained to generate new data similar to the existing data, whereas the discriminator network is trained to distinguish between the existing and generated data. These networks are trained together in an adversarial process, where the generator network tries to create improved data, and the discriminator network tries to identify the generated data. As the network's training proceeds, the generator gets better at generating fake data that is hard to distinguish from real data, and the discriminator becomes more proficient at identifying the fake data.

C. VAEs

VAEs are generative AIs that merge elements from statistics and information theory with the flexibility of deep neural networks to address the generation problem for high-dimensional data (Asperti *et al.*, 2021). VAEs are a type of autoencoder that learn from a compressed representation of the input data, called the latent space, and generate new data by sampling it. By compressing the input data into a lower-dimensional representation, VAEs learn from an efficient and compact representation of the input data, which can be used to generate new data similar to the training data. VAEs are trained using a probabilistic approach, where the model learns to maximize the likelihood of the training data, that is, achieve a good approximation of the true posterior distribution (Notin *et al.*, 2021).

D. Autoregressive models

Autoregressive models generate new data by modeling each data point as a function of the previous data points in the sequence (Postolache *et al.*, 2023). They define a probability distribution over a data-point sequence and then sample it to generate new data points. The probability distribution is learned by maximizing the likelihood of the training data. Unlike GANs and VAEs, autoregressive models do not require any additional networks or representations. They model each data point as a function of the previous data points in the sequence.

The generative pre-trained transformer (GPT) is an autoregressive model that generates new text by modeling each token in the sequence as a function of the previous tokens (Floridi & Chiriatti, 2020). Its foundation lies in the transformer (Vaswani *et al.*, 2017), which enables the prediction model to focus on different parts of the input sequence (Radford *et al.*, 2018). More specifically, the self-attention mechanism in the transformer architecture enables the model to capture dependencies between different tokens in the input sequence, which is crucial for generating coherent and meaningful text.

E. Normalizing flows

Normalizing flows generate new data by transforming a simple base distribution into a complex target distribution using a series of invertible transformations (Papamakarios *et al.*, 2021). These transformations are designed to be computationally efficient and can be used to model complex distributions accurately. Unlike other generative models such as GANs and VAEs, normalizing flows learn a probability distribution over the sequence of data points and then sample it to generate new data points. The probability distribution is learned by maximizing the likelihood of the training data.

2.3. Research gaps and research questions

As noted above, there has been a pronounced interest in applying AI to various OCBA areas, such as personalization and recommendation systems (Kashyap *et al.*, 2022; Necula & Păvăloaia, 2023; Viktoratos & Tsadiras, 2021), sentiment analysis (Almahmood & Tekerek, 2022, Bawack *et al.*, 2022), virtual try-on (Goti *et al.*, 2023), and fraud detection (Policarpo *et al.*, 2021). However, a comprehensive mapping of what generative AI technologies are being utilized in what areas of OCBA has yet to be provided. Moreover, there needs to be a more systematic discussion on how generative AI technologies are expected to enhance the effectiveness of OCBA. Therefore, this paper addresses the following research questions through a systematic literature review.

Research Question 1: In what OCBA areas are generative AI algorithms applied?

Research Question 2: How does the use of generative AI benefit each OCBA area?

3. METHOD

3.1. Review methodology

For the systematic review, this study adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology (Rethlefsen *et al.*, 2021). PRISMA is a validated framework that ensures a consistent, reproducible, and high-quality review process. Following the PRISMA guidelines, the systematic review was conducted in several phases (Figure 1), each of which is elaborated upon in the subsequent sections.

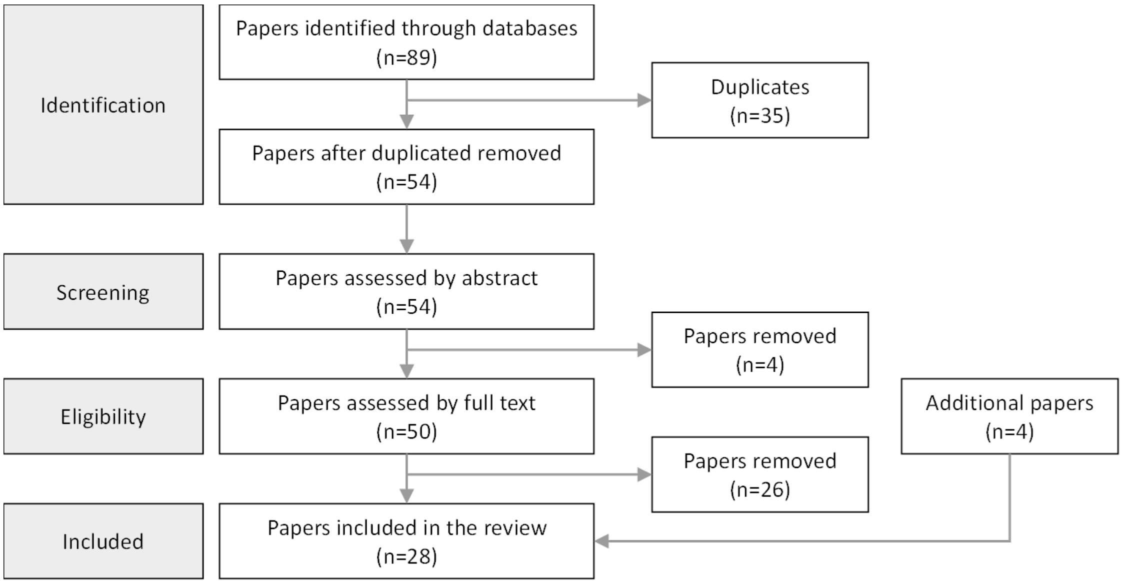

Figure 1
Diagram of identification, screening, and inclusion of studies
Source: This workflow follows (PRISMA) methodology (Rethlefsen *et al.*, 2021).

3.2. Inclusion and exclusion criteria

To provide a structured and exhaustive review, it is pivotal that the studies included are relevant and that non-relevant studies are systematically excluded. The inclusion and exclusion criteria were meticulously crafted to ensure the research questions were addressed adequately.

While various techniques and applications of generative AI have been studied extensively, it is not easy to comprehensively review them in a single paper. Therefore, this study includes only studies that explicitly relate to OCBA. Because of the rapid advancements in generative AI, only papers published within the past five years (i.e., 2019-2023) were included to ensure the review is grounded in contemporary methodologies and insights. Studies had to be published in a peer-reviewed journal to ensure that the current review is based on high-quality and credible research. In addition, the study had to be written in English, the predominant language for scientific communication in this field.

Because the objective of this study was to produce a primary review of model development and evaluation of generative AI for OCBA, theoretical papers, review papers, and editorials were excluded from this review. For the same reason, empirical and experimental research that does not contain generative AI model development was excluded. Further, consistent with the inclusion criteria, conference proceedings, and papers not written in English were excluded.

