Journal of
*Theoretical and Applied Electronic Commerce Research*

*Review*

# A Brief Survey of Machine Learning and Deep Learning Techniques for E-Commerce Research

**Xue Zhang [1], Fusen Guo [2,*], Tao Chen [1], Lei Pan [2,*], Gleb Beliakov [2] and Jianzhang Wu [1,2,*]**

[1] School of Business, Ningbo University, Ningbo 315211, China; 18758334787@163.com (X.Z.); chentao@nbu.edu.cn (T.C.)
[2] School of Information Technology, Deakin University, Burwood, Geelong 3125, Australia; gleb@deakn.edu.au
* Correspondence: dobbysen430@gmail.com (F.G.); l.pan@deakin.edu.au (L.P.); wujianzhang@nbu.edu.cn (J.W.)

**Abstract:** The rapid growth of e-commerce has significantly increased the demand for advanced techniques to address specific tasks in the e-commerce field. In this paper, we present a brief survey of machine learning and deep learning techniques in the context of e-commerce, focusing on the years 2018–2023 in a Google Scholar search, with the aim of identifying state-of-the-art approaches, main topics, and potential challenges in the field. We first introduce the applied machine learning and deep learning techniques, spanning from support vector machines, decision trees, and random forests to conventional neural networks, recurrent neural networks, generative adversarial networks, and beyond. Next, we summarize the main topics, including sentiment analysis, recommendation systems, fake review detection, fraud detection, customer churn prediction, customer purchase behavior prediction, prediction of sales, product classification, and image recognition. Finally, we discuss the main challenges and trends, which are related to imbalanced data, over-fitting and generalization, multi-modal learning, interpretability, personalization, chatbots, and virtual assistance. This survey offers a concise overview of the current state and future directions regarding the use of machine learning and deep learning techniques in the context of e-commerce. Further research and development will be necessary to address the evolving challenges and opportunities presented by the dynamic e-commerce landscape.

**Keywords:** machine learning; deep learning; e-commerce; sentiment analysis; recommendation system; fake review detection; customer churn prediction; financial risk recognition



**Citation:** Zhang, X.; Guo, F.; Chen T.; Pan, L.; Beliakov, G.; Wu, J. A Brief Survey of Machine Learning and Deep Learning Techniques for E-Commerce Research. *J. Theor. Appl. Electron. Commer. Res.* **2023**, *18*, 2188–2216. https://doi.org/10.3390/jtaer18040110

Academic Editor: Ljiljana Kašćelan

Received: 29 July 2023
Revised: 30 October 2023
Accepted: 13 November 2023
Published: 4 December 2023



## 1. Introduction

Machine learning is a subset of artificial intelligence (AI) that focuses on developing algorithms and models capable of automatically learning, identifying patterns, and making predictions or decisions from data [1]. The field of machine learning encompasses a broad array of methods and algorithms. Some prominent examples of supervised learning methods include linear regression, logistic regression, decision tree, random forest, support vector machine, and artificial neural network techniques [2]. On the other hand, unsupervised learning methods typically include K-means clustering, hierarchical clustering, principal component analysis, and matrix factorization [3].

Deep learning is a specialized branch of machine learning that emphasizes training artificial neural networks with multiple hidden layers, enabling them to acquire hierarchical representations of data [4]. Exceptional accomplishments in diverse domains, including image classification, object detection, speech recognition, and language translation, have been achieved through the use of deep learning approaches [5]. The ability to automatically learn intricate features from raw data has positioned deep learning as a pivotal component in modern AI systems [6].

E-commerce refers to the buying and selling of goods and services over the Internet, which involves online transactions, electronic payments, and digital interactions between businesses and customers [7]. E-commerce has become increasingly popular due to its convenience, wide product range, and global accessibility. It also provides a favorable environment for the application of machine learning and deep learning techniques, due to the availability of vast data sets, the need for personalized experiences, the challenges of fraud detection and security, the potential for supply chain optimization, and the importance of customer sentiment analysis [8]. By leveraging these techniques, e-commerce businesses can enhance customer satisfaction, improve operational efficiency, drive sales, and gain a competitive edge in the digital marketplace [9].

To comprehensively explore the integration and application of machine learning and deep learning in e-commerce, we conducted an extensive investigation using Google Scholar. We opted for Google Scholar due to its robust search functionality, citation tracking, and seamless integration with citation management tools such as Zotero, EndNote, or Mendeley. To address any concerns about source quality, we supplemented our research with other databases, including Scopus, Web of Science, and IEEE Xplore, in order to ensure that a thorough review was conducted. A search query combining keywords such as 'machine learning', 'deep learning', and 'e-commerce' was employed. After reviewing the top 300 ranked results output by Google Scholar, 158 journal articles and conference papers were selected for this survey. We endeavor to provide a lucid overview of the central themes and primary challenges investigated in recent studies at the intersection of machine learning, deep learning, and e-commerce. By restricting our focus to the years 2018–2023, we ensure the inclusion of the most recent developments and trends in these dynamically evolving fields. Table 1 presents the number of publications from 2018 to 2023, revealing a notable upward trend. It is essential to highlight that data for 2023 were available only up to August.

**Table 1.** Literature selected for each year (2018–2023).

| Year | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 |
|---|---|---|---|---|---|---|
| Number | 10 | 16 | 22 | 26 | 40 | 54 |

Table 2 presents the 30 most frequently occurring words in the titles of these 158 papers, while Figure 1 illustrates the 30 most frequent 2-gram combinations [10] found in these titles. These visual representations collectively provide valuable insight into the central themes and prevailing methodologies explored within the surveyed papers. Upon examining both the table and figure, it becomes evident that certain terms had a prominent presence in the titles. Single words such as 'learning', 'deep', 'e-commerce', and 'machine', as well as word groups like 'deep learning', 'machine learning', and 'learning approach' featured prominently. This underscores a significant alignment between the surveyed papers and the overarching topic of the survey. Furthermore, an analysis of the top 30 highest-frequency words listed in the table revealed key terms including 'sentiment', 'detection', 'recommendation', 'product', 'fraud', 'reviews', 'customer', and 'churn', along with word groups including 'sentiment analysis', 'customer churn', 'recommendation system', and 'card fraud'. These terms serve as indicators of the primary subjects within the realm of e-commerce research. Additionally, noteworthy words such as 'neural', 'reinforcement', 'network', 'algorithm', 'hybrid', 'deep reinforcement', 'hybrid deep', and 'artificial intelligence' shed light on the prevailing methodologies employed in these studies.

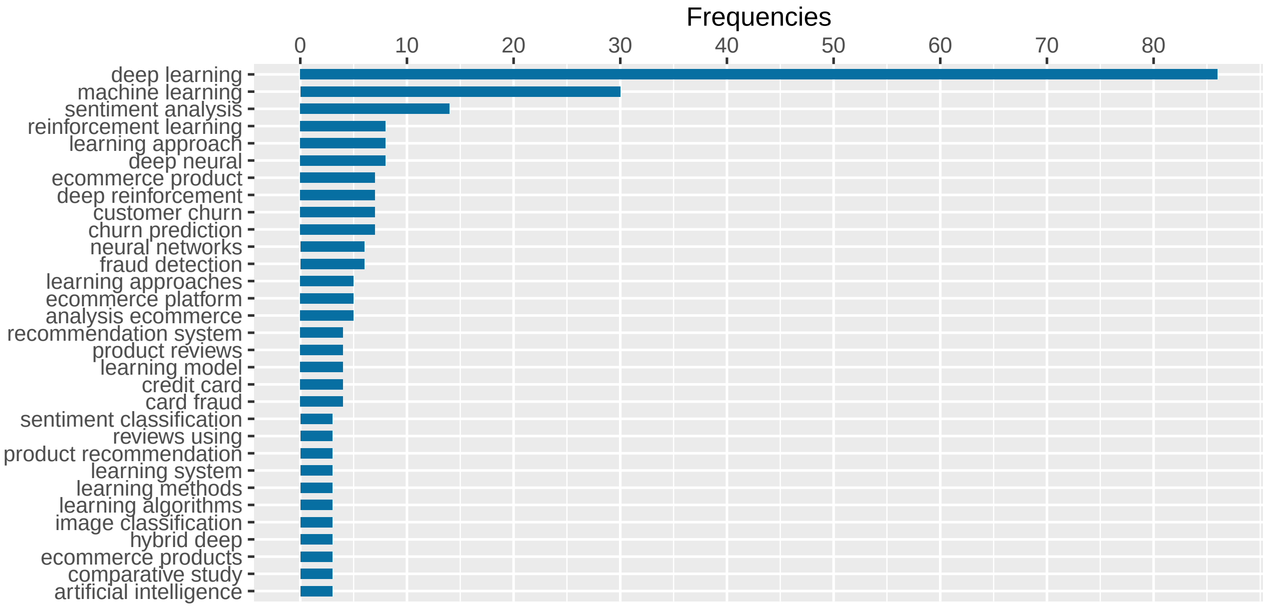


**Figure 1.** The 2-gram combinations most frequently found in the titles of the 158 papers.

**Table 2.** The top 30 most frequent words found in the titles of the 158 papers.

| Ranking | Word | Frequency | Ranking | Word | Frequency | Ranking | Word | Frequency |
|---|---|---|---|---|---|---|---|---|
| 1 | learning | 159 | 11 | product | 19 | 21 | network | 9 |
| 2 | deep | 132 | 12 | fraud | 18 | 22 | churn | 9 |
| 3 | ecommerce | 77 | 13 | approach | 16 | 23 | online | 8 |
| 4 | machine | 51 | 14 | neural | 16 | 24 | reinforcement | 8 |
| 5 | analysis | 38 | 15 | reviews | 16 | 25 | comparative | 8 |
| 6 | sentiment | 28 | 16 | classification | 16 | 26 | hybrid | 8 |
| 7 | detection | 24 | 17 | system | 15 | 27 | card | 8 |
| 8 | model | 24 | 18 | review | 12 | 28 | credit | 8 |
| 9 | recommendation | 21 | 19 | customer | 12 | 29 | algorithm | 7 |
| 10 | prediction | 20 | 20 | data | 11 | 30 | techniques | 7 |

Table 3 showcases the 30 most frequently occurring words in the abstracts of the 158 papers, while Figure 2 illustrates the 30 most frequent 2-gram combinations [10] found within these abstracts. These findings not only corroborate our previous conclusions, but also provide additional insight into the background and methodologies applied in these studies. Of particular note are terms such as 'features', 'traditional', 'techniques', 'accuracy', 'dataset', 'classification', 'information', 'sales', 'CNN', 'convolutional', 'Bayes', 'forecasting', 'training', and 'Amazon', which prominently featured in the abstracts. Additionally, word groups like 'e-commerce', 'credit card', 'convolutional neural', 'random forest', 'product review', and 'reinforcement learning' emphasize that the existing literature has focused on diverse aspects, including feature analysis, traditional techniques, classification accuracy, data set characteristics, sales information, and the utilization of specific methods such as CNN, recurrent networks, random forest, Bayes, and reinforcement learning for forecasting and classification. These insights contribute to a more comprehensive understanding of the research landscape within the e-commerce domain.

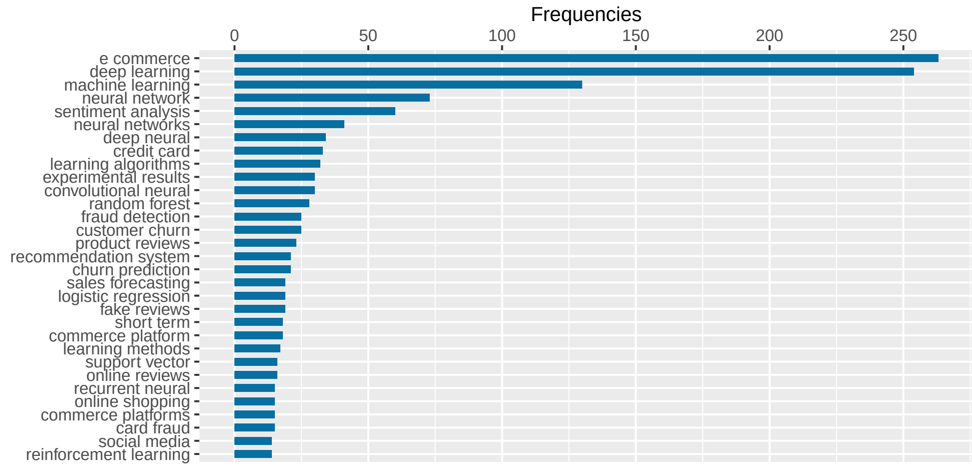


**Figure 2.** The 2-gram combinations found in the abstracts of the 158 papers.

**Table 3.** The top 30 highest frequency words found in the abstracts of the 158 papers.

| Ranking | Word | Frequency | Ranking | Word | Frequency | Ranking | Word | Frequency |
|---|---|---|---|---|---|---|---|---|
| 1 | learning | 381 | 11 | models | 110 | 21 | classification | 83 |
| 2 | deep | 309 | 12 | online | 108 | 22 | products | 82 |
| 3 | ecommerce | 252 | 13 | accuracy | 107 | 23 | system | 77 |
| 4 | model | 246 | 14 | sentiment | 103 | 24 | detection | 75 |
| 5 | data | 175 | 15 | network | 99 | 25 | information | 74 |
| 6 | reviews | 160 | 16 | prediction | 96 | 26 | features | 73 |
| 7 | product | 145 | 17 | proposed | 92 | 27 | cnn | 69 |
| 8 | machine | 144 | 18 | recommendation | 88 | 28 | algorithms | 68 |
| 9 | analysis | 131 | 19 | customer | 84 | 29 | sales | 68 |
| 10 | neural | 125 | 20 | dataset | 84 | 30 | performance | 65 |

## 2. The Utilized Machine Learning And Deep Learning Techniques

In this section, we provide a brief overview of the machine learning and deep learning techniques employed in the 158 surveyed papers.

