Survey on Product Defects Detection using Customer Reviews and Machine Learning

Diana Fayez Shosha Boles
Department of Information Systems and Technology
Faculty of Graduate Studies for Statistical Research
Cairo University Egypt

Nesrine Ali AbdelAzim
Department of Information Systems and Technology
Faculty of Graduate Studies for Statistical Research
Cairo University Egypt

Tarek El Ghazaly
Department of Information Systems and Technology
Faculty of Graduate Studies for Statistical Research Cairo University Egypt

ABSTRACT
This paper focuses on the aspect of product defect detection as an essential capability for manufacturers to guarantee customer satisfaction with the quality of products. Successful product defect detection helps companies in avoiding expensive recalls, improving product quality, and building trust with consumers. The rapid rise in the use of e-commerce platforms results in the production of a large amount of information related to product defects from dissatisfied customers and negative feedback. Specifically, customer reviews contain rich information in the form of unstructured data, providing in-depth information on the practical application and product defect problems affecting purchasing decisions. The current study presents an extensive literature review on extracting information about product defects from customer reviews using machine learning approaches. The literature review includes traditional machine learning models, e.g., logistic regression, support vector machines, random forest, and gradient boosting models, as well as deep learning models, including but not limited to convolutional neural networks and long short-term memory networks. More recently, the study includes the application of transformer-based models, including BERT and its variants, because of their strong ability to understand context. The results of the analyzed papers show that deep learning approaches as well as transformer techniques always excel the performance of traditional approaches in the detection of both explicit as well as implicit product defects. This is because the application of the attention mechanism as well as the transformer technique results in better recall as well as F1 score values, making them more useful in the early detection of product defects in highly imbalanced review datasets. The results also manifest important implications related to machine learning in enhancing the detection of product defects, higher product quality, and optimized customer satisfaction.

General Terms
Machine Learning

Keywords
Machine learning techniques, customer reviews, product defect detection.

1. INTRODUCTION
Currently, research focuses on the crucial aspect of having satisfied customers and enhancing the quality of manufactured goods. Online reviews by customers are an essential source of information for gaining knowledge about customer responses and experiences. Machine learning (ML) algorithms is utilized for automatically identifying defects in manufactured goods in online customer responses. Nevertheless, because of the growing number of ML algorithms being used for this purpose, it becomes more difficult to choose an appropriate algorithm for a specific application. Thus, a review of the literature aims to identify cutting-edge ML algorithms for defect identification in manufactured goods through customer review via the internet.

Detecting the defects of products in the ML domain involves the use of ML algorithms in the detection of possible defects in the product prior to its release to the customer. Thus, it helps organizations get rid of expensive product recalls, enhance the quality of products, and improve the satisfaction of customers. To the organization, the accurate detection of defects in their products is a major concern in the management of product quality. Data about product defects could be obtained from customer complaints and responses [13].

Several pieces of the literature employed ML techniques for identifying data relevant to defects [14][15][16]. Data selection and preprocessing, along with employing appropriate algorithms, are needed for product defect identification within an ML environment. These elements collaborate with techniques of validation in order to make sure that the model is reliable and accurate. They aid companies in identifying the causes of product defects and in taking corrective measures against their recurrence. However, a challenge exists with regard to the extraction of valuable data from online reviews despite their availability for product defect identification. Barriers exist within an ML approach for product defect identification, which include the selection of relevant variables, dealing with noisy data, and addressing issues related to the imbalance of classes.

2. BACKGROUND
2.1 Online shopping
Online shopping is a common way of shopping since the increase in internet users. Its activities are easy, flexible, and accessible [1]. It gives users the opportunity to obtain more information to compare products and prices, better product selection, and more ease and convenience in finding the desired product online [2]. Users who have purchased a product can share their opinions using the review feature. It is necessary to make every positive and negative review [1].Reviews represent a form of user experience information about a product or service that can be used as a reference for potential consumers' preferences for purchasing, using, or consuming a product [3]. Looking at the reviews submitted by other consumers to get an overview of the product is essential to shaping the products customers buy online [4]. Product reviews can increase interest in purchasing and using the product. Users can provide reviews on purchased products using the consumer review feature provided by the marketplace. Sellers can use these reviews as

rating material, and potential buyers also get a general view of the products, they are interested in based on the experiences of other consumers. There are many advantages to online shopping, but there are also disadvantages. The products sold often do not match the actual conditions, such as shape, size, and color because they can only be seen based on the picture. It is not like the original condition, as shown in the images [1]. Customer reviews provide useful information for buyer decision-making, such as numerical ratings, accuracy of product descriptions and advertising claims, product quality and seller service, and overall purchase satisfaction. More than 80% of online buyers research and read reviews before purchasing [5]. Sellers also rely on customer reviews to evaluate the products and services they offer and to analyze consumer buying patterns and trends. Therefore, receiving credible online responses from customers plays key roles for both companies and customers [12]. The websites of companies help customers in reviewing the product lists and filtering the quality of ratings given, thereby facilitating and giving credibility to the process of making decisions compared to promotion campaigns [5].