3.3. Information sources and search strategy

To ensure comprehensive coverage of relevant studies, the review sources in this study were set according to the DARE criteria proposed by the York University Centre for Reviews and Dissemination (CDR). The DARE criteria recommend searching four or more digital libraries and adding additional searches. Therefore, this study conducted article searches using Scopus, Web of Science, IEEE Xplore, and ACM Library, the major article databases in AI research. In addition, articles identified during the review process and meeting this study's objectives were added to the body of work to be reviewed.

In the search, conducted on October 22, 2023, three keywords were used to identify studies related to OCBA: "e-commerce," "online shopping," and "online consumer." To find studies related to generative AI, "energy-based model," "generative adversarial network," "variational autoencoder," "autoregressive model," "normalizing flow," "generative model," "transformer-based model," "transformer architecture," "generative pre-trained transformer," "generative neural network," "generative AI," and "generative artificial intelligence" were used as keywords.

Search formulas were compiled using Boolean operators, with the condition for review inclusion of at least one of the keywords related to OCBA and at least one related to generative AI in either the title, keywords, or abstract. The query was structured as follows:

> ("e-commerce" OR "online shopping" OR "online consumer") AND "energy-based model" OR "generative adversarial network" OR "variational autoencoder" OR "autoregressive model" OR "normalizing flow" OR "transformer-based model" OR "transformer architecture" OR "generative pre-trained transformer" OR "generative model" OR "generative neural network" OR "generative AI" OR "generative artificial intelligence")

3.4. Screening

A total of 89 papers were extracted following the search. The 89 papers were screened to ensure the selected literature was relevant to the research questions. The screening was implemented as a multi-step process to refine the initial pool into a curated set of relevant studies.

First, duplicates, which were single articles extracted from multiple sources, were removed. This step removed 35 duplicates, leaving 54 papers. The next step was to read the abstracts of the remaining articles. At this stage, studies that were unrelated or did not address the integration of generative AI with OCBA were removed. Two editorials and two literature reviews were removed at this stage.

The remaining 50 articles that survived the initial screening were then subjected to a full-text review. Each study was evaluated using the inclusion and exclusion criteria. Of the 50 articles included in the full-text review, 16 did not concern generative AI, and 10 were studies not directly related to OCBA. Following their exclusion, 24 articles were selected for the next step, quality assessment.

3.5. Quality assessment

Quality assessment was conducted using the 14 evaluation items for quantitative studies proposed by Kmet *et al.* (2004). Of these 14 items, five items (3, 4, 5, 6, 7), suitable for clinical trials and controlled experiments but not relevant to AI model development, were excluded. Following the author's scoring method, the remaining nine items were rated on a scale from 0 to 18, with 2 points for Yes, 1 for Partial, and 0 for No for each item. A score of 13 or higher, corresponding to a scoring rate of at least 70% of the 18 points, was set as the cutoff score value for inclusion.

All 24 articles rated as relevant to the full-text review achieved a score above the cutoff in the quality assessment. In addition, four articles were included following reference-list searches, resulting in 28 articles in total for the final review. All of the newly added articles were registered in the Scopus database.

3.6. Data extraction and synthesis

Data extraction and synthesis followed the methodology of Webster and Watson (2002). First, an author matrix was created for each article. The author matrix included the author's name, article title, journal title, publication year, generative AI algorithm, and application area (e.g., recommender system, sentiment analysis). We then created multiple concept matrices based on each research question. The concept matrix summarizes the information extracted from the author matrix and the full-text review.

4. RESULTS

Table 1 groups reviewed papers by the generative AI algorithms used in those papers. Among the 28 papers analyzed, 20 employed GANs, six used VAEs, and three utilized autoregressive models. Notably, Ding *et al.* (2023) combined GANs and VAEs to create VAEGANs. The table outlines the specific areas of OCBA where generative AI was applied and the role of generative AI in enhancing OCBA's effectiveness in each area. The upcoming sections delve into the application areas of each generative AI algorithm (Research Question 1) and the ways in which generative AI improves OCBA in these areas (Research Question 2).

Table 1
Author matrix

Generative AI Algorithm	Papers
GAN	Du *et al.* (2023), Fincato *et al.* (2022), Fiore *et al.* (2019), Guo *et al.* (2023), Hao *et al.* (2022), Li *et al.* (2023), Lu *et al.* (2022), Nabeel *et al.* (2019), Rabbi *et al.* (2023), Roy *et al.* (2022), Singh *et al.* (2020), Terzioğlu *et al.* (2022), Wang *et al.* (2020), Wang & Yang (2021), Wei *et al.* (2023), Zhang, Sun *et al.* (2020), Zhang Sun *et al.* (2021), Zhang, Wang *et al.* (2020), Zhao *et al.* (2020)
VAE	Laenen & Moens (2022), Li *et al.* (2022), Nguyen & Cho (2020), Yu & Grauman (2020), Zhang, Wong *et al.* (2021)
Autoregressive model	Deshai & Bhaskara Rao (2023), Han *et al.* (2023), Perez-Castro *et al.* (2023)
GAN and VAE	Ding *et al.* (2023)

Source: Own elaboration.

4.1. Application areas of generative AI

Table 2 organizes the reviewed studies by algorithm and application area in OCBA. Generative AI was developed and validated in the following 14 OCBA areas.

A. Ad creative generation

Ad creative generation is the process of automatically creating marketing messages and visuals for advertising campaigns. It uses generative AI algorithms to craft ad copy relevant to a product that resonates with the target audience. The ad creative generation process is a text summarization task, where information from product landing pages is condensed into compelling ad copy. This process utilizes abstractive summarization to include novel phrases not found on the landing page but relevant to the content.

B. Anomaly detection

In e-commerce settings, anomaly detection models are developed to detect unusual consumer patterns, which could indicate fraud or market shifts. For instance, Guo *et al.* (2023) developed and tested a generative AI model to identify unusual patterns or outliers in dynamic graphs, which can be observed in activities in social networks and e-commerce. In another study, Li *et al.* (2022) addressed outlier detection, where the model tries to identify data points that significantly deviate from the norm.