### *2.1. Machine Learning Techniques*

- Support vector machine (SVM) [11] is a machine learning model used for classification and regression. An SVM operates by identifying an optimal hyperplane that maximizes the margin between distinct classes, which is determined by critical data points known as support vectors. It can handle both linearly separable and non-linearly separable data using the kernel trick, such as the linear, polynomial, radial basis function, and sigmoid kernels. It is particularly effective for binary and even multi-class classification problems [12].
- Decision Tree [13] is a model used for prediction tasks, functioning by segmenting the predictor space into simple regions for analysis. It uses a tree-like structure to make decisions based on feature values. At each internal node of the tree, a decision or splitting criterion is applied to determine the best feature and threshold for splitting the data [14]. In classification tasks, each leaf node represents a class label while, in regression tasks, the leaf nodes contain the predicted continuous value in that subset.
- Random Forest [15,16] is an ensemble learning method that combines multiple decision trees to make predictions. It enhances classification and regression tasks by training multiple trees on various sub-samples of the data set and aggregating the predictions of individual trees to improve accuracy and prevent over-fitting [17].
- Naïve Bayes [18] is based on the assumption that features are independently and naïvely unrelated to each other. It utilizes the Bayes theorem to calculate the posterior

probabilities of classes based on observed feature values. Depending on the assumed distribution type of the features, there are Gaussian, Multinomial, and Bernoulli Naïve Bayes algorithms. Naïve Bayes is widely recognized for its simplicity and efficiency in training and prediction tasks, making it popular for various applications [19].

- Logistic regression [20] utilizes the logistic function or the sigmoid function to estimate the probabilities of inputs belonging to different classes. This method can be extended to softmax regression or multinomial logistic regression by replacing the sigmoid function with the softmax function. Logistic and softmax regression provide straightforward and interpretable approaches to classification problems, allowing for accurate and probabilistic predictions [21].
- Principal component analysis (PCA) [22] is a linear modeling technique used to map high-dimensional input features to a lower-dimensional space, typically referred to as latent factors or principal components. PCA aims to transform the original data into a set of orthogonal components that explain the maximum variance in the data [23].
- Matrix factorization algorithms [24,25] work by decomposing the original matrix into two or more lower-dimensional matrices that represent latent factors. These algorithms aim to find lower-rank representations of the data by uncovering the underlying structure or patterns within the matrix [26].
- K-nearest neighbors (KNN) [27] is a non-parametric algorithm that predicts the class label (for classification) or the target value (for regression) of a test instance based on its similarity to its K nearest neighbors in the training data. In classification, the majority vote among the neighbors determines the class label while, in regression, the average (or weighted average) of the target values is taken [28].

### *2.2. Deep Learning Techniques*

Deep learning approaches continue to evolve rapidly, with new architectures, algorithms, and techniques having been developed to address various challenges in different domains. Their ability to learn complex representations from data has significantly advanced the field of artificial intelligence and contributed to various groundbreaking applications [29].

- An Artificial Neural Network (ANN) [30] is a computational model inspired by the structure and functionality of biological neural networks in the human brain. It is composed of interconnected artificial neurons or nodes, organized into layers including the input layer, hidden layers, and output layer. The connections between neurons have associated weights, which are adjusted iteratively by propagating the error from the output layer back to the input layer, guided by a defined objective or loss function [31].
- A Convolutional Neural Network (CNN) [32,33] consists of convolutional layers that apply filters to extract features from input data, followed by pooling layers to reduce the spatial dimensions. They have demonstrated exceptional performance in image classification, object detection, and image segmentation [34].
- The Visual Geometry Group network (VGG) [35] is a deep convolutional neural network architecture (e.g., with 16–19 convolutional layers) developed by the Visual Geometry Group. It showcases the effectiveness of deep convolutional neural networks in capturing complex image features and hierarchies [36].
- A Temporal Convolutional Network (TCN) [37] utilizes dilated convolutional layers to capture temporal patterns and dependencies in the input data. These dilated convolutions enable an expanded receptive field without significantly increasing the number of parameters or computational complexity.
- Recurrent Neural Networks (RNNs) [30] are designed to process sequential data and utilize recurrent connections that enable information to be carried across different time steps. The key characteristic of an RNN is its recurrent connections, which create a loop-like structure and allow information to flow in cycles, enabling the network

to maintain a form of memory or context to process and remember information from previous steps [38].

- Long Short-Term Memory (LSTM) [39] is a type of RNN architecture that excels at capturing long-term dependencies and processing sequential data. It utilizes a memory cell and a set of gates that regulate the flow of information; in particular, the memory cell retains information over time, the input gate determines which values to update in the memory cell, the forget gate decides what information to discard from the memory cell, and the output gate selects the relevant information to be output at each time step [38].
- Bidirectional Long Short-Term Memory (BiLSTM) [40] combines two LSTMs that process the input sequence in opposite directions: one LSTM processes the sequence in the forward direction, while the other processes it in the backward direction. This bidirectional processing allows the model to capture information from both past and future contexts, providing a more comprehensive understanding of the input sequence. It has demonstrated strong performance in various natural language processing tasks.
- The Gated Recurrent Unit (GRU) [41] is a simplified alternative to the LSTM network, offering comparable performance with fewer parameters and less computation. In GRU, the update gate determines the amount of the previous hidden state to retain and the extent to which the new input is incorporated. The reset gate controls how much of the previous hidden state is ignored and whether the hidden state should be reset, based on the current input [42].
- The BiGRU [42,43] is an extension of the standard GRU, which processes the input sequence in both forward and backward directions simultaneously, resulting in a more comprehensive understanding of the sequence.
- The attention-based BiGRU [43,44] adopts attention mechanisms to dynamically assign different weights to different time steps of the sequence, allowing the model to attend to more informative or salient parts of the input. It has demonstrated superior performance in various natural language processing tasks [45].
- Reinforcement Learning (RL) [46,47] involves an agent learning through interactions with an environment, receiving feedback in the form of rewards or punishments based on its actions, and learning a mapping from states to actions that maximize the expected cumulative reward over time [48].
- Deep Q-Networks (DQN) [47] combine reinforcement learning and deep learning, utilizing the deep neural network to approximate the Q-function and then learn optimal policies in complex environments. The Q-function—also known as the action-value or quality function—represents the expected cumulative reward an agent can achieve by taking a specific action in a given state and following a certain policy. In recent years, Deep RL has gained substantial attention and success in various domains, including robotics, game playing, and autonomous systems [49].
- A Generative Adversarial Network (GAN) [50] is composed of a generator network and a discriminator network, which engage in a competitive game. The generator aims to produce synthetic data samples, while the discriminator tries to discern between real and fake samples. Through iterative training in this adversarial process, GANs have exhibited remarkable capabilities in tasks such as image generation, image-to-image translation, and text generation [51,52].
- Transformers [53,54] are neural networks that use self-attention to capture relationships between words or tokens in a sequence. Self-attention involves calculating attention scores based on the relevance of each element to others, obtaining attention weights through the softmax function, and computing weighted sums using these attention weights. In transformers, the encoder computes representations for each element using self-attention, capturing dependencies and relationships, while the decoder uses this information to generate an output sequence [55].
- Bidirectional Encoder Representations from Transformers (BERT) [56] is a powerful pre-trained language model introduced by Google in 2018. BERT is trained in a

bidirectional manner, learning to predict missing words by considering both the preceding and succeeding context, resulting in a better understanding of the overall sentence or document. BERT's ability to capture contextual information and leverage pre-training has paved the way for advancements in understanding and generating human language [57].

- Autoencoders [58,59] are neural networks that learn to reconstruct their input data. They consist of an encoder network that maps input data to a compressed latent space and a decoder network that reconstructs the original data from the latent representation. They can be employed for tasks such as dimensionality reduction, anomaly detection, and generative modeling [60].
- A Stack Denoising Autoencoder (SDAE) [61] is a deep neural network composed of multiple layers of denoising autoencoders. These autoencoders are designed to reconstruct the input data from a corrupted or noisy version, enabling the model to learn robust and informative representations [62].
- A Deep Belief Network (DBN) [63,64] is a type of generative deep learning model that consists of multiple layers of stochastic unsupervised restricted Boltzmann machines (RBMs). An RBM is a two-layer neural network with binary nodes that learns representations by minimizing the energy between visible and hidden nodes [65].
- Graph Neural Networks (GNNs) [66–68] are a class of deep learning model designed to learn node representations by aggregating information from neighboring nodes in a graph, which are typically used to capture and propagate information through the graph structure, enabling effective learning and prediction tasks on graph-structured data [69].
- A Directed Acyclic Graph Neural Network (DAGNN) [70] is an architecture specifically designed for directed acyclic graphs, where the nodes represent entities or features, and edges denote dependencies or relationships. DAGNNs can effectively capture complex dependencies and facilitate learning and inference in domains with intricate relationships among variables.

### 2.3. Optimization Techniques for Machine and Deep Learning

Optimization techniques play a crucial role in machine learning and deep learning algorithms, helping to find the optimal set of parameters that minimize a loss function or maximize a performance metric with the aim of improving the model's accuracy and generalization ability. Some popular optimization techniques are detailed below [71].

- Gradient Descent is an iterative algorithm that updates the model's parameters by moving in the direction of steepest descent of the loss function.
- Stochastic Gradient Descent (SGD) is a variant of the Gradient Descent algorithm that is particularly suitable for large-scale data sets. It is widely used in deep learning, where it updates the network parameters based on a randomly selected subset of training examples, called a mini-batch.
- Adaptive Moment Estimation is an extension of gradient descent that incorporates adaptive learning rates for different parameters. It dynamically adjusts the learning rate based on the first and second moments of the gradients.
- Root Mean Square Propagation is an optimization algorithm that adapts the learning rate individually for each parameter based on the average of past squared gradients.
- Adagrad adapts the learning rate for each parameter based on their historical gradients. It places more emphasis on less frequent features by reducing the learning rate for frequently occurring features.

Researchers and practitioners often experiment with different optimization algorithms to achieve better training outcomes.

### 2.4. Ensemble Techniques for Machine and Deep Learning

Ensemble techniques for machine and deep learning approaches involve combining multiple individual models to create a more powerful and accurate predictive model.

By leveraging the strengths and diversity of different models, ensemble techniques often present improved performance and robustness when compared to using a single model [72].

Some common ensemble techniques for machine and deep learning are as follows.

- Bagging (Bootstrap Aggregating) [73] involves training multiple models independently on different subsets of the training data, typically using the same learning algorithm. The final prediction is obtained by averaging or voting the predictions of the individual models. Random Forest is an example of a popular ensemble method that utilizes bagging [74].
- AdaBoost (Adaptive Boosting) [75] sequentially trains multiple homogeneous weak models and adjusts the weights of the training examples to emphasize misclassified instances. The final prediction is a weighted combination of the predictions from the individual models, with more weight given to more accurate models [76].
- Gradient Boosting [77] is an advanced boosting methodology that incorporates the principles of gradient descent for optimization purposes. It assembles an ensemble of weak learners in a sequential manner. The primary objective during this iterative process is for each subsequent model to specifically address and minimize the residual errors—also referred to as gradients—with respect to a pre-determined loss function [78].
- XGBoost (Extreme Gradient Boosting) [79] is an optimized and highly efficient implementation of gradient boosting. It introduces regularization techniques to control model complexity and prevent over-fitting and uses a more advanced construction to provide parallel processing capabilities to accelerate training on large data sets. It also offers built-in functionality for handling missing values, feature importance analysis, and early stopping [80].
- Stacking [81,82] enhances the predictive accuracy by integrating heterogeneous weak learners. These base models are trained in parallel to provide a range of predictions, upon which a meta-model is subsequently trained, synthesizing them into a unified final output. This not only leverages the strengths of individual models, but also reduces the risk of over-fitting.

Ensemble techniques can enhance model performance by reducing over-fitting, increasing model stability, and capturing diverse aspects of the data. They are widely used in various domains and have been shown to improve performance in tasks such as classification, regression, and anomaly detection.

### *2.5. Techniques To Prevent Over-Fitting And Improve Generalization*

To prevent over-fitting and improve the generalization capability of individual or ensemble models, beside the above-mentioned ensemble methods, several other techniques can be employed [83], as detailed below.