2.2 Machine Learning in Product Defect Detection

This entails employing ML algorithms to detect possible weaknesses in items prior to their distribution to consumers. It assists businesses in circumventing expensive recalls, enhancing the quality of their goods, and elevating customer loyalty. For producers, the prompt and precise detection of manufacturing flaws is a fundamental responsibility of the quality of their goods management, with data on defects gathered by means of critiques and comments from consumers[30]. Research sometimes utilizes ML to determine the data related to defects [14][21][22].

Detecting the defects of products using ML necessitates meticulously selecting and preparing data, alongside the application of suitable algorithms and validation methods in order to guarantee model accuracy and reliability. It may assist firms in identifying the underlying causes of faults and implementing corrective measures to avert their recurrence.

2.3 Machine Learning

ML is an academic investigation of algorithms and statistical models that computer systems utilize to execute certain tasks without explicit programming. These algorithms are employed for many applications like data mining, predictive analytics, and image processing. ML is employed to instruct machines in the more efficient management of data [23].

ML is a rapidly advancing domain that facilitates numerous novel strategies for addressing real-world challenges. It allows the machine to gain knowledge from data without human involvement and is utilized in various fields, including fraud detection, systems for recommendations, and imaging in medicine [24]. ML processes are categorized into three primary types, as seen in Figure 1: supervised learning, unsupervised learning, and reinforcement learning.

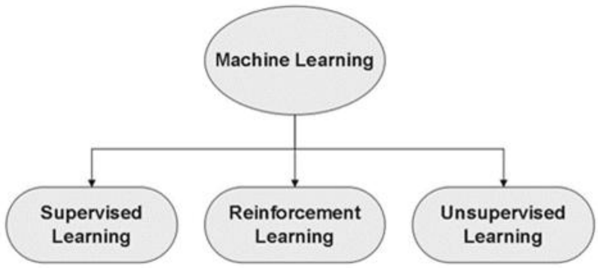

Figure 1: Major Categories of ML [25]

For supervised learning, the model is trained on labeled data. Once trained, it is then used for the prediction of data that are not labeled [25].

For unsupervised learning, the model is trained on the unlabeled data. Using this approach, it is able to learn from the data on its own by identifying patterns [26].

For reinforcement learning, the agent undergoes training within the setting, which facilitates the discovery of optimal solutions and the attainment of goals in complicated scenarios[26].

Supervised learning includes pre-training data of the model on a labeled dataset and involves two different types of learning: classification and regression. Regression is employed when the output is continuous, and classification is used when the output is categorical [25].

3. MACHINE LEARNING ALGORITHMS

A diverse collection of machine learning algorithms as shown in Figure2 was implemented to allow comprehensive comparison.

3.1 Classical Machine Learning Models

3.1.1 Logistic Regression (LR)

Logistic Regression is widely used as a baseline classifier in text classification tasks due to its simplicity and interpretability. When combined with TF-IDF features, LR performs efficiently on high-dimensional sparse text data. However, its linear decision boundary limits its ability to capture complex linguistic patterns present in defect-related reviews [17].

3.1.2 Support Vector Machine (SVM)

Support Vector Machines are effective for high-dimensional text representations and have shown strong performance in binary defect classification tasks. Through optimizing the separation between categories, SVMs ensure strong generalizations. However, the computational expense escalates considerably when dealing with extensive datasets [17].

3.1.3 Random Forest (RF)

Random forest (RF) is a collaborative learning technique that integrates numerous decision trees with the purpose of enhancing classification efficacy and resilience. Its models can manage non-linear relationships and noisy data, rendering them appropriate for identifying defect tasks characterized by complicated attribute associations [27], [28].

3.1.4 XGBoost

XGBoost is an algorithm used to boost gradients that has exhibited exceptional efficacy in numerous detection of defects research investigations. Its capacity to predict unpredictable connections and manage absent data renders it exceptionally successful. XGBoost necessitates meticulous hyperparameter optimization and entails significant computation expenses throughout training [20].