Table 2
Application areas of generative AI in OCBA

Application Area	Algorithm		
	GAN	**VAE**	**Autoregressive model**
Ad creative generation	Terzioğlu *et al.* (2022)		
Anomaly detection	Guo *et al.* (2023)	Li *et al.* (2022)	
Credit card fraud detection	Ding *et al.* (2023) Fiore *et al.* (2019)	Ding *et al.* (2023)	
Customer behavior prediction	Hao *et al.* (2022)		
Dialogue generation	Nabeel *et al.* (2019)		
Fake review detection (& generation)			Deshai & Bhaskara Rao (2023) Perez-Castro *et al.* (2023)
Identification of authentic product	Rabbi *et al.* (2023)		
Image-based product retrieval	Zhang, Sun *et al.* (2020) Zhan, Sun *et al.* (2021) Zhang, Wang *et al.* (2020)		
Multi-human parsing	Zhao *et al.* (2020)		
Online review summarization			Han *et al.* (2023)
Recommendation system	Lu *et al.* (2022) Singh *et al.* (2020) Wang *et al.* (2020)	Laenen & Moens (2022) Nguyen & Cho (2020) Yu & Grauman (2020) Zhang, Wong *et al.* (2021)	
Sales prediction	Wang & Yang (2021)		
User identity alignment prediction	Wei *et al.* (2023)		
Virtual try-on	Du *et al.* (2023) Fincato *et al.* (2022) Li *et al.* (2023) Roy *et al.* (2022)		

Source: Own elaboration.

C. Credit card fraud detection

Credit card fraud detection can identify fraudulent activities in credit card transactions. This type of fraud can be seen as a binary classification problem, where the model classifies a transaction as fraudulent or non-fraudulent. It is usually an imbalanced classification problem where fraudulent transactions are rare compared to non-fraudulent transactions.

D. Customer behavior prediction

In a study on consumer behavior prediction, Hao *et al.* (2002) developed a model that predicted four customer behaviors: browsing, commenting, reposting, and another unspecified type. These behaviors are influential in the promotion of product information in e-commerce marketing.

E. Dialogue generation

A dialogue management system is a system that manages a conversation with a user, handling tasks like online shopping and booking. The quality of dialogues generated by the system can influence user experience and purchase decision-making.

F. Fake review detection and generation

Fake review detection refers to identifying and classifying reviews that are not genuine or have been manipulated to mislead consumers. The fake review detection model uses machine learning and natural language processing techniques to analyze online reviews for authenticity. In contrast, fake review generation is the creation of misleading opinions that do not reflect the genuine opinion of the author. Improvements in fake review generation by novel generative AI make it increasingly difficult to detect fake reviews.

G. Identification of authentic products

This image-processing task involves verifying the authenticity of a product to prevent fraud. It is especially important in e-commerce where consumers cannot physically examine products.

H. Image-based product retrieval

Image-based product retrieval is a technique to retrieve products from images. It uses computer vision and machine learning techniques to analyze the visual features of images and retrieve visually-similar products. The studies reviewed here processed images to detect and retrieve clothing items from images containing human models.

I. Multi-human parsing

Multi-human parsing is a computer vision task that separates and identifies different persons and objects within an image. Multi-human parsing can enhance the accuracy of product recommendations by analyzing the clothing items worn by people in online images.

J. Online review summarization

Online review summarization refers to the systematic generation of concise summaries of online user reviews that retain essential information. In OCBA, the process is used to extract key themes and sentiments from user feedback to inform product design and development stages.

K. Recommendation system

The recommendation system provides personalized recommendations to customers based on their past behavior, interest, and preferences. The goal of recommendation systems is to improve user experience by providing relevant and useful recommendations tailored to the users' needs.

L. Sales prediction

Sales prediction is a practice of time series forecasting that estimates the future sales volumes of products or services based on historical data and other influencing factors. It can inform business decisions such as inventory management and marketing strategies.

M. User identity alignment prediction

User identity alignment prediction refers to matching users' accounts across e-commerce platforms. It enables a comprehensive view of a user's activities, improving personalized marketing and user experience.

N. Virtual try-on

Virtual try-on is a technique that enables users to try on products digitally using their phones or tablets. It aims to enable clothing trials on e-commerce websites by generating a rendition of a person wearing the clothes that they are viewing on line. Virtual try-on can be used to improve the accuracy of product recommendations by analyzing the clothing items worn by people in the images.

4.2. How generative AI improves OCBA

Table 3 summarizes the generative AI functions that improve the model's performance in each OCBA area. In Table 3, the rows represent the functionality, the columns are the algorithms used in each study, and the application areas and references are noted in the cells. As shown here, the relationship between functionality and algorithm is not one-to-one but many-to-many. One functionality is utilized in multiple application areas and vice versa. The current review identified 12 functionalities. In other words, generative AI can provide 12 OCBA-specific benefits. The following subsections explain these 12 generative AI functionalities, summarizing findings and insights obtained in the review process.

A. Abstractive summarization

Abstractive summarization is a method that leverages language models to generate text in an advanced fashion, similar to human interpretation, rather than simply extracting and linking key sentences or paragraphs from the original text (Nallapati *et al.*, 2016). Han *et al.* (2023) applied a transformer-based model called T5 (text-to-text transfer transformer) to online review summarization. The model leverages a synthetic dataset for fine-tuning, enabling it to handle various granularities and polarities in summarization. In their study, T5 was fine-tuned to generate abstractive summaries that capture user opinions from online reviews, aiding in product design and development.

B. Alleviation of data sparsity problem

Data sparsity refers to a situation where a large proportion of the data is missing or zero. The data sparsity problem can lead to poor model performance because a lack of data can lead to inaccurate predictions (Guo *et al.*, 2022). Notably, collaborative filtering, one of the most popular methods for building recommendation systems, suffers from data sparsity problems (Hu *et al.*, 2023; Margaris *et al.*, 2022).

Lu *et al.* (2022) applied GANs to develop a recommendation system. In their study, GANs created denser datasets by generating samples from the original dataset to enhance the data space, compensating for the lack of information. In another study, Wang *et al.* (2020) applied GANs to improve the prediction of links between items. GAN-based link prediction generates new links in the training data to alleviate data sparsity with synthetic data.

Yu and Grauman (2020) and Zhang, Wong *et al.* (2021) utilized VAEs to address data sparsity problems in building a recommendation system. Like GANs, VAEs were used to generate synthetic training data to compensate for data sparsity. Yu and Grauman (2020) developed Conditional VAE (CVAE) that incorporates conditional variables into the encoder and decoder, enabling it to generate data specific to given conditions. Unlike traditional VAEs, CVAEs can generate targeted data modifications, making them suitable for tasks like attribute manipulation in images. Zhang, Wong *et al.* (2021) developed a hybrid VAE (HVAE), which is an integration of VAE and collaborative filtering. Unlike traditional VAEs that may focus on one-sided content information, HVAE jointly learns the latent representations of content information for both users and items.

Data sparsity can also be a problem in user identity alignment prediction. Wei *et al.* (2023) developed Double-GAN to address data sparsity. Their model utilized data from a base platform and also from heterogeneous e-commerce platforms. Its two-layer iterative mechanism enabled the model to match users' accounts across different platforms effectively, enhancing cross-platform consumer behavior analysis.