- Cross-validation [84,85] is a widely used technique to estimate the performance of a model on unseen data. It involves partitioning the available data into multiple subsets, training the model on a subset, and evaluating its performance on the remaining subset which can guide the selection of hyperparameters and model architecture.
- Regularization methods [86,87], such as L1 and L2 regularization, add a penalty term to the loss function during training. This discourages the model from fitting the training data too closely and encourages simpler and more robust models [88].
- Dropout [89] is a technique commonly used in deep learning models. It randomly deactivates a fraction of the neurons during training, effectively creating an ensemble of smaller sub-networks. This encourages the network to learn more robust and less dependent representations, reducing over-fitting and improving generalization.
- Early stopping [90,91] involves monitoring the model's performance on a validation set during training and stopping the training process when the performance on the validation set starts to degrade. This prevents the model from over-optimizing the

training data and helps to find an optimal point that balances training accuracy and generalization.

- Data augmentation [92] involves artificially increasing the size of the training set by applying various transformations to the existing data. This introduces diversity into the training data, reducing the risk of over-fitting and helping the model to better generalize to unseen examples.

These techniques, either used individually or in combination, can help to mitigate over-fitting and improve the generalization ability of machine learning and deep learning models, leading to better performance on unseen data.

## 3. The Main Research Topics Of Machine And Deep Learning In E-Commerce

According to the main focus areas and applications of machine learning and deep learning approaches in the field of e-commerce, the 158 retrieved papers in the e-commerce literature could be categorized into the following categories: Sentiment analysis, recommendation systems, fake review detection, fraud detection, customer churn prediction, customer purchase behavior prediction, sales prediction, product classification and image recognition, and other directions. The distribution of the papers across these different categories is shown in Figure 3.

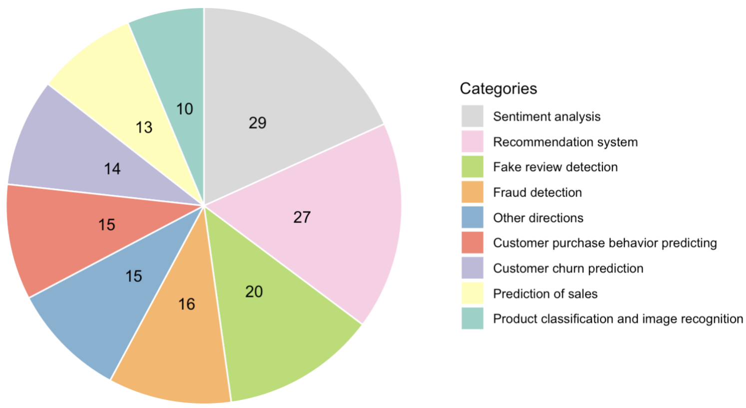


**Figure 3.** Pie chart of categories for the 158 retrieved papers.

### 3.1. *Sentiment Analysis*

Sentiment analysis—also known as opinion mining—is the process of determining the sentiment or emotional tone (e.g., positive, negative, or neutral) behind a piece of text, such as a review, comment, or social media post. Machine learning and deep learning techniques have been widely applied in sentiment analysis to automate the process of sentiment classification. A notable resource that provides an in-depth exploration of these methods for sentiment analysis is the book referenced in [93,94].

Deep learning has demonstrated its superiority over traditional machine learning algorithms in sentiment analysis tasks. One study [95] has explored various machine learning techniques, including linear SVM, random forest, multi-nomial naïve Bayes, Bernoulli naïve Bayes, and logistic regression, for the analysis of product ratings and text reviews in the context of Amazon's data set. The results revealed that the random forest classifier achieved an impressive accuracy of 91.90%, while the combination of RNN and LSTM models achieved a maximum accuracy of 97.52%, highlighting the superior performance of deep learning approaches in sentiment classification tasks. Another example is the continuous naïve Bayes learning framework [96], which achieved high accuracy ratios, ranging from 78.31% to 85.58% for domain-specific sentiment classification and from 70.97% to 75.68% for cross-domain sentiment classification, using Amazon's product and movie review sentiment data sets. An SVM-based prediction model applied to reviews of beauty products and musical instruments on Amazon achieved a precise accuracy of 87.88%, a recall rate of 99.98%, and an F1 score of 93.54% [97]. Research conducted on hotel reviews [98] has demonstrated that LSTM and GRU models achieved accuracy rates of 86% and 84%, respectively, out-performing decision tree, random forest, naïve Bayes, and SVM models (with accuracy of 71%, 71%, 75%, and 82%, respectively). Another study [99] has demonstrated the varying accuracy of different classifiers for sentiment analysis: Naïve Bayes classifiers achieved an accuracy of 68%, SVM classifiers achieved 71%, and deep learning models out-performed the others, achieving the highest accuracy of 75%. A 3-classification sentiment analysis on an Amazon data set [100] showed that RBM and DBN achieved an accuracy of 90%, while SVM only achieved an accuracy of 64%. These results indicate that deep learning methods are significantly better than classical machine learning methods for sentiment analysis. A study [101] has shown that, among the evaluated classifiers (including random forest, gradient boosting, XGBoost, multinomial naïve Bayes, and decision tree), the Bi-LSTM model consistently out-performed the others in terms of categorizing customer sentiment on product reviews in the Amazon data set. Furthermore, the research undertaken in [102] provided evidence that, in the context of aspect-based sentiment analysis, random forest out-performed other conventional machine learning techniques while, among the deep learning approaches examined, the multi-channel CNN model exhibited an accuracy of 96.23%, the highest among them, indicating its exceptional capability in accurately classifying sentiment across various aspects. The findings presented in [103] illustrated that the combined utilization of deep learning models out-performed a wider range of machine learning algorithms when applied to the data set gathered from Amazon.

Regarding the comparison of deep learning models, the findings of the study [104] substantiated that CNN out-performed RNNs when evaluated on the IMDB Movie Review (50k) and Amazon Product Review (10k) data sets, while ensemble models combining CNN and RNN (LSTM or BiLSTM) exhibited significant performance improvements compared to CNN alone. A study using LSTM and CNN-LSTM algorithms indicated that they achieved an accuracy of 94% and 91%, respectively, when testing reviews of cameras, laptops, mobile phones, tablets, televisions, and video surveillance products from Amazon [105]. Another study [106] has emphasized that the combination of CNN and LSTM out-performed their individual performances in sentiment analysis, with CNN excelling in spatial feature extraction and LSTM demonstrating effective data classification. The research findings in [107] indicated that the ensemble of CNN and attention-based BiGRU out-performed other models, including CNN, GRU, attention-based CNN, and attention-based GRU. These conclusions were drawn from the analysis of real book reviews obtained from dangdang.com. Another hybrid model consisting of Bert, BiGRU, and Softmax function models presented superior performance when compared to RNN, BiGRU, and Bert-BiLSTM models, achieving an accuracy of over 95.5% and lower loss in the analysis of e-commerce reviews [108]. One study showed that a TCN attention model out-performed other models, such as LSTM, Bi-directional LSTM, LSTM Attention, and TCN, for sentiment analysis of Chinese commodity comment phrases, achieving higher accuracy and F1-score as well as presenting superior ability to recognize noise and avoid over-fitting [109]. A hybrid model of Bidirectional CNN and spotted hyena-optimized LSTM out-performed LSTM, CNN, BIGRU, and other attention classifiers, as reported in [110]. The survey in [111] highlighted the limitations of machine learning-based sentiment analysis, due to heavy sentiment dictionaries and a large amount of labeled data, while deep learning models provide effective solutions, along with a comprehensive comparison of different deep learning approaches with regard to binary, multi-modal, and aspect-level sentiment analysis.

Some deep learning models and the used experimental data sets for sentiment analysis from 2018 to 2023 are presented in Table 4.

**Table 4.** Deep learning models proposed for sentiment analysis.

| References | Models | Data Set |
|---|---|---|
| [98] | LSTM, GRU (compared with Naïve Bayes, Decision Tree, Random Forest, and SVM) | Indian Hotel booking data from booking.com |
| [104] | Comparative study of LSTM, Bi-LSTM, GRU-CNN, CNN-RNN, CNN-LSTM, and CNN-BiLSTM | IMDB Movie Review, Amazon Product Review |
| [105] | LSTM, CNN-LSTM | Reviews of cameras, laptops, mobile phones, tablets, televisions, and video surveillance products from the Amazon website |
| [106] | LSTM, CNN, CNN-LSTM | Vietnamese VLSP 2018 data set |
| [107] | Combination sentiment lexicon, CNN, and attention-based BiGRU | Book evaluation of dangdang.com |
| [108] | RNN, Bert, BiGRU, Bert-BiLSTM, Softmax function | COAE2014-task4, ChnSentiCorp-Htl-ba-6000, Reviews about mobile phone products from Sunning and Taobao |
| [109] | LSTM, Bi-directional LSTM, LSTM Attention, TCN, TCN Attention model | 10,679 positive comments and 10,428 negative comments collected from Chinese e-commerce platforms |
| [110] | Hybrid model of spotted hyena optimized LSTM and BiCNN (compared with CNN, BIGRU, CNN-LSTM, and Attention LSTM), | Amazon Product Reviews (Text Data), Twitter Emoji data sets (Emojis), Shopping Customer data (Text + Emojis) |
| [112] | LSTM encoder–decoder (compared with LSTM, Bi-LSTM, and attention-based LSTM) | Reviews on electronics products from the Amazon website |
| [113] | Combination of SVM, logistic regression, naïve Bayes model with LSTM, RNN | Comments of smartphones on Amazon and Flipkart, Kaggle |
| [114] | CNN attention, RNN attention (compared with CNN and RNN) | 8 cross-border e-commerce APP reviews from APP Store |
| [115] | Combination of LSTM and Sigmoid kernel | Amazon-based customer reviews |

### 3.2. Recommendation System

A recommendation system in e-commerce is a technology that suggests personalized product recommendations to users based on their preferences, past behavior, and other relevant data, with the aim of enhancing the user experience, increasing sales, and improving customer satisfaction. Machine learning and deep learning techniques can analyze large amounts of user data, such as browsing history, purchase patterns, and demographic information, in order to generate personalized recommendations, which play a crucial role in building recommendation systems in the e-commerce field.

In the works selected from 2018–2023, several key perspectives on recommendation system studies were observed, which are summarized as follows:

Deep learning models have demonstrated improved accuracy, personalization, and performance in recommendation systems, when compared to traditional machine learning approaches. Furthermore, hybrid deep learning models have the potential to enhance the recommendation quality even further. For instance, in a study on personalized size and fit recommendations [116], the proposed methodology surpassed traditional methods, including a Bayesian approach, resulting in enhanced customer satisfaction and reduced returns. Another study on session-based recommendation [117] has shown that advanced neural-based models and nearest neighbor algorithms out-performed baseline techniques in various scenarios, while combining hybrid models with baseline algorithms yielded promising results. Additionally, seeking to address the challenge of sparse rating data in recommendation systems [118], a hybrid deep learning model integrating SDAE, LSTM, and Probabilistic Matrix Factorization out-performed previous deep learning works, leading to more accurate rating predictions.

Matrix factorization plays a crucial role in enhancing deep learning-based recommendation systems, as evidenced by the findings from the following papers. In a study on predicting size and fit in fashion e-commerce [116], a deep learning system incorporated matrix factorization techniques to optimize a global set of parameters and learn population-level abstractions of size and fit information from customer-article interactions. This approach out-performed existing methods on publicly available and proprietary data sets. Another paper [119] addressed data sparsity and user cold-start issues in cross-domain recommendation systems and utilized matrix factorization and deep learning to learn user features, establish trust relationships, and generate personalized recommendations. The model was evaluated on a real data set from an e-commerce retail service, showing significant improvements compared to baseline models. Additionally, in the study of [120], deep neural networks were employed to extract deep features from customer reviews, while matrix factorization was used for collaborative filtering. The proposed methodology out-performed baseline methods in terms of rating prediction accuracy on the Amazon data set. Another research focused on handling sparse rating matrices in e-commerce recommendation systems [118], and proposed a hybrid deep learning approach combining LSTM, SDAE, and latent factor models, which led to significant improvements over previous works.

Real-time feedback, proper display, and session-based recommendations have emerged as important research areas in the field of recommendation systems. One approach is to develop a deep learning-based online system that captures the real-time purchasing intentions of users, helping to improve the recommendation process [121]. Another method involves utilizing deep reinforcement learning techniques to optimize item display on a 2D page, leveraging real-time user feedback to enhance page-wise recommendations [122]. Additionally, hybrid frameworks that combine deep learning with baseline algorithms have been proposed to address challenges associated with long sessions and limited training data [117,123].

Some machine learning and deep learning model-based works on recommendation systems from 2018 to 2023 are presented in Table 5.