3.1.5 Long Short-Term Memory Networks (LSTM)

Long Short-Term Memory (LSTM) networks are designed for recognizing the sequential relationships that exist in the text data, which makes them ideal for the task of analyzing consumer reviews. LSTMs are effective at learning context-based patterns, e.g., the temporal description of defects, from the data. Although LSTMs are accurate, they impose high costs in terms of computation and always necessitate the use of GPU accelerators [18], [19].

3.1 Transformer-Based Models

Transformer models, including BERT, have shown high-tech use for detecting the defects using customer reviews. The ability of these models to utilize self-attention to overcome long-range dependencies has proven superior to recurrent networks in learning semantic meaning. Yet, their large model capacity, lack of interpretability, and complexity pose significant obstacles to practical applications [18], [20].

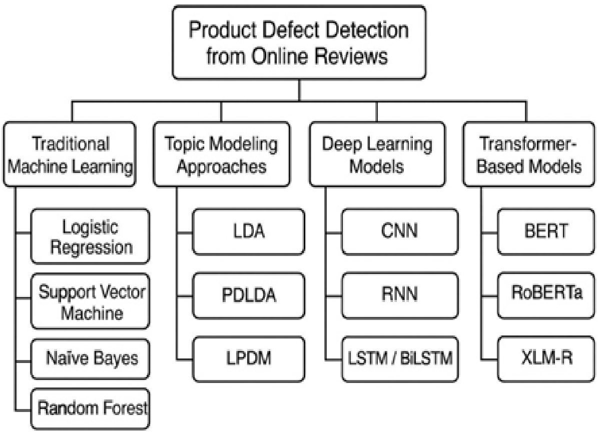

Figure 2: Machine Learning Algorithms for Defect Detection

Recently, there has been an increasing trend of using ML and deep learning approaches for the automatic detection of the defects of products based on web-based responses of customers. Traditional approaches for these early studies included the usage of basic text mining and topic modeling methods. Zhang et al. [9], for example, introduced a "Product Defect Latent Dirichlet Allocation" (PDLDA) model that concurrently identified product defect components, symptoms, and solutions for massive customer-generated content. On similar lines, Qiao et al. [10] presented a domain-specific LDA model that outperformed the basic LDA model for identifying product defects based on vehicle complaints.

Alongside the development of supervised learning techniques, there has been a growing adoption of traditional ML classifiers for defect detection. For instance, Liu et al. [14] adopted contextual and multi-view ensemble learning for detecting defects in online forums for improved accuracy. The effectiveness of feature development and domain-related terminology was highlighted in automotive defect isolation through mining data from social media platforms, as highlighted by Abrahams et al. [15]. Although the results are fruitful, there are limitations in handling noisy data in customer reviews.

Recently carried out research has addressed deep learning models that are capable of extracting semantic and context-based information from consumer feedback. For example, Fong [8] used RNN models in combination with Latent Dirichlet allocation in order to determine negative feedback that pointed out defects in consumer products, while also deriving meaningful inferences about these defects. Williams [7] proved that CNN models are capable of accurately identifying possible consumer product recalls based on feedback gathered from major retail websites. These papers demonstrated that neural models are significantly superior to ML models in modeling complex patterns in language.

The development of models based on the transformer architecture has significantly pushed the high-tech performance for the detection of defects based on texts. For instance, a hybrid model combining BERT with CNN and BiGRU networks was proposed by Yang et al. [18]. It achieved state-of-the-art performance for sentiment classification on electronic commerce product feedback. Moreover, Zhang et al. [20] proposed a hybrid ML approach that incorporates gradient boosting with a deep neural network. This approach achieved better robust and accurate performance in detecting defects. Later, Torres and Singh [19] applied LSTM models with topic models for the automatic discovery of consumer safety alerts with high recall for early defect discovery.

However, the literature has limits regarding scalability, class imbalance, domain dependency, and interpretability. Most current approaches are evaluated on domain-specific datasets, limiting their generalizability across different product categories and platforms. These gaps motivate the need for enhanced machine learning frameworks that balance performance, interpretability, and practical deployment considerations.