Table 3
Major Functionalities of Generative AI

Functionality	GANs	VAEs	Autoregressive model
Abstract summarization			Online review summarization (Han *et al.*, 2023)
Alleviation of data sparsity problem	Recommendation system (Lu *et al.*, 2022; Wang *et al.*, 2020) User identity alignment prediction (Wei *et al.*, 2023)	Recommendation system (Yu & Grauman, 2020; Zhang, Wong *et al.*, 2021)	
Alleviation of exposure bias problem	Ad creative generation (Terzioğlu *et al.*, 2022)		
Controllable image generation	Virtual Try-On (Du *et al.*, 2023; Fincato *et al.*, 2022; Li *et al.*, 2023; Roy *et al.*, 2022)		
Cross-domain retrieval	Image-based product retrieval (Zhang, Sun *et al.*, 2020, 2021; Zhang, Wang *et al.*, 2020)		
Feature extraction		Anomaly detection (Li *et al.*, 2022)	Fake review detection (Deshai & Bhaskara Rao, 2023)
Identification of independent generative factors		Recommendation system (Laenen & Moens, 2022)	
Learning of Normal Data Distribution	Anomaly detection (Guo *et al.*, 2023) Customer behavior prediction (Hao *et al.*, 2022) Authentic product identification (Rabbi *et al.*, 2023) Sales forecasting (Wang & Yang, 2021)"	Recommendation system (Nguyen & Cho, 2020)	Fake review generation (Perez-Castro *et al.*, 2023)
New item generation	Recommendation system (Singh *et al.*, 2020)		
Optimal policy selection	Dialogue generation (Nabeel *et al.*, 2019)		
Rebalancing dataset	Credit card fraud detection (Ding *et al.*, 2023; Fiore *et al.*, 2019)	Credit card fraud detection (Ding *et al.*, 2023)	
Regularization for realistic output	Multi-human parsing (Zhao *et al.*, 2020)		

Source: Own elaboration.

C. Alleviation of exposure bias problem

Exposure bias is a problem where a machine learning model performs well on the training data but poorly on the testing data because there is a large discrepancy between the distribution of training data and that of test data (Yang *et al.*, 2022). In a setting of ad creative generation, Terzioğlu *et al.* (2022) applied GAN and reinforcement learning to address the exposure bias problem. GAN enables the generator to create results by learning to confuse the discriminator, reducing reliance on ground truth data for generating outputs. Moreover, reinforcement learning addresses the exposure bias problem by using rewards instead of ground truth data, further lessening the generator's dependence on it.

D. Controllable image generation

Controllable image generation is the ability of a machine learning model to manipulate and control specific aspects of an image, such as the appearance and location of objects, in order to generate desired outputs (Casanova *et al.*, 2023). In OCBA, controllable image generation can be applied to virtual try-on. A notable challenge for a virtual try-on system is accurately detecting and tracking the user's body and facial features. In a virtual try-on setting, controllable image generation allows customers to generate images of themselves with specific attributes or features, such as the position of the eyes, nose, mouth, and the shape of the face. It can be useful in generating realistic object images and creating personalized content.

Du *et al.* (2023) and Fincato *et al.* (2022) applied GAN to develop a virtual try-on system. GAN can achieve high-quality controllable image generation by providing specific inputs to the generator network to generate images with specific attributes or features. Roy *et al.* (2022) proposed landmark guided virtual try-on (LGVTON), which uses GAN to approximate data distribution, leading to better output in a virtual try-on system. Li *et al.* (2023) developed precise outfit visualization net (POVNet), which uses GANs for learned rendering procedures that ensure the accurate reflection of fine details like shading. In POVNet, adversarial loss from GANs is applied to ensure high-resolution rendering and fine shading accuracy.

E. Cross-domain retrieval

Cross-domain retrieval is a task to find relevant data in one domain using a query from another (Zhou *et al.*, 2022). For example, if we have an image from one domain and want to retrieve similar images from another, we can use cross-domain retrieval techniques. For instance, when analysts have a dataset of cloth images and need to retrieve similar cloth images from a separate dataset, they can use cross-domain retrieval techniques to identify similar images. Cross-domain retrieval is a challenging problem because the data in different domains can differ in distributions and feature representations.

Previous studies have utilized GAN-based models for cross-domain retrieval to translate clothes from the human body to a tiled image. Zhang, Wang *et al.* (2020) used GAN to transform images of clothes worn on human bodies into tiled clothing images, resulting in high accuracy and efficiency in clothing item retrieval. Zhang, Sun *et al.* (2020) applied CascadeGAN (Cas-GAN) in a similar task and showed that Cas-GAN outperformed traditional methods by generating higher-quality images that improved clothing retrieval performance. A notable strength of their model is that it requires only a single human body image for one-to-one clothing image translation, simplifying the retrieval process.

Zhang, Sun *et al.* (2021) proposed and applied triple supervised GAN (TripleGAN), which consists of three components: a generator, a discriminator, and a classifier. The model used triplet loss in the discriminator to ensure the generated images are more similar to the real images than fake ones. Their experiment showed that the TripleGAN model outperforms other GAN models in generating images with delicate details, which is beneficial for cross-domain clothing retrieval.

F. Feature extraction

Feature extraction is a technique to reduce the number of resources required for processing data without losing important information. It involves selecting and transforming relevant features from raw data to create new features that can be used to train machine learning models. It can improve the efficiency and accuracy of machine learning models by reducing the amount of redundant data. Therefore, it is highly beneficial when dealing with large datasets, as it can help reduce the computational complexity of machine learning algorithms. It can also help improve the interpretability of prediction models by identifying the most important predictor variables.

Li *et al.* (2022) proposed a variational autoencoder and genetic algorithm (VAEGA), an integration of VAEs and genetic algorithm (GA) for anomaly detection. In VAEGA, VAEs create a probabilistic dimensionality reduction, encoding high-dimensional data into a lower-dimensional latent space. The latent space characterizes the high-dimensional inputs, effectively capturing the data's essential features. A GA is then applied to analyze the abnormal subspace of outliers, enhancing the interpretability of the detection results. Their experiment showed that VAEGA outperforms principal component analysis (PCA) and other traditional dimensionality reduction methods. The VAE's ability to capture nonlinear relationships and analyses of abnormal subspace by GA are not inherent in PCA.

Deshai and Bhaskara Rao (2023) applied GPT to fake review detection. GPT captures deep contextualized word representations, enhancing the semantic understanding of review texts. Attention mechanisms enable GPT to focus on relevant parts of the text, improving the feature extraction process. Additionally, by leveraging its pre-trained knowledge, GPT can better distinguish between genuine and deceptive language patterns commonly found in fake reviews. Moreover, the model's ability to generate embeddings reflecting language nuances leads to accurate review detection.