**Table 5.** Deep learning models proposed for recommendation systems.

| References | Models | System Types |
|---|---|---|
| [116] | Matrix factorization, deep MLP, SGD | Personalized size and fit recommendation |
| [117] | Hybrid model of Matrix Factorization, RNN-GRU, attention mechanism, and GAN | Session-based recommendation |
| [118] | Hybrid model of SDAE, LSTM, and Probabilistic Matrix Factorization | E-commerce recommendation system from sparse rating matrix |
| [119] | Generalized matrix factorization, deep MLP | Cross-domain recommendations |
| [120] | Latent Dirichlet Allocation, deep neural network | E-commerce recommendation based on customer reviews analysis |
| [121] | Deep neural network | Production online recommendation |
| [122] | RNN, GRU | Page-wise recommendation and interaction, 2D page real time feedback |
| [123] | Combination of CNN, GAN, RL, and Deep Q-Network | Session-based interactive recommendation |
| [124] | Time window-based RNN | Product sequence recommendation |
| [125] | Comparative analysis of CNN, RNN, LSTM, and GRU | Product recommendation for online shopping |
| [126] | CNN | Image retrieval and visual recommendation |
| [127] | Comparative review of KNN, SVM, random forest, CNN, LSTM-RNN, and GNN | Search engine recommendation review |
| [128] | Pairwise deep RL | Recommendations with negative feedback |
| [129] | CNN, LSTM | Cross-border Niche product recommendation |
| [130] | RL | Product recommendation in online advertising |
| [131] | CNN | Fashion collocation recommendation model |
| [132] | Deep neural network, collaborative filtering | Recommendation engine of social media websites |
| [133] | CNN, RNN | Cold start and data sparsity of recommendation system |
| [134] | Transformer | Sequential signals underlying user behavior sequences for recommendation in Alibaba |
| [135] | GAN | Content-based recommendation system |
| [136] | CNN, autoencoders | E-commerce fashion market recommendation |
| [137] | CNN, RNN | Shopping basket recommendation |

### 3.3. Fake Review Detection

Fake review detection refers to the process of identifying and distinguishing fraudulent or deceptive reviews from genuine ones. With the rise of online platforms and e-commerce websites, fake reviews have become a prevalent issue, as they can mislead consumers and impact their purchasing decisions. Machine Learning (ML) and Deep Learning (DL) techniques have been widely applied in the field of fake review detection, in order to automate the process and improve accuracy.

In the works selected from 2018–2023, the key perspectives on fake review detection studies can be summarized as follows:

Different machine learning methods have their own advantages in various environments. In the field of sentiment analysis-based detection models, SVM has been found to out-perform KNN and decision tree, achieving the highest precision both with and without the use of stop words (with accuracies of 81.75% and 81.35%, respectively) in the empirical study of Elmurngi et al. [138]. Similarly, in a separate study by Elmurngi et al. [139], logistic regression was identified as the best classifier in a comparison with NB, decision tree, and SVM, based on Amazon sample data sets related to clothing, shoes, and jewelry. Logistic regression achieved the highest accuracy of 81.61%, out-performing NB (80.61%), decision tree (81.45%), and SVM (80.90%). In the domain of detection methods based on semantic and behavioral features, a study by Wang et al. [140] demonstrated that logistic regression out-performed KNN, naïve Bayes, and SVM classifiers. The logistic regression classifier achieved an impressive accuracy of 97.2%. These findings highlight the variability in performance among different machine learning methods, emphasizing the importance of selecting the most appropriate approach based on the specific task and data set at hand.

Deep learning methods have consistently demonstrated superior performance, compared to traditional machine learning methods, in the field of fake review detection. Researchers have shown that deep learning models such as CNN, Bi-LSTM, and CNN-LSTM generally out-perform traditional machine learning models, including NB, SVM, KNN, decision tree, and logistic regression methods, when it comes to detecting fake reviews [141]. In an emotional expression and extreme rating-based detection model, the CNN-LSTM model achieved a higher accuracy (of 78.4%) compared to SVM and a multi-layer ANN [142]. In a rule-based fake review identification system, the CNN-BiLSTM model provided better performance than random forest, in terms of accuracy [143]. Moreover, in an opinion spam detection work, experimental results demonstrated a significant improvement in accuracy when using CNN, compared to traditional machine learning approaches such as SVM, logistic regression, and naïve Bayes [144]. Similarly, in the context of spam review detection, deep learning models like CNN and LSTM consistently out-performed naïve Bayes, KNN, and SVM [145]. These findings collectively support the notion that deep learning methods—particularly CNN and LSTM—exhibit superior performance when compared to traditional machine learning methods in the detection of fake reviews and spam reviews, as well as identifying opinions.

The word2vec and GloVe (Global Vectors for Word Representation) methods are two popular techniques for word embeddings associated with machine learning and deep learning for fake review detection and rating [141,144,146,147].

Some machine learning and deep learning model-based works on fake review detection from 2018 to 2023 are presented in Table 6.

**Table 6.** Machine and deep learning-based works on fake review detection.

| References | Models | Auxiliary Techniques or Application Fields |
| --- | --- | --- |
| [138] | Comparative analysis of Naïve Bayes, SVM, KNN, and decision tree | Weka Text Classification, film reviews data set |
| [139] | Comparison analysis of Naïve Bayes, SVM, logistic regression, and decision tree | Tokenization and removing stop words, Amazon data sets |
| [140] | Comparison analysis of KNN, SVM, naïve Bayes, and logistic regression | Yelp data set |
| [141] | CNN, Bi-LSTM (compared with Naïve Bayes, SVM, KNN, and decision tree) | GloVe embedding method, bag of words model, Amazon e-commerce reviews |
| [142] | Hybrid model of SVM, MLP, and CNN-LSTM | Emotional expressions, extreme rating |
| [143] | CNN-BiLSTM (compared with Random forest) | Amazon platform |
| [144] | CNN (compared with SVM, logistic regression, and naïve Bayes) | GloVe word embedding, Ott data set |
| [145] | Comparative analysis of CNN, MLP, LSTM, naïve Bayes, KNN, and SVM | Yelp Database |
| [146] | Comparative Study of CNN, Bi-LSTM, CNN-Bi-LSTM, logistic regression, random forest, naïve Bayes, and SVM | Word2Vec, FastText, and GloVe embeddings, E-commerce website called DarazBD |
| [147] | Bi-GRU | Amazon, Yelp |
| [148] | CNN + LSTM (compared with MLP, naïve Bayes, and SVM) | Ott and Yelp data sets |
| [149] | RNN+CNN (compared with BiLSTM and SVM) | Amazon product reviews data set |
| [150] | BiLSTM+Attention, CNN-BiLST (compared with CNN, BiLSTM, logistic regression, naïve Bayes, and BERT) | Real-world data sets from http://Yelp.com (accessed on 2 November 2022) |
| [151] | Hybrid model of CNN, RNN, and attention mechanism | Cross-domain spam detection, Hotel, restaurant, and doctor reviews |
| [152] | Comparison analysis of MLP, CNN, LSTM, Naïve Bayes, KNN, and SVM | Ott Data set, Yelp Data set |

### *3.4. Fraud Detection*

Fraud detection in e-commerce refers to the process of identifying and preventing fraudulent activities within online retail transactions. Machine learning and deep learning techniques play a crucial role in combating e-commerce fraud by analyzing large amounts of data, detecting patterns, and identifying suspicious behaviors.

In the selected works from 2018–2023 focused on fraud detection, it was observed that deep learning methods did not consistently out-perform machine learning methods, due to their requirement for large quantities of data to fully learn features. The performance of machine learning methods varied across individual business cases, with factors such as the number of features, transactions, and feature correlations influencing the model's effectiveness [153]. For example, in a comparative analysis of Credit Card Fraud Detection, a random forest model showed slightly better accuracy than a deep neural network [154]. However, in another study on credit card fraud detection, BiLSTM and BiGRU out-performed naïve Bayes, Adaboost, random forest, decision tree, and logistic regression [155].

Another noteworthy concern is the vulnerability of deep fraud detectors to slight perturbations in input transactions. Studies have shown that these deployed detectors can be highly susceptible to attacks, with small perturbations significantly reducing the average precision from nearly 90% to as low as 20%. However, models trained with an adversarial training process exhibit remarkable robustness against such attacks, ensuring better performance even in the presence of adversarial perturbations.

Some machine learning and deep learning model-based works on fraud detection from 2018 to 2023 are presented in Table 7.

**Table 7.** Machine and deep learning-based works on fraud detection.

| References | Models | Research Topic |
|---|---|---|
| [153] | Comparison analysis of Decision tree, random forest, SVM, logistic regression, XGBoost, CNN, LSTM, RNN, GAN, and RBM | Credit card fraud detection method comparison |
| [154] | Comparative analysis Random forest and deep neural network | Credit card fraud detection |
| [155] | Hybrid model of BiLSTM and BiGRU (compared with naïve Bayes, Adaboost, random forest, decision tree, and logistic regression) | Credit card fraud detection |
| [156] | Deep network, RL | Vulnerability of deep fraud detector |
| [157] | Hybrid model of Markov Decision Process and RL | Impression allocation for combating Fraud in e-commerce |
| [158] | RL | Order fraud evaluation |
| [159] | Machine learning methods | Comprehensive survey on fraud detection |
| [160] | Comparative analysis of Random forest, SVM, KNN, KNN-SVM-CNN, and RBM | Fraudulent transaction tracing |
| [161] | Hybrid model of Encoders, SVM and CNN | Financial fraud detection |
| [162] | Comparative analysis of Random forest and Adaboost | Credit card fraud detection |

### *3.5. Customer Churn Prediction*

Customer churn prediction in e-commerce refers to the process of forecasting which customers are likely to discontinue their engagement or stop making purchases. By analyzing historical consumer data, transactions, and browsing patterns, machine learning and deep learning models can be employed to predict potential churners, improve overall customer satisfaction, and optimize the profitability and success of e-commerce businesses [163].

In the selected works from 2018–2023 focused on customer churn prediction, it was observed that the imbalance in churned and unchurned customers can bias machine and deep learning models, thus hindering generalization and increasing misclassification errors [164]. To address this, techniques such as re-sampling, class weighting, data augmentation, and ensemble methods can be used to mitigate the effect of imbalanced class data [165–167]. Additionally, the accuracy of ML and DL methods for churn prediction can be influenced by the presence of multiple types of data, such as numerical data on orders, textual after-purchase reviews, and socio-geo-demographic data [168]. Careful feature engineering and appropriate model architectures are vital for leveraging the unique information in different data types, thereby improving the accuracy and predictive capability of the churn prediction system [168].

Some machine learning and deep learning model-based works on customer churn prediction from 2018 to 2023 are detailed in Table 8.

**Table 8.** Machine and deep learning-based works on customer churn prediction.

| References | Models | Research Topic or Application Field |
|---|---|---|
| [165] | AdaBoost, deep network | Imbalanced data processing |
| [166] | RNN | Imbalanced classes of real e-commerce data |
| [167] | Hybrid model of PCA, AdaBoost, and decision tree | High-dimensional and unbalanced data |
| [168] | Deep neural network | Telco data set, Churn factor analysis |
| [169] | Comparative analysis of PCA, SVM, naïve Bayes, random forest, and deep network | Brazilian e-commerce data set |
| [170] | CNN | Telco Customer, Distributed model |
| [171] | Hybrid model of CNN, decision tree, and grid search optimization | Diagnosis of employee churn |
| [172] | Comparative analysis of Naïve Bayes, SVM, decision tree, random forest, and logistic regression | IBM Watson Analytics HR data, Employee attrition prediction |

### 3.6. Customer Purchase Behavior Prediction

Predicting purchase behavior is a vital aspect of marketing and business strategy in e-commerce. By analyzing historical purchase data, customer demographics, browsing behaviors, and other relevant factors, businesses can anticipate the future purchasing decisions of individual customers or customer segments. This enables personalized marketing, targeted promotions, and customer retention strategies, ultimately leading to enhanced customer satisfaction and increased sales. Some of the studies in the realm of e-commerce from 2018 to 2023 focused on predicting and understanding customer behavior to optimize marketing strategies and enhance user experience. One such approach employed RNNs, considering the sequential nature of clickstream data, to obtain accurate predictions and cost-saving marketing techniques [173]. Another study has explored the use of DBNs and SDAE against traditional machine learning methods in effectively handling severe class imbalances [174]. Additionally, the time of the next purchase is predicted using the deep network, out-performing conventional methods such as random forest and SVM, showcasing the potential of deep neural networks for time-series forecasting in the e-commerce field [175–178]. Moreover, the identification of purchase intentions from implicit queries has been investigated, employing a model combining Word2Vec and LSTM to enhance feature extraction and improve the accuracy of purchase intention identification [179]. These studies collectively highlight the importance of employing advanced deep learning techniques to predict customer behaviors and improve business performance in the dynamic landscape of e-commerce.

### *3.7. Prediction of Sales*

Effective sales prediction in e-commerce empowers companies to stay competitive, anticipate market changes, and provide a seamless shopping experience to customers. Machine learning and deep learning techniques play a vital role in this domain, utilizing historical sales data, consumer behaviors, product attributes, and various other factors to predict future sales trends. Several works on sales prediction for online products using deep learning have emerged, showcasing the adaptability and effectiveness of CNNs in various domains. One study has compared full connection models with the CNN training results, highlighting the accuracy and generalization capabilities of CNNs [180]. Deep learning approaches, including CNNs, have been shown to out-perform traditional machine learning methods in terms of sales forecasting accuracy [181]. Another hybrid model, combining BiLSTM and CNN, considered product attributes and sentiment analysis from customer comments to achieve accurate sales predictions [182]. Additionally, a sales forecasting method used a DAGNN dynamically to predict daily sales revenue for different product categories, offering scalability and generalization [183]. Furthermore, a deep neural framework has demonstrated significant performance gains over traditional and other deep learning models in forecasting E-commerce sales, considering promotion campaigns and competing relationships between products [184]. These studies collectively contribute valuable insights into sales prediction and demonstrate the potential of deep learning techniques to revolutionize the e-commerce industry.