G. Identification of independent generative factors

Independent generative factors are the underlying variables that can be manipulated to change specific attributes of data observations without affecting others (Laenen & Moens, 2022). These factors are the building blocks of the data set and, therefore, determine the specific attributes of the data. In OCBA, identifying these factors allows for more interpretable and explainable recommendation systems.

Laenen and Moens (2022) employed an explainable VAE framework (E-VAE), a variant of VAE. The E-VAE uses a two-level alignment to steer the disentanglement process toward discovering relevant factors of variations, including fine-grained attributes, and to ignore irrelevant visual variations. The E-VAE also aligns visual and textual attributes, ensuring an accurate and fine-grained visual contextualization of the textual attributes. The authors showed that the E-VAE can improve e-commerce search and recommendation systems through explainable and interpretable item representations.

H. Learning of normal data distribution

Generative AI can approximate probability distributions of the original data so that the model improves its performance by learning the distribution of normal data. GANs and VAEs are two popular models used for this purpose. While GANs are particularly useful for generating high-quality data similar to the original, VAEs are better suited for tasks requiring more structured and interpretable data representations.

Guo *et al.* (2023) applied GANs to anomaly detection to enable the model to learn the distribution of normal data, which aids in generating data similar to non-anomalies for more effective anomaly detection. They developed and used a novel graph GANs called RegraphGAN, which incorporated encoders to map real data to latent space, improving the training stability and efficiency.

Rabbi *et al.* (2023) employed GAN to identify authentic tribal products. In their study, GANs augment the dataset with synthetic images that resemble authentic tribal dresses. The augmented data improved the model's ability to identify tribal dresses accurately. In another study, Wang and Yang (2021) proposed M-GAN-XGBoost, which is a combination of LSTM, GAN, and XGBoost. They applied the model to sales forecasting, where GANs are used to replicate the distribution of consumer traffic data, aiming to produce data that resembles real user behavior patterns. In this process, GANs minimize the difference between generated and actual traffic data, refining the model's ability to predict consumer actions. Hao *et al.* (2022) applied deep convolutional generative adversarial nets (DCGAN; Radford *et al.*, 2015) to customer behavior prediction. The generator in DCGAN generates synthetic data that is distributed similarly to the real user behavior data, which is categorized into four types. Different patterns in the data represent the four types of user behaviors. By mimicking these patterns, the generator predicts the likelihood of each behavior type. In addition, the model creates a 2D color image representing user behaviors, facilitating visual analysis and prediction of emerging behaviors. In a further study, Nguyen and Cho (2020) introduced a mixture model for online behavior recommendation as a new approach to social media mining. Their model integrates VAE to capture the latent preferences of users based on their online activity data. It learns the underlying distribution of user behaviors, including repeat and new activities, to enhance recommendation quality. In a separate study, Perez-Castro *et al.* (2023) applied GPT to fake online review generation, highlighting the challenge that fake reviews generated by AI pose to existing fake review detection classifiers. They utilized GPT to generate text that closely resembles human writing. The model's effectiveness in mimicking the input text demonstrates its ability to learn from normal data distributions.

I. New item generation

One of the strengths of generative AI is its ability to generate new and unique outputs based on the training data. In developing a recommendation system, non-generative AI models typically rely on existing data to make recommendations but cannot generate new data. GANs can learn the patterns and characteristics of the existing data and generate new data similar to the existing data but with unique variations. It allows GANs to provide highly diverse and personalized recommendations to users. Singh *et al.* (2020) proposed a GAN-based model to generate new clothes by combining the latest fashion trends with the user's previous purchases. It allows for a highly personalized and tailored recommendation for each user. Moreover, the proposed method can establish cross-domain relationships between different types of clothes, such as vintage and new fashion clothes. It enables the model to generate new clothes that combine features from both domains, resulting in unique and fashionable designs.

J. Optimal policy selection

A policy, as a technical term in machine learning, refers to a function that maps the current state of the environment to an action. Reinforcement learning aims to find the optimal policy that maximizes the expected cumulative reward. In the development of a dialogue management system, Nabeel *et al.* (2019) applied GAN in dialogue generation to improve the fluency and diversity of generated dialogues. They proposed Cascade GAN (Cas-GAN), which combines GAN and reinforcement learning to model the relations between dialogues utilizing graph convolutional networks.

K. Rebalancing dataset

An imbalanced dataset is a dataset that consists of one or more classes that are significantly underrepresented in the training data (Kuhn & Johnson, 2013). When classification models are used on such data, they tend to undervalue the minority classes, resulting in a higher misclassification rate of minority class samples than those from majority classes (Sun *et al.*, 2009). Previous studies have used the synthetic minority oversampling technique (SMOTE) and its extensions to address the class imbalance problem (Kimura, 2022). SMOTE oversamples the minority class by creating synthetic samples from the minority class instead of creating copies of existing samples (Chawla *et al.*, 2002). Although less popular than SMOTE as an oversampling method, GAN and VAE can also be used for oversampling because they can generate more credible and varied synthetic examples that are indistinguishable from real data. Fiore *et al.* (2019) applied GAN to credit card fraud detection and showed that the GAN-based model outperforms previous methods like SMOTE regarding sensitivity and f-1 score. Ding *et al.* (2023) proposed VAEGAN, which integrates VAE and GAN. VAEGAN uses VAE's ability to learn a latent space representation and GAN's adversarial training to generate more realistic data. This combination enhances the generation of realistic and diverse minority-class data for imbalanced datasets. According to their results, VAEGAN outperforms previous methods, such as GAN, VAE, and SMOTE, in terms of precision, f1-score, and other indicators.

L. Regularization for realistic output

Overfitting occurs when a machine learning model learns the training data too well and fails to generalize to new data. In GAN, overfitting means that the generator network produces data that is too similar to the training data and not sufficiently diverse. GAN uses the adversarial losses as auxiliary regularization terms to prevent overfitting. The regularization enforces the generator network to learn the underlying distribution of the real data, which leads to more diverse and realistic output. Zhao *et al.* (2020) applied GAN to multi-human parsing. Multi-human parsing is a challenging task because it involves multiple people interacting with each other. Auxiliary regularization in GAN can help refine the realism of the multi-human parsing results by acting as a regularization term in the adversarial learning process. This approach can improve the quality of the generated images.

5. DISCUSSION

5.1. Summary of findings

Generative AI is currently utilized in various areas of OCBA and can improve OCBA performance. This review identified 14 utilization areas and 12 functions of generative AI in OCBA. These 12 functions can be grouped into three broader functions: personalization, prediction accuracy, and modeling efficiency.