### *3.8. Product Classification and Image Recognition*

Efficient and precise product classification in the realm of e-commerce poses a significant challenge, due to the constant influx of new products and the ever-changing nature of categories. To address this challenge, a decision-level fusion approach has been proposed, combining text and image neural network classifiers to reduce human editorial effort and improve classification accuracy on a large-scale product data set from Walmart.com [185]. Another innovative solution involves machine translation-based e-commerce product categorization, employing a sequence-to-sequence neural network and the transformer model to out-perform traditional algorithms [186]. Additionally, a random forest model utilizing product titles achieved superior accuracy, compared to SVM and CNN, for the classification of e-commerce products [187]. On the other hand, image recognition techniques have been extensively explored in the field of fashion and clothing product classification, employing various machine learning and deep learning models such as CNN, CNN-RNN, transfer learning, SVM-CNN, and LSTM to accurately categorize products [188–196]. These studies emphasize the vital role of advanced machine learning and deep learning techniques in enhancing product classification, image recognition, and categorization in the e-commerce industry.

### *3.9. Other Directions*

In the studies retrieved from 2018 to 2023, machine learning and deep learning techniques have been utilized in other lines of e-commerce research, including financial risk prediction, capacity allocation, entrepreneur identification, tax evasion detection, marketing promotion, task scheduling, automatic pricing, and entity recognition. These studies are briefly detailed in Table 9.

**Table 9.** Machine and deep learning-based works in other directions.

| References | Models | Direction or Application Fields |
|---|---|---|
| [197,198] | Machine learning optimization | Last-mile delivery, third-party logistics, and bin packing problem |
| [199] | Comparative analysis of Random Forest, XGBoost, Logistic Regression, and Neural Network | Detection and prediction of company short-, middle-, and long-term defaults and bankruptcy |
| [200] | Deep neural network | Financial early warning model |
| [201] | LSTM | Financial risk prediction for Chinese e-commerce enterprises from 2012 to 2022 |
| [202] | CNN-LSTM | Service capacity allocation for cross-border e-commerce |
| [203] | Deep neural network | Identification of entrepreneurs in rural e-commerce |
| [204] | Deep neural network | Social e-commerce tax evasion detection |
| [205] | Hybrid model of AdaBoost and Deep neural network | E-commerce industry marketing promotion |
| [206] | CNN | Task scheduling based on deadline and cost |
| [207] | CNN | Cross-border e-commerce platform for commodity automatic pricing |
| [208] | LSTM | Named Entity Recognition |

## 4. The Main Challenges and Trends For Machine And Deep Learning In E-Commerce

Based on the 158 selected papers, we provide six main challenges and trends for machine learning and deep learning in e-commerce research, as follows:

- Imbalanced data pose a significant challenge for both machine learning and deep learning-based classification tasks in e-commerce. This issue is prevalent in fields such as fraud detection, fake review detection, customer churn prediction, and re-purchase behavior classification [164–166,174], where one class significantly outweighs the others, leading to biased models with poor performance on the minority class. To address this issue, various methods can be applied, including re-sampling techniques, weighted training, and transfer learning, helping to enhance model performance and achieve more accurate predictions in e-commerce applications.
- Preventing over-fitting and achieving robust generalization is another challenge for machine learning and deep learning in e-commerce [109,180,183]. Ensembling techniques, such as bagging and boosting, combine multiple models to improve the overall performance and reduce over-fitting. Hybrid models integrate different types of machine learning and deep learning algorithms, leveraging their respective strengths to enhance the overall generalization ability. Regularization techniques, data augmentation, dropout, cross-validation, transfer learning, and early stopping also collectively contribute to building more reliable and accurate models for e-commerce tasks.
- Multi-modal learning poses significant challenges for machine learning and deep learning approaches in e-commerce [185,209]. Integrating data from diverse sources such as text, images, and audio requires careful alignment and pre-processing: feature extraction becomes complex and resource-intensive, labeling and annotating multi-modal data is time-consuming, and the development of fusion strategies to combine modalities for accurate prediction becomes challenging. Despite these chal-

lenges, multi-modal learning has great promise for enhancing e-commerce applications such as product classification [185], recommendation [210], sentiment analysis [211], and customer behavior prediction [212].

- Model interpretability poses a significant challenge for both machine learning and deep learning approaches in the context of e-commerce [6,213]. Due to the complexity of deep learning architectures, these models are often considered "black boxes", making it difficult to understand the reasoning behind their decisions. In e-commerce applications, the ability to interpret why a model makes a specific recommendation or classification is crucial for building trust with users. Interpretability techniques, such as feature visualization [214], attention mechanisms [215], and gradient-based methods [216], are being explored to shed light on the inner workings of machine and deep learning models, enabling better transparency and accountability in the e-commerce decision-making process.
- Personalization is a prominent research area in e-commerce, aiming to enhance the user experience on various platforms [217,218]. AI-powered customer services employing machine learning and deep learning-driven chatbots and virtual assistants not only provide support, but can also help to predict customer needs and deliver tailored assistance and recommendations. Real-time inference capabilities are essential for e-commerce platforms, offering instantaneous recommendations and predictions to users. Reinforcement learning holds great potential in the realm of personalized marketing, allowing for the tailoring of promotions and advertisements to align with individual customer preferences [158]. Additionally, transfer learning stands out as a valuable strategy for refining pre-trained models, thereby enhancing their performance in specialized e-commerce tasks and mitigating the need for extensive data collection and training efforts [194]. These research-driven trends signify the growing importance of personalization and customer-centric strategies in the dynamic e-commerce landscape.
- Machine learning and deep learning-enabled chatbots and virtual assistants are, indeed, emerging as a new trend in e-commerce [219,220]. These technologies harness the power of natural language processing (NLP) and conversational AI to deliver efficient and personalized customer support [219]. AI-driven chatbots analyze customer queries and interactions, providing real-time assistance and recommendations, thereby enhancing the overall shopping experience [220]. The integration of machine learning and deep learning with chatbots can facilitate continuous learning and adaptation to evolving user behaviors, enhancing their effectiveness in addressing customer needs. As e-commerce platforms endeavor to elevate customer engagement and streamline support processes, the adoption of machine and deep learning for chatbots and virtual assistants is poised to gain traction within the e-commerce industry.

## 5. Conclusions

The multi-faceted domains within e-commerce, including review analysis, product classification, recommendation systems, and customer retention, offer abundant opportunities for the exploration of machine learning and deep learning techniques. Nonetheless, they present various challenges such as imbalanced data, over-fitting, generalization, and interpretability. In the dynamic e-commerce landscape, the supremacy of either machine learning or deep learning across all tasks remains unclear. Given the varying characteristics of data, tailored approaches are vital for optimal classification and prediction. Conducting methodical experiments and comparisons is, thus, essential for determination of the most suitable approach for specific e-commerce tasks, emphasizing the fundamental nature of e-commerce research. Meanwhile, conducting regular independent ethical audits to identify and rectify potential biases and privacy concerns in e-commerce machine and deep learning systems can ensure ongoing compliance and safeguard user trust and data integrity.

Indeed, considering the symbiotic relationship between learning techniques such as machine learning and deep learning, their application in e-commerce is poised to foster continuous growth and mutual reinforcement. Researchers and practitioners are motivated to explore cutting-edge methods, experiment with different models, and adapt existing techniques to cater to the ever-changing demands of the industry. As the digital marketplace continues to expand and evolve, the synergy between learning techniques and e-commerce applications will undoubtedly remain a driving force, fostering innovation and propelling the industry forward into new frontiers of growth and success.

**Author Contributions:** Conceptualization, X.Z. and J.W.; methodology, F.G., L.P. and J.W.; formal analysis, G.B. and J.W.; resources, X.Z. and T.C.; writing—original draft preparation, X.Z., L.P. and J.W.; writing—review and editing, F.G., T.C., G.B. and J.W.; visualization, F.G., L.P. and J.W.; supervision, G.B. and J.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| AI | Artificial intelligence |
| SVM | Support vector machine |
| PCA | Principal component analysis |
| KNN | K-nearest neighbor |
| ANN | Artificial neural network |
| CNN | Convolutional Neural Network |
| VGG | Visual Geometry Group Network |
| TCN | Temporal Convolutional Network |
| RNN | Recurrent Neural Network |
| LSTM | Long Short-Term Memory |
| BiLSTM | Bidirectional Long Short-Term Memory |
| GRU | Gated Recurrent Unit |
| RL | Reinforcement Learning |
| DQN | Deep Q-Network |
| GAN | Generative Adversarial Network |
| BERT | Bidirectional Encoder Representations from Transformers |
| SDAE | Stack Denoising Autoencoder |
| DBN | Deep Belief Network |
| GNN | Graph Neural Network |
| DAGNN | Directed Acyclic Graph Neural Network |
| SGD | Stochastic Gradient Descen |
| AdaBoost | Adaptive Boosting |
| XGBoost | Extreme Gradient Boosting |

## References


1. Sarker, I.H. Machine learning: Algorithms, real-world applications and research directions. *SN Comput. Sci.* **2021**, *2*, 160. [CrossRef] [PubMed]
2. Burkart, N.; Huber, M.F. A survey on the explainability of supervised machine learning. *J. Artif. Intell. Res.* **2021**, *70*, 245–317. [CrossRef]
3. Bi, Q.; Goodman, K.E.; Kaminsky, J.; Lessler, J. What is machine learning? A primer for the epidemiologist. *Am. J. Epidemiol.* **2019**, *188*, 2222–2239. [CrossRef] [PubMed]
4. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [CrossRef]
5. Dong, S.; Wang, P.; Abbas, K. A survey on deep learning and its applications. *Comput. Sci. Rev.* **2021**, *40*, 100379. [CrossRef]
6. Sarker, I.H. Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput. Sci.* **2021**, *2*, 420. [CrossRef] [PubMed]
7. Laudon, K.C.; Traver, C.G. *E-Commerce*; Pearson Boston: Boston, MA, USA, 2013.

8. Nosratabadi, S.; Mosavi, A.; Duan, P.; Ghamisi, P.; Filip, F.; Band, S.S.; Reuter, U.; Gama, J.; Gandomi, A.H. Data science in economics: Comprehensive review of advanced machine learning and deep learning methods. *Mathematics* **2020**, *8*, 1799. [CrossRef]
9. Song, X.; Yang, S.; Huang, Z.; Huang, T. The application of artificial intelligence in electronic commerce. *Proc. J. Phys. Conf. Ser.* **2019**, *1302*, 032030. [CrossRef]
10. Di Corso, E.; Proto, S.; Vacchetti, B.; Bethaz, P.; Cerquitelli, T. Simplifying text mining activities: Scalable and self-tuning methodology for topic detection and characterization. *Appl. Sci.* **2022**, *12*, 5125. [CrossRef]
11. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
12. Noble, W.S. What is a support vector machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [CrossRef]
13. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Cart. In *Classification and Regression Trees*; Wadsworth Publishing Group: Belmont, CA, USA, 1984.
14. Charbuty, B.; Abdulazeez, A. Classification based on decision tree algorithm for machine learning. *J. Appl. Sci. Technol. Trends* **2021**, *2*, 20–28. [CrossRef]
15. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; Volume 1, pp. 278–282.
16. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
17. Speiser, J.L.; Miller, M.E.; Tooze, J.; Ip, E. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst. Appl.* **2019**, *134*, 93–101. [CrossRef]
18. Hand, D.J.; Yu, K. Idiot's Bayes—Not so stupid after all? *Int. Stat. Rev.* **2001**, *69*, 385–398.
19. Chen, S.; Webb, G.I.; Liu, L.; Ma, X. A novel selective naïve Bayes algorithm. *Knowl.-Based Syst.* **2020**, *192*, 105361. [CrossRef]
20. Cramer, J.S. *The Origins of Logistic Regression*; University of Amsterdam and Tinbergen Institute: Amsterdam, The Netherlands, 2002.
21. Pampel, F.C. *Logistic Regression: A Primer*; Number 132; Sage Publications: New York, NY, USA, 2020.
22. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **1933**, *24*, 417. [CrossRef]
23. Xie, X. Principal component analysis. *Wiley Interdiscip. Rev.* **2019**, *12*, 273.
24. Koren, Y.; Bell, R.; Volinsky, C. Matrix factorization techniques for recommender systems. *Computer* **2009**, *42*, 30–37. [CrossRef]
25. Lee, D.; Seung, H.S. Algorithms for non-negative matrix factorization. *Adv. Neural Inf. Process. Syst.* **2000**, *13*.
26. Wang, Y.X.; Zhang, Y.J. Nonnegative matrix factorization: A comprehensive review. *IEEE Trans. Knowl. Data Eng.* **2012**, *25*, 1336–1353. [CrossRef]
27. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [CrossRef]
28. Isnain, A.R.; Supriyanto, J.; Kharisma, M.P. Implementation of K-Nearest Neighbor (K-NN) Algorithm For Public Sentiment Analysis of Online Learning. *Indones. J. Comput. Cybern. Syst.* **2021**, *15*, 121–130. [CrossRef]
29. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
30. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [CrossRef]
31. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Umar, A.M.; Linus, O.U.; Arshad, H.; Kazaure, A.A.; Gana, U.; Kiru, M.U. Comprehensive review of artificial neural network applications to pattern recognition. *IEEE Access* **2019**, *7*, 158820–158846. [CrossRef]
32. Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **1980**, *36*, 193–202. [CrossRef]
33. Lawrence, S.; Giles, C.L.; Tsoi, A.C.; Back, A.D. Face recognition: A convolutional neural-network approach. *IEEE Trans. Neural Netw.* **1997**, *8*, 98–113. [CrossRef]
34. Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 24–49. [CrossRef]
35. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
36. Paymode, A.S.; Malode, V.B. Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG. *Artif. Intell. Agric.* **2022**, *6*, 23–33. [CrossRef]
37. Lea, C.; Vidal, R.; Reiter, A.; Hager, G.D. Temporal convolutional networks: A unified approach to action segmentation. In Proceedings of the 2016 Computer Vision Workshops, Amsterdam, The Netherlands, 8–10. 15–16 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 47–54.
38. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [CrossRef]
39. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
40. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **2005**, *18*, 602–610. [CrossRef] [PubMed]
41. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
42. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555.

43. Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A.; Hovy, E. Hierarchical attention networks for document classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, CA, USA, 11–12 June 2016; pp. 1480–1489.
44. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv* **2014**, arXiv:1409.0473.
45. Liu, J.; Yang, Y.; Lv, S.; Wang, J.; Chen, H. Attention-based BiGRU-CNN for Chinese question classification. *J. Ambient. Intell. Humaniz. Comput.* **2019, 10**, 1–12. [CrossRef]
46. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. *J. Artif. Intell. Res.* **1996**, *4*, 237–285. [CrossRef]
47. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M. Playing atari with deep reinforcement learning. *arXiv* **2013**, arXiv:1312.5602.
48. Arulkumaran, K.; Deisenroth, M.P.; Brundage, M.; Bharath, A.A. Deep reinforcement learning: A brief survey. *IEEE Signal Process. Mag.* **2017**, *34*, 26–38. [CrossRef]
49. Huang, Y. Deep Q-networks. *Deep Reinforcement Learning: Fundamentals, Research and Applications*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 135–160.
50. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; Advances in Neural Information Processing Systems 27 (NIPS 2014), Volume 2.
51. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.
52. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [CrossRef]
53. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; Advances in Neural Information Processing Systems 30 (NIPS 2017).
54. Luong, M.T.; Pham, H.; Manning, C.D. Effective approaches to attention-based neural machine translation. *arXiv* **2015**, arXiv:1508.04025.
55. Khan, S.; Naseer, M.; Hayat, M.; Zamir, S.W.; Khan, F.S.; Shah, M. Transformers in vision: A survey. *ACM Comput. Surv.* **2022**, *54*, 1–41. [CrossRef]
56. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
57. Alaparthi, S.; Mishra, M. Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey. *arXiv* **2020**, arXiv:2007.01127.
58. Kingma, D.P.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**, arXiv:1312.6114.
59. Kramer, M.A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.* **1991**, *37*, 233–243. [CrossRef]
60. Pinaya, W.H.L.; Vieira, S.; Garcia-Dias, R.; Mechelli, A. Autoencoders. In *Machine Learning*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 193–208.
61. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.A.; Bottou, L. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *J. Mach. Learn. Res.* **2010**, *11*,3371–3408.
62. Zhang, Z.; Jiang, T.; Li, S.; Yang, Y. Automated feature learning for nonlinear process monitoring–An approach using stacked denoising autoencoder and k-nearest neighbor rule. *J. Process Control* **2018**, *64*, 49–61. [CrossRef]
63. Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef] [PubMed]
64. Hinton, G.E. Deep belief networks. *Scholarpedia* **2009**, *4*, 5947. [CrossRef]
65. Krizhevsky, A.; Hinton, G. Convolutional deep belief networks on CIFAR-10. *Unpubl. Manuscr.* **2010**, *40*, 1–9.
66. Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv* **2016**, arXiv:1609.02907.
67. Scarselli, F.; Gori, M.; Tsoi, A.C.; Hagenbuchner, M.; Monfardini, G. The graph neural network model. *IEEE Trans. Neural Netw.* **2008**, *20*, 61–80. [CrossRef] [PubMed]
68. Scarselli, F.; Yong, S.L.; Gori, M.; Hagenbuchner, M.; Tsoi, A.C.; Maggini, M. Graph neural networks for ranking web pages. In Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05), Washington, DC, USA, 19–22 September 2005; pp. 666–672.
69. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57–81. [CrossRef]
70. Shen, W.; Wu, S.; Yang, Y.; Quan, X. Directed acyclic graph network for conversational emotion recognition. *arXiv* **2021**, arXiv:2105.12907.
71. Curtis, F.E.; Scheinberg, K. Optimization methods for supervised machine learning: From linear models to deep learning. In *Leading Developments from INFORMS Communities*; INFORMS: Catonsville, MD, USA, 2017; pp. 89–114.
72. Ganaie, M.A.; Hu, M.; Malik, A.; Tanveer, M.; Suganthan, P. Ensemble deep learning: A review. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105151. [CrossRef]
73. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
74. Altman, N.; Krzywinski, M. Ensemble methods: Bagging and random forests. *Nat. Methods* **2017**, *14*, 933–935. [CrossRef]

75. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]
76. Schapire, R.E.; Singer, Y. Improved boosting algorithms using confidence-rated predictions. In Proceedings of the Eleventh Annual Conference on Computational Learning Theory, Madison, WI, USA, 24–26 July 1998; pp. 80–91.
77. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
78. Natekin, A.; Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurorobot.* **2013**, *7*, 21. [CrossRef] [PubMed]
79. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
80. Rusdah, D.A.; Murfi, H. XGBoost in handling missing values for life insurance risk prediction. *SN Appl. Sci.* **2020**, *2*, 1–10. [CrossRef]
81. Wolpert, D.H. Stacked generalization. *Neural Netw.* **1992**, *5*, 241–259. [CrossRef]
82. Breiman, L. Stacked regressions. *Mach. Learn.* **1996**, *24*, 49–64. [CrossRef]
83. Ying, X. An overview of overfitting and its solutions. *Proc. J. Phys. Conf. Ser.* **2019**, *1168*, 022022. [CrossRef]
84. Stone, M. Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc. Ser. B* **1974**, *36*, 111–133. [CrossRef]
85. Bergmeir, C.; Benítez, J.M. On the use of cross-validation for time series predictor evaluation. *Inf. Sci.* **2012**, *191*, 192–213. [CrossRef]
86. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1996**, *58*, 267–288. [CrossRef]
87. Hoerl, A.E.; Kennard, R.W. Ridge regression: Applications to nonorthogonal problems. *Technometrics* **1970**, *12*, 69–82. [CrossRef]
88. Ghojogh, B.; Crowley, M. The theory behind overfitting, cross validation, regularization, bagging, and boosting: Tutorial. *arXiv* **2019**, arXiv:1905.12787.
89. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
90. Morgan, N.; Bourlard, H. Generalization and parameter estimation in feedforward nets: Some experiments. In Proceedings of the Advances in Neural Information Processing Systems 2 (NIPS 1989), Cambridge, MA, USA, 1 January 1989.
91. Prechelt, L. Early stopping-but when? In *Neural Networks: Tricks of the Trade*; Springer: Berlin/Heidelberg, Germany, 2002; pp. 55–69.
92. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 1–48. [CrossRef]
93. Agarwal, B.; Nayak, R.; Mittal, N.; Patnaik, S. *Deep Learning-Based Approaches for Sentiment Analysis*; Springer: Berlin/Heidelberg, Germany, 2020.
94. Cao, J.; Li, J.; Yin, M.; Wang, Y. Online reviews sentiment analysis and product feature improvement with deep learning. *Trans. Asian-Low-Resour. Lang. Inf. Process.* **2023**, *22*, 1–17. [CrossRef]
95. Gope, J.C.; Tabassum, T.; Mabrur, M.M.; Yu, K.; Arifuzzaman, M. Sentiment analysis of Amazon product reviews using machine learning and deep learning models. In Proceedings of the 2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), IEEE, Gazipur, Bangladesh, 24–26 February 2022; pp. 1–6.
96. Xu, F.; Pan, Z.; Xia, R. E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework. *Inf. Process. Manag.* **2020**, *57*, 102221. [CrossRef]
97. Jabbar, J.; Urooj, I.; JunSheng, W.; Azeem, N. Real-time sentiment analysis on e-commerce application. In Proceedings of the 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), IEEE, Banff, AB, Canada, 9–11 May 2019; pp. 391–396.
98. Hossen, M.S.; Jony, A.H.; Tabassum, T.; Islam, M.T.; Rahman, M.M.; Khatun, T. Hotel review analysis for the prediction of business using deep learning approach. In Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Piscataway, NJ, USA, 21–25 March 2021; pp. 1489–1494.
99. Vaidhya, M. Sentiment Analysis of Different e-Commerce Platform Reviews Using Machine Learning Algorithm. Ph.D. Thesis, IOE Pulchowk Campus, Patan, Nepal, 2022.
100. Wang, J.; Wang, W.M.; Tian, Z.; Li, Z. Classification of multiple affective attributes of customer reviews: Using classical machine learning and deep learning. In Proceedings of the 2nd International Conference on Computer Science and Application Engineering, New York, NY, USA, 22–24 October 2018; pp. 1–5.
101. Agarwal, A.; Meena, S. A comparative study of deep learning and machine learning algorithm for sentiment analysis. In Proceedings of the 2022 3rd International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), Ghaziabad, India, 11–12 November 2022; pp. 1–7.
102. Nagelli, A.; Saleena, B. A comparative review of sentimental analysis using machine learning and deep learning approaches. *J. Inf. Knowl. Manag.* **2023**, *22*, 2350003. [CrossRef]
103. Ben, T.L.; Alla, P.C.R.; Komala, G.; Mishra, K. Detecting sentiment polarities with comparative analysis of machine learning and deep learning algorithms. In Proceedings of the 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), Mohali, India, 5–6 May 2023; pp. 186–190.
104. Tahseen, T.; Kabir, M.M.J. A comparative study of deep learning neural networks in sentiment classification from texts. In Proceedings of the Machine Learning and Autonomous Systems (ICMLAS 2021), Tamil Nadu, India, 24–25 September 2021; Springer: Berlin/Heidelberg, Germany, 2022; pp. 289–305.