Generative AI algorithms are better at personalization than traditional machine learning because they can create entirely new data, while traditional machine learning models rely on pre-existing data for making predictions. The distinct feature of generative AI in crafting new data is pivotal to produce effective personalization. It enables the model to produce data specifically tailored to individual users, moving away from the constraints of using existing data that might not be relevant or accurate for the user. This tailored approach can achieve high levels of personalization, improving the customer experience and increasing customer engagement.

Generative AI techniques such as GAN and VAE improve prediction accuracy compared to traditional machine learning because they can generate synthetic samples that mimic real data distribution. This ability of generative AI improves prediction performance in various ways, such as rebalancing datasets, learning normal data distribution, feature extraction, and addressing data sparsity and exposure bias problems.

Generative Al algorithms can develop more efficient and cost-effective models compared to traditional machine learning methods because generative AI decreases the need for human intervention. Traditional methods often require manual work to label and annotate data, which can be time-consuming and expensive. For instance, Han *et al.* (2023) noted that generative AI with pre-trained language models can automate the generation of synthetic training data, eliminating the need for manual annotation.

5.2. Limitations

It must be stated that this review has several limitations. First, the studies examined in this review are not exhaustive. Only articles published in peer-reviewed journals were included and so conference proceedings and book chapters were excluded. Therefore, this review may not capture some of the state-of-the-art methods. However, this study focused on rigorously scrutinized, high-quality studies to draw conclusions from robust findings.

The application areas of generative AI identified in this review did not include popular consumer behavior analytics, such as customer churn prediction and segmentation. Although this study collected a wide range of articles from four databases using various search terms, some studies may have been missed because the search was conducted without specifying the areas of OCBA. Because this study aimed to conduct an exploratory review of the use of generative AI in OCBA, we did not search for papers focusing on specific practices. Future studies can review the use of generative AI in specific OCBA domains, such as customer churn prediction and customer segmentation.

5.3. Implications

In future research and practice, analysts can develop generative AI models for popular OCBAs, such as customer segmentation, market basket analysis, customer lifetime value prediction, and customer churn prediction. In these analytics, traditional machine learning models face problems such as high-dimensional data handling, data sparsity, and class imbalance. These problems make it difficult for machine learning models to achieve high performance. Additionally, in supervised learning, collecting a large amount of labeled data is usually difficult, and the manual labeling process is arduous. Furthermore, there is a cold start problem where models cannot make inferences about customers or items for which sufficient data has yet to be obtained. Generative AI can address these problems by generating synthetic data that is similar to but slightly different from the original data.

Future research could also examine the impact of generative AI in OCBA on consumer experience from more diverse perspectives. The studies reviewed in this study assumed that generative AI models provide customers with positive experiences, such as high levels of satisfaction and fulfillment of needs. However, using generative AI can also result in negative customer experiences. For example, prior research suggests that customers prefer interacting with humans rather than virtual agents (Oshrat *et al.*, 2022). This tendency has been shown to intensify when customers perceive dissatisfaction with their interactions with virtual agents (Ashfaq *et al.*, 2020). Therefore, whether advanced generative AI enables virtual assistants to provide higher-quality services, and if so, do customers have positive attitudes toward their interactions with virtual agents, are questions that should be addressed in future research.

Since Open AI released ChatGPT in November 2022, public attention on generative AI seems to have concentrated on information gathering, sentence generation, and chatting with AI using large-scale language models. However, as this study has shown, generative AI has various functions contributing to OCBA. Practitioners can explore ways to make OCBA more effective in the future by integrating these various functions of generative AI with other disruptive technologies.

For example, in the medical industry, virtual reality is being used to train for surgery (Lungu *et al.*, 2021), and in the construction industry, augmented reality simulations of three-dimensional spaces are used to design buildings (Adebowale, &

Agumba, 2022). These technologies give consumers a more realistic experience of using those products and services. Consumer responses to this realistic experience can be used as high-quality data for OCBA by generative AI. Such integration would contribute to developing more effective recommendation systems and producing more attractive advertisements. Besides, generative AI will also contribute to efficiently generating virtual and augmented reality.

Although generative AI has much potential for OCBA, researchers and practitioners should be aware of the challenges related to its use. First, generative AI is not bias-free. If trained on biased data, a generative AI model will produce biased outcomes, resulting in unreliable predictions, potentially unfair and raising ethical concerns. Additionally, as generative AI's use in OCBA grows, there could be an increase in consumer resistance to the continuous data collection in online platforms. To address this, businesses must establish and communicate clear data privacy policies. These policies should detail how consumer data is collected, shared, and used, ensuring the protection of consumer privacy. It is also crucial for businesses to provide consumers with options regarding their data contribution, including mechanisms to opt-out of data sharing and requests for data deletion.

6. REFERENCES

Adebowale, O. J., & Agumba, J. N. (2022). Applications of augmented reality for construction productivity improvement: a systematic review. *Smart and Sustainable Built Environment*, (ahead-of-print). https://doi.org/10.1108/SASBE-06-2022-0128

Almahmood, R. J. K., & Tekerek, A. (2022). Issues and solutions in deep learning-enabled recommendation systems within the e-commerce field. *Applied Sciences*, *12*(21), 11256. https://doi.org/10.3390/app122111256

Anshari, M., Almunawar, M. N., Lim, S. A., & Al-Mudimigh, A. (2019) Customer relationship management and big data enabled: Personalization & customization of services. *Applied Computing and Informatics*, *15*(2), 94-101. https://doi.org/10.1016/j.aci.2018.05.004

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics*, *54*, 101473. https://doi.org/10.1016/j.tele.2020.101473

Asperti, A., Evangelista, D., & Loli Piccolomini, E. (2021). A survey on variational autoencoders from a green AI perspective. *SN Computer Science*, *2*(4), 301. https://doi.org/10.1007/s42979-021-00702-9

Baek, T. H. (2023). Digital advertising in the age of generative AI. *Journal of Current Issues and Research in Advertising*, *44*(3), 249-251. https://doi.org/10.1080/10641734.2023.2243496

Bawack, R. E., Wamba, S. F., Carillo, K. D. A., & Akter, S. (2022). Artificial intelligence in e-commerce: A bibliometric study and literature review. *Electronic Markets*, *32*(1), 297-338. https://doi.org/10.1007/s12525-022-00537-z

Bond-Taylor, S., Leach, A., Long, Y., & Willcocks, C. G. (2022). Deep generative modelling: A comparative review of vaes, gans, normalizing flows, energy-based and autoregressive models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(11), 7327-7347. https://doi.org/10.1109/TPAMI.2021.3116668

Breen, R. L. (2006). A practical guide to focus-group research. *Journal of Geography in Higher Education*, *30*(3), 463-475. https://doi.org/10.1080/03098260600927575

Casanova, A., Careil, M., Romero-Soriano, A., Pal, C. J., Verbeek, J., & Drozdzal, M. (2023). *Controllable image generation via collage representations. arXiv preprint arXiv:2304.13722*.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321-357. https://doi.org/10.1613/jair.953