105. Alzahrani, M.E.; Aldhyani, T.H.; Alsubari, S.N.; Althobaiti, M.M.; Fahad, A. Developing an intelligent system with deep learning algorithms for sentiment analysis of e-commerce product reviews. *Comput. Intell. Neurosci.* **2022**, *2022*, 3840071. [CrossRef]
106. Hung, B.T.; Semwal, V.B.; Gaud, N.; Bijalwan, V. Hybrid deep learning approach for aspect detection on reviews. In Proceedings of the 2020 Integrated Intelligence Enable Networks and Computing (IIENC 2020), Chamoli, India, 5–7 September 2020; Springer: Berlin/Heidelberg, Germany, 2021; pp. 991–999.
107. Yang, L.; Li, Y.; Wang, J.; Sherratt, R.S. Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access* **2020**, *8*, 23522–23530. [CrossRef]
108. Liu, Y.; Lu, J.; Yang, J.; Mao, F. Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax. *Math. Biosci. Eng.* **2020**, *17*, 7819–7837. [CrossRef] [PubMed]
109. Zhou, H.; Wu, G. Research on sentiment analysis of chinese e-commerce comments based on deep learning. *Proc. J. Phys. Conf. Ser.* **2019**, *1237*, 022002. [CrossRef]
110. Venkatesan, R.; Sabari, A. Deepsentimodels: A novel hybrid deep learning model for an effective analysis of ensembled sentiments in e-commerce and s-commerce platforms. *Cybern. Syst.* **2023**, *54*, 526–549. [CrossRef]
111. Lin, P.; Luo, X.; Fan, Y. A survey of sentiment analysis based on deep learning. *Int. J. Comput. Inf. Eng.* **2020**, *14*, 473–485.
112. Jain, S.; Roy, P.K. E-commerce review sentiment score prediction considering misspelled words: A deep learning approach. *Electron. Commer. Res.* **2022**, 1–25. [CrossRef]
113. Sharm, N.; Jain, T.; Narayan, S.S.; Kandakar, A.C. Sentiment analysis of Amazon smartphone reviews using machine learning & deep learning. In Proceedings of the 2022 IEEE International Conference on Data Science and Information System (ICDSIS), Hassan, India, 29–30 July 2022; pp. 1–4.
114. Wang, C.; Zhu, X.; Yan, L. Sentiment analysis for e-commerce reviews based on deep learning hybrid model. In Proceedings of the 2022 5th International Conference on Signal Processing and Machine Learning, New York, NY, USA, 4–6 August 2022; pp. 38–46.
115. Mehul, A.R.; Mahmood, S.M.; Tabassum, T.; Chakraborty, P. Sentiment polarity detection using machine learning and deep learning. In Proceedings of the 2023 International Conference on Electrical, Computer and Communication Engineering (ECCE), Chittagong, Bangladesh, 23–25 February 2023; pp. 1–5.
116. Sheikh, A.S.; Guigourès, R.; Koriagin, E.; Ho, Y.K.; Shirvany, R.; Vollgraf, R.; Bergmann, U. A deep learning system for predicting size and fit in fashion e-commerce. In Proceedings of the 13th ACM Conference on Recommender Systems, New York, NY, USA, 16–20 September 2019; pp. 110–118.
117. Maher, M.; Ngoy, P.M.; Rebriks, A.; Ozcinar, C.; Cuevas, J.; Sanagavarapu, R.; Anbarjafari, G. Comprehensive empirical evaluation of deep learning approaches for session-based recommendation in e-commerce. *Entropy* **2022**, *24*, 1575. [CrossRef]
118. Pujastuti, E.; Laksito, A.; Hardi, R.; Perwira, R.; Arfriandi, A. Handling sparse rating matrix for e-commerce recommender system using hybrid deep learning based on LSTM, SDAE and latent factor. *Int. J. Intell. Eng. Syst.* **2022**, *15*, 379–393.
119. Ahmed, A.; Saleem, K.; Khalid, O.; Rashid, U. On deep neural network for trust aware cross domain recommendations in E-commerce. *Expert Syst. Appl.* **2021**, *174*, 114757. [CrossRef]
120. Shoja, B.M.; Tabrizi, N. Customer reviews analysis with deep neural networks for e-commerce recommender systems. *IEEE Access* **2019**, *7*, 119121–119130. [CrossRef]
121. Yan, Y.; Liu, Z.; Zhao, M.; Guo, W.; Yan, W.P.; Bao, Y. A practical deep online ranking system in e-commerce recommendation. In Proceedings of the Conference on Machine Learning and Knowledge Discovery in Databases, Dublin, Ireland, 10–14 September 2018; Springer: Berlin/Heidelberg, Germany, 2019; pp. 186–201.
122. Zhao, X.; Xia, L.; Zhang, L.; Ding, Z.; Yin, D.; Tang, J. Deep reinforcement learning for page-wise recommendations. In Proceedings of the 12th ACM Conference on Recommender Systems, New York, NY, USA, 2–4 October 2018; pp. 95–103.
123. Gao, R.; Xia, H.; Li, J.; Liu, D.; Chen, S.; Chun, G. DRCGR: Deep reinforcement learning framework incorporating CNN and GAN-based for interactive recommendation. In Proceedings of the 2019 IEEE International Conference on Data Mining (ICDM), Beijing, China, 8–11 November 2019; pp. 1048–1053.
124. Zhou, L. Product advertising recommendation in e-commerce based on deep learning and distributed expression. *Electron. Commer. Res.* **2020**, *20*, 321–342. [CrossRef]
125. Almahmood, R.J.K.; Tekerek, A. Issues and solutions in deep learning-enabled recommendation systems within the e-commerce field. *Appl. Sci.* **2022**, *12*, 11256. [CrossRef]
126. Shankar, D.; Narumanchi, S.; Ananya, H.; Kompalli, P.; Chaudhury, K. Deep learning based large scale visual recommendation and search for e-commerce. *arXiv* **2017**, arXiv:1703.02344.
127. Lalitha, T.; Sreeja, P. Recommendation system based on machine learning and deep learning in varied perspectives: A systematic review. In Proceedings of the ICT Conference on Information and Communication Technology for Competitive Strategies, Jaipur, India, 2–6 June 2021; pp. 419–432.
128. Zhao, X.; Zhang, L.; Ding, Z.; Xia, L.; Tang, J.; Yin, D. Recommendations with negative feedback via pairwise deep reinforcement learning. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, New York, NY, USA, 19–23 August 2018; pp. 1040–1048.
129. Li, L. Cross-border e-commerce intelligent information recommendation system based on deep learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 6602471. [CrossRef] [PubMed]

130. Rohde, D.; Bonner, S.; Dunlop, T.; Vasile, F.; Karatzoglou, A. Recogym: A reinforcement learning environment for the problem of product recommendation in online advertising. *arXiv* **2018**, arXiv:1808.00720.
131. Wang, S.; Qiu, J. A deep neural network model for fashion collocation recommendation using side information in e-commerce. *Appl. Soft Comput.* **2021**, *110*, 107753. [CrossRef]
132. Guha, R. Improving the performance of an artificial intelligence recommendation engine with deep learning neural nets. In Proceedings of the 2021 6th International Conference for Convergence in Technology (I2CT), Mumbai, India, 2–4 April 2021; pp. 1–7.
133. Da'u, A.; Salim, N. Recommendation system based on deep learning methods: A systematic review and new directions. *Artif. Intell. Rev.* **2020**, *53*, 2709–2748. [CrossRef]
134. Chen, Q.; Zhao, H.; Li, W.; Huang, P.; Ou, W. Behavior sequence transformer for e-commerce recommendation in Alibaba. In Proceedings of the 1st International Workshop on Deep Learning Practice for High-dimensional Sparse Data, New York, NY, USA, 5–6 August 2019; pp. 1–4.
135. Ay, B.; Aydın, G.; Koyun, Z.; Demir, M. A visual similarity recommendation system using generative adversarial networks. In Proceedings of the 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), Istanbul, Turkey, 26–28 August 2019; pp. 44–48.
136. Vieira, A.; Ribeiro, B.; Vieira, A.; Ribeiro, B. Recommendation algorithms and e-commerce. *Introd. Deep. Learn. Bus. Appl. Dev. Conversational Bots Cust. Serv. Med. Image Process.* **2018**, *1*, 171–184.
137. Huang, G. E-commerce intelligent recommendation system based on deep learning. In Proceedings of the 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), Dalian, China, 14–16 April 2022; pp. 1154–1157.
138. Elmurngi, E.; Gherbi, A. An empirical study on detecting fake reviews using machine learning techniques. In Proceedings of the 2017 Seventh International Conference on Innovative Computing Technology (INTECH), Luton, UK, 16–18 August 2017; pp. 107–114.
139. Elmurngi, E.I.; Gherbi, A. Unfair reviews detection on amazon reviews using sentiment analysis with supervised learning techniques. *J. Comput. Sci.* **2018**, *14*, 714–726. [CrossRef]
140. Wang, X.; Zhang, X.; Jiang, C.; Liu, H. Identification of fake reviews using semantic and behavioral features. In Proceedings of the 2018 4th International Conference on Information Management (ICIM), Oxford, UK, 25–27 May 2018; pp. 92–97.
141. Baishya, D.; Deka, J.J.; Dey, G.; Singh, P.K. SAFER: Sentiment analysis-based FakE review detection in E-commerce using deep learning. *SN Comput. Sci.* **2021**, *2*, 1–12. [CrossRef]
142. Aono, T.V.M. Fake review detection focusing on emotional expressions and extreme rating. *Assoc. Nat. Lang. Process.* **2019**, *13*, 422–425.
143. Alsubari, S.N.; Deshmukh, S.N.; Aldhyani, T.H.; Al Nefaie, A.H.; Alrasheedi, M. Rule-based classifiers for identifying fake reviews in e-commerce: A deep learning system. In *Fuzzy, Rough and Intuitionistic Fuzzy Set Approaches for Data Handling: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 257–276.
144. Bathla, G.; Kumar, A. Opinion spam detection using Deep Learning. In Proceedings of the Proceedings of the 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 26–27 August 2021; pp. 1160–1164.
145. Ashraf, M.S.; Rehman, F.; Sharif, H.; Aqeel, M.; Arslan, M.; Rida, A. Spam Consumer's Reviews Detection for E-Commerce Website using Linguistic Approach in Deep Learning. In Proceedings of the 2022 3rd International Conference on Innovations in Computer Science & Software Engineering (ICONICS), Karachi, Pakistan, 14–15 December 2022; pp. 1–7.
146. Chowdhury, R.; Zaman, F.U.; Sharker, A.; Rahman, M.; Shah, F.M. Rate Insight: A Comparative Study on Different Machine Learning and Deep Learning Approaches for Product Review Rating Prediction in Bengali Language. In Proceedings of the IEEE 2022 25th International Conference on Computer and Information Technology (ICCIT), Bazar, Bangladesh, 17–19 December 2022; pp. 406–411.
147. Ahmed, B.H.; Ghabayen, A.S. Review rating prediction framework using deep learning. *J. Ambient Intell. Humaniz. Comput.* **2022**, *13*, 3423–3432. [CrossRef]
148. Bathla, G.; Singh, P.; Singh, R.K.; Cambria, E.; Tiwari, R. Intelligent fake reviews detection based on aspect extraction and analysis using deep learning. *Neural Comput. Appl.* **2022**, *34*, 20213–20229. [CrossRef]
149. Alsmadi, A.; AlZu'bi, S.; Hawashin, B.; Al-Ayyoub, M.; Jararweh, Y. Employing deep learning methods for predicting helpful reviews. In Proceedings of the 2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 7–9 April 2020; pp. 7–12.
150. Zhang, D.; Li, W.; Niu, B.; Wu, C. A deep learning approach for detecting fake reviewers: Exploiting reviewing behavior and textual information. *Decis. Support Syst.* **2023**, *166*, 113911. [CrossRef]
151. Ren, Y.; Zhang, Y. Deceptive opinion spam detection using neural network. In Proceedings of the 26th International Conference on Computational Linguistics, Osaka, Japan, 16–17 December 2016; pp. 140–150.
152. Shahariar, G.; Biswas, S.; Omar, F.; Shah, F.M.; Hassan, S.B. Spam review detection using deep learning. In Proceedings of the 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 17–19 October 2019; pp. 27–33.
153. Alarfaj, F.K.; Malik, I.; Khan, H.U.; Almusallam, N.; Ramzan, M.; Ahmed, M. Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms. *IEEE Access* **2022**, *10*, 39700–39715. [CrossRef]

154. Ashraf, M.; Abourezka, M.A.; Maghraby, F.A. A comparative analysis of credit card fraud detection using machine learning and deep learning techniques. In *Digital Transformation Technology*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 267–282.
155. Najadat, H.; Altiti, O.; Aqouleh, A.A.; Younes, M. Credit card fraud detection based on machine and deep learning. In Proceedings of the 2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 7–9 April 2020; pp. 204–208.
156. Guo, Q.; Li, Z.; An, B.; Hui, P.; Huang, J.; Zhang, L.; Zhao, M. Securing the deep fraud detector in large-scale e-commerce platform via adversarial machine learning approach. In Proceedings of the 2019 the World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 616–626.
157. Zhao, M.; Li, Z.; An, B.; Lu, H.; Yang, Y.; Chu, C. Impression allocation for combating fraud in e-commerce via deep reinforcement learning with action norm penalty. In Proceedings of the 2018 IJCAI, Stockholm, Sweden, 13–19 July 2018; pp. 3940–3946.
158. Degirmenci, S.; Jones, C. Benchmarking offline reinforcement learning algorithms for e-commerce order fraud evaluation. *arXiv* **2022**, arXiv:2212.02620.
159. Tax, N.; de Vries, K.J.; de Jong, M.; Dosoula, N.; van den Akker, B.; Smith, J.; Thuong, O.; Bernardi, L. Machine learning for fraud detection in e-commerce: A research agenda. In Proceedings of the 2nd International Workshop on Deployable Machine Learning for Security Defense, Virtual Event, 15 August 2021; Springer: Berlin/Heidelberg, Germany, 2021; pp. 30–54.
160. Raghavan, P.; El Gayar, N. Fraud detection using machine learning and deep learning. In Proceedings of the 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 11–12 December 2019; pp. 334–339.
161. Liu, J.; Gu, X.; Shang, C. Quantitative detection of financial fraud based on deep learning with combination of e-commerce big data. *Complexity* **2020**, *2020*, 6685888. [CrossRef]
162. Sailusha, R.; Gnaneswar, V.; Ramesh, R.; Rao, G.R. Credit card fraud detection using machine learning. In Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 13–15 May 2020; pp. 1264–1270.
163. Santharam, A.; Krishnan, S.B. Survey on customer churn prediction techniques. *Int. Res. J. Eng. Technol.* **2018**, *5*, 3.
164. Dingli, A.; Marmara, V.; Fournier, N.S. Comparison of deep learning algorithms to predict customer churn within a local retail industry. *Int. J. Mach. Learn. Comput.* **2017**, *7*, 128–132. [CrossRef]
165. Dhote, S.; Vichoray, C.; Pais, R.; Baskar, S.; Mohamed Shakeel, P. Hybrid geometric sampling and AdaBoost based deep learning approach for data imbalance in E-commerce. *Electron. Commer. Res.* **2020**, *20*, 259–274. [CrossRef]
166. Pondel, M.; Wuczyński, M.; Gryncewicz, W.; Łysik, Ł.; Hernes, M.; Rot, A.; Kozina, A. Deep learning for customer churn prediction in e-commerce decision support. In Proceedings of the 24th International Conference on Business Information Systems, Hannover, Germany, 14–17 June 2021; pp. 3–12. [CrossRef]
167. Wu, Z.; Jing, L.; Wu, B.; Jin, L. A PCA-AdaBoost model for e-commerce customer churn prediction. *Ann. Oper. Res.* **2022**, *308*, 1–18. [CrossRef]
168. Agrawal, S.; Das, A.; Gaikwad, A.; Dhage, S. Customer churn prediction modelling based on behavioural patterns analysis using deep learning. In Proceedings of the 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE), Shah Alam, Malaysia, 11–14 July 2018; pp. 1–6.
169. Baghla, S.; Gupta, G. Performance evaluation of various classification techniques for customer churn prediction in e-commerce. *Microprocess. Microsystems* **2022**, *94*, 104680. [CrossRef]
170. Tariq, M.U.; Babar, M.; Poulin, M.; Khattak, A.S. Distributed model for customer churn prediction using convolutional neural network. *J. Model. Manag.* **2022**, *17*, 853–863. [CrossRef]
171. Pekel Ozmen, E.; Ozcan, T. A novel deep learning model based on convolutional neural networks for employee churn prediction. *J. Forecast.* **2022**, *41*, 539–550. [CrossRef]
172. Mohbey, K.K. Employee's attrition prediction using machine learning approaches. In *Machine Learning and Deep Learning in Real-Time Applications*; IGI Global: Hershey, PA, USA, 2020; pp. 121–128.
173. Koehn, D.; Lessmann, S.; Schaal, M. Predicting online shopping behaviour from clickstream data using deep learning. *Expert Syst. Appl.* **2020**, *150*, 113342. [CrossRef]
174. Vieira, A. Predicting online user behaviour using deep learning algorithms. *arXiv* **2015**, arXiv:1511.06247.
175. Utku, A.; Akcayol, M.A. Deep learning based prediction model for the next purchase. *Adv. Electr. Comput. Eng.* **2020**, *20*, 35–44. [CrossRef]
176. Zhang, W.; Wang, M. An improved deep forest model for prediction of e-commerce consumers' repurchase behavior. *PLoS ONE* **2021**, *16*, e0255906. [CrossRef]
177. Ali Hakami, N.; Hosni Mahmoud, H.A. Deep learning analysis for reviews in Arabic e-commerce sites to detect consumer behavior towards sustainability. *Sustainability* **2022**, *14*, 12860. [CrossRef]
178. Feng, L. Data analysis and prediction modeling based on deep learning in e-commerce. *Sci. Program.* **2022**, *2022*, 1041741. [CrossRef]
179. Ma, J.; Guo, X.; Zhao, X. Identifying purchase intention through deep learning: Analyzing the Q&D text of an e-commerce platform. *Ann. Oper. Res.* **2022**, *314*, 1–20.
180. Yin, X.; Tao, X. Prediction of merchandise sales on E-commerce platforms based on data mining and deep learning. *Sci. Program.* **2021**, *2021*, 2179692. [CrossRef]