Deshai, N., & Bhaskara Rao, B. (2023). Unmasking deception: A CNN and adaptive PSO approach to detecting fake online reviews. *Soft Computing*, 27(16), 11357-11378. https://doi.org/10.1007/s00500-023-08507-z

Ding, Y., Kang, W., Feng, J., Peng, B., & Yang, A. (2023). Credit card fraud detection based on improved Variational Autoencoder Generative Adversarial Network. *IEEE Access*, 11, 83680-83691. https://doi.org/10.1109/ACCESS.2023.3302339

Du, C., Yu, F., Jiang, M., Hua, A., Wei, X., Peng, T., & Hu, X. (2023). Vton-scfa: A virtual try-on network based on the semantic constraints and flow alignment. *IEEE Transactions on Multimedia*, *25*, 777-791. https://doi.org/10.1109/TMM.2022.3152367

Fincato, M., Cornia, M., Landi, F., Cesari, F., & Cucchiara, R. (2022). Transform, warp, and dress: A new transformation-guided model for virtual try-on. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *18*(2), 1-24. https://doi.org/10.1145/3491226

Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, *479*, 448-455. https://doi.org/10.1016/j.ins.2017.12.030

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, *30*(4), 681-694. https://doi.org/10.1007/s11023-020-09548-1

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139-144. https://doi.org/10.1145/3422622

Goti, A., Querejeta-Lomas, L., Almeida, A., de la Puerta, J. G., & López-de-Ipiña, D. (2023). Artificial intelligence in business-to-customer fashion retail: A literature review. *Mathematics*, *11*(13), 2943. https://doi.org/10.3390/math11132943

Grove, S. J., & Fisk, R. P. (1992). Observational data collection methods for services marketing: An overview. *Journal of the Academy of Marketing Science*, *20*(3), 217-224. https://doi.org/10.1007/BF02723408

Guo, D., Liu, Z., & Li, R. (2023). RegraphGAN: A graph generative adversarial network model for dynamic network anomaly detection. *Neural Networks*, *166*, 273-285. https://doi.org/10.1016/j.neunet.2023.07.026

Guo, J., Zheng, C., Li, S., Jia, Y., & Liu, B. (2022). BiInfGCN: Bilateral information augmentation of graph convolutional networks for recommendation. *Mathematics*, *10*(17), 3042. https://doi.org/10.3390/math10173042

Han, Y., Nanda, G., & Moghaddam, M. (2023). Attribute-sentiment-guided summarization of user opinions from online reviews. *Journal of Mechanical Design*, *145*(4), 041402. https://doi.org/10.1115/1.4055736

Hao, Y., Zhao, D., Li, H., Ip, W. H., & Liu, Y. (2022). User Behavior Prediction Based on DCGAN: The Case of Sina Weibo. *Journal of Internet Technology*, 23(6), 1367-1376. https://doi.org/10.53106/160792642022112306019

Hu, K., Qiu, L., Zhang, S., Wang, Z., Fang, N., & Zhou, H. (2023). A novel neighbor selection scheme based on dynamic evaluation towards recommender systems. *Science Progress*, *106*(2), 368504231180090. https://doi.org/10.1177/00368504231180090

Jiang, L. A., Yang, Z., & Jun, M. (2013). Measuring consumer perceptions of online shopping convenience. *Journal of Service Management*, *24*(2), 191-214. https://doi.org/10.1108/09564231311323962

Kashyap, D. A. K., Sahu, I., & Kumar, D. A. (2022). Artificial intelligence and its applications in e-commerce —A review analysis and research agenda. *Journal of Theoretical and Applied Information Technology*, *100*(24), 7347-7465.

Kimura, T. (2022). Customer churn prediction with hybrid resampling and ensemble learning. *Journal of Management Information & Decision Sciences*, *25*(1).

Kimura, T. (2023). Assessment of personal values for data-driven human resource management. *Data Science Journal*, *22*. https://doi.org/10.5334/dsj-2023-022

Kmet, L., M., Lee, R., C., Cook, L., & S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields*. Alberta Heritage Foundation for Medical Research (Alberta Heritage Foundation for Medical Research).

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (26th ed). Springer.

Laenen, K., & Moens, M. F. (2022). Learning explainable disentangled representations of e-commerce data by aligning their visual and textual attributes. *Computers*, *11*(12), 182. https://doi.org/10.3390/computers11120182

Li, J., Zhang, J., Bah, M. J., Wang, J., Zhu, Y., Yang, G., Li, L., & Zhang, K. (2022). An auto-encoder with genetic algorithm for high dimensional data: Towards accurate and interpretable outlier detection. *Algorithms*, *15*(11), 429. https://doi.org/10.3390/a15110429

Li, K., Zhang, J., & Forsyth, D. (2023). POVNet: Image-based virtual try-on through accurate warping and residual. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(10), 12222-12235. https://doi.org/10.1109/TPAMI.2023.3283302

Lu, X., Liu, J., Gan, S., Li, T., Xiao, Y., & Liu, Y. (2022). Recommendation model based on dynamic interest group identification and data compensation. *IEEE Transactions on Network and Service Management*, *19*(1), 89-99. https://doi.org/10.1109/TNSM.2021.3112702

Lungu, A. J., Swinkels, W., Claesen, L., Tu, P., Egger, J., & Chen, X. (2021). A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: an extension to different kinds of surgery. *Expert Review of Medical Devices*, 18(1), 47-62. https://doi.org/10.1080/17434440.2021.1860750

Margaris, D., Vassilakis, C., & Spiliotopoulos, D. (2022). On producing accurate rating predictions in sparse collaborative filtering datasets. *Information*, *13*(6), 302. https://doi.org/10.3390/info13060302

Nabeel, M., Riaz, A., & Zhenyu, W. (2019). Cas-gans: An approach of dialogue policy learning based on gan and rl techniques. *International Journal of Advanced Computer Science and Applications*, *10*(7). https://doi.org/10.14569/IJACSA.2019.0100766

Nallapati, R., Zhou, B., Gulcehre, C., & Xiang, B. (2016). *Abstractive text summarization using sequence-to-sequence rnns and beyond. arXiv preprint arXiv:1602.06023*.

Necula, S. C., & Păvăloaia, V. (2023). AI-driven recommendations: A systematic review of the state of the art in e-commerce. *Applied Sciences*, *13*(9), 5531. https://doi.org/10.3390/app13095531

Ngai, E. W., & Wu, Y. (2022). Machine learning in marketing: A literature review, conceptual framework, and research agenda. Journal of Business Research, 145, 35-48. https://doi.org/10.1016/j.jbusres.2022.02.049

Nguyen, M. D., & Cho, Y. S. (2020). A variational autoencoder mixture model for online behavior recommendation. *IEEE Access*, *8*, 132736-132747. https://doi.org/10.1109/ACCESS.2020.3010508

Notin, P., Hernández-Lobato, J. M., & Gal, Y. (2021). Improving black-box optimization in VAE latent space using decoder uncertainty. *Advances in Neural Information Processing Systems*, *34*, 802-814.