181. Dharshini, M.P.A.; Vijila, S.A. Survey of machine learning and deep learning approaches on sales forecasting. In Proceedings of the 2021 4th International Conference on Computing and Communications Technologies (ICCCT), Chennai, India, 16–17 December 2021; pp. 59–64.
182. Zhu, H. A deep learning based hybrid model for sales prediction of e-commerce with sentiment analysis. In Proceedings of the 2021 2nd International Conference on Computing and Data Science (CDS), Stanford, CA, USA, 19–20 January 2021; pp. 493–497.
183. Petroșanu, D.M.; Pîrjan, A.; Căruţaşu, G.; Tăbușcă, A.; Zirra, D.L.; Perju-Mitran, A. E-Commerce Sales Revenues Forecasting by Means of Dynamically Designing, Developing and Validating a Directed Acyclic Graph (DAG) Network for Deep Learning. *Electronics* **2022**, *11*, 2940. [CrossRef]
184. Qi, Y.; Li, C.; Deng, H.; Cai, M.; Qi, Y.; Deng, Y. A deep neural framework for sales forecasting in e-commerce. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; pp. 299–308.
185. Zahavy, T.; Krishnan, A.; Magnani, A.; Mannor, S. Is a picture worth a thousand words? A deep multi-modal architecture for product classification in e-commerce. In Proceedings of the 2018 AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
186. Gupta, P.; Raman, S. Translate2Classify: Machine translation for e-commerce product categorization in comparison with machine learning & deep learning classification. In Proceedings of the 2021 Data Analytics and Management (ICDAM 2021), Polkowise, Poland, 26 June 2021; Springer: Berlin/Heidelberg, Germany, 2022; pp. 769–788.
187. Dai, J.; Wang, T.; Wang, S. A deep forest method for classifying e-commerce products by using title information. In Proceedings of the 2020 International Conference on Computing, Networking and Communications (ICNC), Big Island, HI, USA, 17–20 February 2020; pp. 1–5.
188. Samia, B.; Soraya, Z.; Malika, M. Fashion images classification using machine learning, deep learning and transfer learning models. In Proceedings of the 2022 7th International Conference on Image and Signal Processing and Their Applications (ISPA), Mostaganem, Algeria, 8–9 May 2022; pp. 1–5.
189. Sharma, R.; Vishvakarma, A. Retrieving similar e-commerce images using deep learning. *arXiv* **2019**, arXiv:1901.03546.
190. Mohammad, M.N.; Kumari, C.U.; Murthy, A.S.D.; Jagan, B.O.L.; Saikumar, K. Implementation of online and offline product selection system using FCNN deep learning: Product analysis. *Mater. Today: Proc.* **2021**, *45*, 2171–2178. [CrossRef]
191. Tuinhof, H.; Pirker, C.; Haltmeier, M. Image-based fashion product recommendation with deep learning. In Proceedings of the 4th International Conference on Machine Learning, Optimization, and Data Science, Volterra, Italy, 13–16 September 2018; Springer: Berlin/Heidelberg, Germany, 2019; pp. 472–481.
192. Vijayaraj, A.; Raj, V.; Jebakumar, R.; Gururama Senthilvel, P.; Kumar, N.; Suresh Kumar, R.; Dhanagopal, R. Deep learning image classification for fashion design. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 7549397. [CrossRef]
193. Rui, C. Research on classification of cross-border E-commerce products based on image recognition and deep learning. *IEEE Access* **2020**, *9*, 108083–108090. [CrossRef]
194. Bhoir, S.; Patil, S. Transfer learning with deep neural networks for image classification in the e-commerce industry. In Proceedings of the 2022 IEEE 7th International conference for Convergence in Technology (I2CT), Pune, India, 7–9 April 2022; pp. 1–8.
195. Xu, J.; Wei, Y.; Wang, A.; Zhao, H.; Lefloch, D. Analysis of clothing image classification models: A comparison study between traditional machine learning and deep learning models. *Fibres Text. East. Eur.* **2022**, *30*, 66–78. [CrossRef]
196. Zhang, P. E-commerce products recognition based on a deep learning architecture: Theory and implementation. *Future Gener. Comput. Syst.* **2021**, *125*, 672–676. [CrossRef]
197. Bruni, M.E.; Fadda, E.; Fedorov, S.; Perboli, G. A machine learning optimization approach for last-mile delivery and third-party logistics. *Comput. Oper. Res.* **2023**, *157*, 106262. [CrossRef]
198. Perboli, G.; Brotcorne, L.; Bruni, M.E.; Rosano, M. A new model for Last-Mile Delivery and Satellite Depots management: The impact of the on-demand economy. *Transp. Res. Part E Logist. Transp. Rev.* **2021**, *145*, 102184. [CrossRef]
199. Perboli, G.; Arabnezhad, E. A Machine Learning-based DSS for mid and long-term company crisis prediction. *Expert Syst. Appl.* **2021**, *174*, 114758. [CrossRef]
200. Cao, Y.; Shao, Y.; Zhang, H. Study on early warning of E-commerce enterprise financial risk based on deep learning algorithm. *Electron. Commer. Res.* **2022**, *22*, 21–36. [CrossRef]
201. Chen, X.; Long, Z. E-commerce enterprises financial risk prediction based on FA-PSO-LSTM neural network deep learning model. *Sustainability* **2023**, *15*, 5882. [CrossRef]
202. Ren, S.; Choi, T.M.; Lee, K.M.; Lin, L. Intelligent service capacity allocation for cross-border-E-commerce related third-party-forwarding logistics operations: A deep learning approach. *Transp. Res. Part E Logist. Transp. Rev.* **2020**, *134*, 101834. [CrossRef]
203. Xie, G.; Huang, L.; Bin, H.; Apostolidis, C.; Jiang, Y.; Li, G.; Cai, W. Sustainable entrepreneurship in rural e-commerce: Identifying entrepreneurs in practitioners by using deep neural networks approach. *Front. Environ. Sci.* **2022**, *10*, 840479. [CrossRef]
204. Zhang, L.; Nan, X.; Huang, E.; Liu, S. Social e-commerce tax evasion detection using multi-modal deep neural networks. In Proceedings of the 2021 Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, Australia, 29–30 November 2021; pp. 1–6.
205. Gao, P.; Zhao, L. Study on deep learning technology to construct e-commerce industry marketing promotion model. *Secur. Commun. Netw.* **2022**, *2022*, 9958398. [CrossRef]

206. Rangra, A.; Sehgal, V.K.; Shukla, S. A Novel approach of cloud based scheduling using deep-learning approach in e-commerce domain. *Int. J. Inf. Syst. Model. Des.* **2019**, *10*, 59–75. [CrossRef]
207. Guo, L. Cross-border e-commerce platform for commodity automatic pricing model based on deep learning. *Electron. Commer. Res.* **2022**, *22*, 1–20. [CrossRef]
208. Wen, M.; Vasthimal, D.K.; Lu, A.; Wang, T.; Guo, A. Building large-scale deep learning system for entity recognition in e-commerce search. In Proceedings of the 6th IEEE/ACM International Conference on Big Data Computing, Applications and Technologies, Auckland, New Zealand, 2–5 December 2019; pp. 149–154.
209. Cai, W.; Song, Y.; Wei, Z. Multimodal data guided spatial feature fusion and grouping strategy for e-commerce commodity demand forecasting. *Mob. Inf. Syst.* **2021**, *2021*, 5568208. [CrossRef]
210. Symeonidis, P.; Tiakas, E.; Manolopoulos, Y. Product recommendation and rating prediction based on multi-modal social networks. In Proceedings of the Fifth ACM Conference on Recommender Systems, Chicago, IL, USA, 23–27 October 2011; pp. 61–68.
211. Jiming, L.; Peixiang, Z.; Ying, L.; Weidong, Z.; Jie, F. Summary of Multi-modal Sentiment Analysis Technology. *J. Front. Comput. Sci. Technol.* **2021**, *15*, 1165.
212. Herberz, M.; Hahnel, U.J.; Brosch, T. The importance of consumer motives for green mobility: A multi-modal perspective. *Transp. Res. Part A Policy Pract.* **2020**, *139*, 102–118. [CrossRef]
213. Wanner, J.; Herm, L.V.; Heinrich, K.; Janiesch, C. Stop ordering machine learning algorithms by their explainability! An empirical investigation of the tradeoff between performance and explainability. In Proceedings of the 2021 Conference on e-Business, e-Services and e-Society, Newcastle upon Tyne, UK, 13–14 September 2022; Springer: Berlin/Heidelberg, Germany, 2021; pp. 245–258.
214. Zhang, Q.s.; Zhu, S.C. Visual interpretability for deep learning: A survey. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, 27–39. [CrossRef]
215. Yang, Z.B.; Zhang, J.P.; Zhao, Z.B.; Zhai, Z.; Chen, X.F. Interpreting network knowledge with attention mechanism for bearing fault diagnosis. *Appl. Soft Comput.* **2020**, *97*, 106829. [CrossRef]
216. Nielsen, I.E.; Dera, D.; Rasool, G.; Ramachandran, R.P.; Bouaynaya, N.C. Robust explainability: A tutorial on gradient-based attribution methods for deep neural networks. *IEEE Signal Process. Mag.* **2022**, *39*, 73–84. [CrossRef]
217. Tan, K.S.; Subramanian, P. Proposition of machine learning driven personalized marketing approach for E-commerce. *J. Comput. Theor. Nanosci.* **2019**, *16*, 3532–3537. [CrossRef]
218. Goldenberg, D.; Kofman, K.; Albert, J.; Mizrachi, S.; Horowitz, A.; Teinemaa, I. Personalization in practice: Methods and applications. In Proceedings of the 14th ACM International Conference on Web Search and Data Mining, Jerusalem, Israel, 8–12 March 2021; pp. 1123–1126.
219. Rakhra, M.; Gopinadh, G.; Addepalli, N.S.; Singh, G.; Aliraja, S.; Reddy, V.S.G.; Reddy, M.N. E-commerce assistance with a smart chatbot using artificial intelligence. In Proceedings of the 2021 2nd International Conference on Intelligent Engineering and Management, London, UK, 28–30 April 2021; pp. 144–148.
220. Landim, A.; Pereira, A.; Vieira, T.; de B. Costa, E.; Moura, J.; Wanick, V.; Bazaki, E. Chatbot design approaches for fashion E-commerce: An interdisciplinary review. *Int. J. Fash. Des. Technol. Educ.* **2022**, *15*, 200–210. [CrossRef]