Oshrat, Y., Aumann, Y., Hollander, T., Maksimov, O., Ostroumov, A., Shechtman, N., & Kraus, S. (2022). *Efficient customer service combining human operators and virtual agents. arXiv preprint arXiv:2209.05226*.

Pan, Z., Yu, W., Yi, X., Khan, A., Yuan, F., & Zheng, Y. (2019). Recent progress on generative adversarial networks (GANs): A survey. *IEEE Access*, *7*, 36322-36333. https://doi.org/10.1109/ACCESS.2019.2905015

Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., & Lakshminarayanan, B. (2021). Normalizing flows for probabilistic modeling and inference. *Journal of Machine Learning Research*, *22*(1), 2617-2680.

Perez-Castro, A., Martínez-Torres, M. R., & Toral, S. L. (2023). Efficiency of automatic text generators for online review content generation. *Technological Forecasting and Social Change*, *189*, 122380. https://doi.org/10.1016/j.techfore.2023.122380

Policarpo, L. M., da Silveira, D. E., da Rosa Righi, R., Stoffel, R. A., da Costa, C. A., Barbosa, J. L. V., ... & Arcot, T. (2021). Machine learning through the lens of e-commerce initiatives: An up-to-date systematic literature review. Computer Science Review, 41, 100414. https://doi.org/10.1016/j.cosrev.2021.100414

Postolache, E., Mariani, G., Mancusi, M., Santilli, A., Cosmo, L., & Rodola, E. (2023). *Latent autoregressive source separation. arXiv preprint arXiv:2301.08562*.

Rabbi, M. F., Sultan, M. N., Hasan, M., & Islam, M. Z. (2023). Tribal dress identification using convolutional neural network. *Journal of Information Hiding and Multimedia Signal Processing*, *14*(3), 72-80.

Radford, A., Metz, L., & Chintala, S. (2015). *Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434*.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pretraining*.

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., & Koffel, J. B. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews*, *10*, 1-19.

Roy, D., Santra, S., & Chanda, B. (2022). LGVTON: A landmark guided approach for model to person virtual try-on. *Multimedia Tools and Applications*, *81*(4), 5051-5087. https://doi.org/10.1007/s11042-021-11647-9

Sajid, S., Rashid, R. M., & Haider, W. (2022). Changing trends of consumers' online buying behavior during COVID-19 pandemic with moderating role of payment mode and gender. *Frontiers in Psychology*, *13*, 919334. https://doi.org/10.3389/fpsyg.2022.919334

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, *3*(3), 210-229. https://doi.org/10.1147/rd.33.0210

Singh, M., Bajpai, U., V., V., & Prasath, S. (2020). Generation of fashionable clothes using generative adversarial networks: A preliminary feasibility study. *International Journal of Clothing Science and Technology*, *32*(2), 177-187.

Sun, R., Dai, H., Li, L., Kearnes, S., & Dai, B. (2021). Towards understanding retrosynthesis by energy-based models. *Advances in Neural Information Processing Systems*, *34*, 10186-10194.

Sun, Y., Wong, A. K. C., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence*, *23*(4), 687-719. https://doi.org/10.1142/S0218001409007326

Terzioğlu, S., Çoğalmış, K. N., & Bulut, A. (2022). Ad creative generation using reinforced generative adversarial network. *Electronic Commerce Research*, 1-17. https://doi.org/10.1007/s10660-022-09564-6

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*.

Viktoratos, I., & Tsadiras, A. (2021). Personalized advertising computational techniques: A systematic literature review, findings, and a design framework. *Information*, *12*(11), 480. https://doi.org/10.3390/info12110480

Wang, S., & Yang, Y. (2021). M-GAN-XGBOOST model for sales prediction and precision marketing strategy making of each product in online stores. *Data Technologies and Applications*, *55*(5), 749-770. https://doi.org/10.1108/DTA-11-2020-0286

Wang, X. W., Chen, Y., & Liu, Y. Y. (2020). Link prediction through deep generative model. *iScience*, *23*(10), 101626. https://doi.org/10.1016/j.isci.2020.101626

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, xiii-xxiii.

Wei, S., Zhou, X., An, X., Yang, X., & Xiao, Y. (2023). A heterogeneous e-commerce user alignment model based on data enhancement and data representation. *Expert Systems with Applications*, *228*, 120258. https://doi.org/10.1016/j.eswa.2023.120258

Yang, Y., Ding, H., Liu, Q., & Quan, X. (2022). *UBARv2: Towards mitigating exposure bias in task-oriented dialogs. arXiv preprint arXiv:2209.07239*.

Yu, A., & Grauman, K. (2020). Densifying supervision for fine-grained visual comparisons. *International Journal of Computer Vision*, *128*(10-11), 2704-2730. https://doi.org/10.1007/s11263-020-01344-9

Zhang, H., Sun, Y., Liu, L., & Xu, X. (2020). CascadeGAN: A category-supervised cascading generative adversarial network for clothes translation from the human body to tiled images. *Neurocomputing*. Cascade, *382*, 148-161. https://doi.org/10.1016/j.neucom.2019.11.085

Zhang, H., Wang, X., Liu, L., Zhou, D., & Zhang, Z. (2020). Warpclothingout: A stepwise framework for clothes translation from the human body to tiled images. *IEEE Multimedia*, *27*(4), 58-68. https://doi.org/10.1109/MMUL.2020.3014037

Zhang, H., Wong, R. K., & Chu, V. W. (2021). Hybrid variational autoencoder for recommender systems. *ACM Transactions on Knowledge Discovery from Data*, *16*(2), 1-37. https://doi.org/10.1145/3470659

Zhang, X., Sun, Y., & Liu, L. (2021). An improved generative adversarial network for translating clothes from the human body to tiled image. *Neural Computing and Applications*, *33*(14), 8445-8457. https://doi.org/10.1007/s00521-020-05598-9

Zhao, J., Li, J., Liu, H., Yan, S., & Feng, J. (2020). Fine-grained multi-human parsing. *International Journal of Computer Vision*, *128*(8-9), 2185-2203. https://doi.org/10.1007/s11263-019-01181-5

Zhou, X., Han, X., Li, H., Wang, J., & Liang, X. (2022). Cross-domain image retrieval: Methods and applications. *International Journal of Multimedia Information Retrieval*, *11*(3), 199-218. https://doi.org/10.1007/s13735-022-00244-7