Table 1. shows The Comparison of Machine Learning Approaches for Product Defect Detection

Study	Methodology	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Li et al. 2025 [30]	Fine-tuned BERT Transformer	93.4	91.8	92.9	92.3
Torres & Singh2025 [19]	LSTM + Topic Modeling	91.0	88.3	92.6	90.4
Singh et al. 2024 [36]	Topic Modeling + Deep Learning	88.9	85.7	90.4	87.9
Zhao et al. 2024 [35]	Online Learning	89.8	86.5	91.1	88.7

	Neural Network				
Hernandez & Kim 2024 [34]	BiLSTM + Attention + SHAP	90.6	88.9	90.2	89.5
Wang et al. 2024 [33]	Multilingual BERT (XLM-R)	91.3	88.7	89.5	89.1
Alvarez et al 2024. [32]	BERT / RoBERTa Transformers	92.8	90.6	91.9	91.2
Liu et al. 2023 [14]	Ensemble (SVM + RF + GB)	90.2	87.9	86.8	87.3
Fong et al. 2023 [38]	Temporal RNN Model	87.6	84.1	88.3	86.1
Chen et al. 2023 [37]	Attention-based LSTM	88.1	85.2	86.5	85.8
Zhang et al. 2023 [20]	XGBoost + Deep Neural Network	89.4	86.7	83.9	85.3

4. RELATED WORK

[Hammad, A. S., Tajammul, A., Dergaa, I., & Al-Asmakh, M. 2025] [29] investigated (Machine learning applications in the analysis of sedentary behavior and associated health risks). The researchers combined different types of algorithms, threshold-based classification, Hierarchical methods, Decision trees, k-nearest neighbor (KNN), Artificial neural networks (ANN), support vector machines (SVM), Naive Bayes, Gaussian mixture models (GMM), Fuzzy logic, Markov chains, Hidden, Markov models (HMM), combining classifiers (Ensemble learning), unsupervised learning. The researchers obtained data from PubMed articles (2004–2024) Scopus articles (2004–2024),Citation-tracked studies (13 additional articles),Wearable sensor data (accelerometers, smart shirts, smartwatches),Physical activity trackers data (step counts, activity logs),Time-series sensor features (orientation angles, acceleration values, and postural transitions),Individual factors (demographics, health metrics).The results show that sedentary behavior is linked to cardiovascular diseases, cancer, osteoporosis ,metabolic disorders ,musculoskeletal issues, and depression; machine learning algorithms such as KNN, ANN, Decision Trees, HMM,SVM , Fuzzy Logic, and ensemble methods improve activity classification and sedentary behavior prediction, achieving 84–92.4% accuracy, with ensemble and deep learning models (CNN, RNN) further enhancing performance, while wearable sensors combined with ML enable precise monitoring without restricting daily activities, and future directions include causal ML and explainable AI to better understand health behaviors and improve model transparency. The study provides a comprehensive review of 46 recent articles, highlights various health impacts of sedentary behavior, summarizes 11 machine learning techniques for activity classification, demonstrates the effectiveness of ML in predicting sedentary behavior using wearable sensors, discusses emerging trends like ensemble learning, deep learning, causal ML, and explainable AI, and offers practical insights for designing interventions to reduce inactivity.

[Torres, P., & Singh, K. 2025] [19] presented research on (Automated Mining of Consumer Safety Signals Using LSTM And Topic Modeling]. They proposed an automated framework for mining consumer safety signals from online customer reviews using LSTM networks combined with probabilistic topic modeling. The paper examined the early identification of signals for defects before a product recall. The authors used consumer safety data sets and electronic commerce feedback to assess their proposed method. The results of their experiments showed a high recall for rare but highly important reviews related to defects. The importance of early detection of defects has been highlighted by the authors to significantly lower the costs of recalls and enhance consumer safety.

[Li, H., Zhang, D., & Wang, H. 2025] [30] carried out a study on (Transformer-based defect detection from online customer reviews using BERT). They explored the use of transformer models for online customer review–based defect detection using BERT. As mentioned in their publication in expert systems with applications, their work introduces a method for product defect detection that utilizes fine-tuned BERT models for the detection of both implicit and explicit product defects in the feedback of customers using multiple languages online. The method tackles one of the most important challenges of implicit defects, where consumers air their grievances without pointing out the actual defect. The method was tested on electronic commerce data with highly imbalanced classes. The results showed a considerable improvement in precision, recall, and F1-measure compared with conventional machine learning models and models based on LSTMs.

Zhao, Q., Sun, M., and Zhou, Y. (2024) [31] conducted a study on (Real-time defect detection from streaming customer reviews using online learning). They studied online learning for real-time defect detection in streaming consumer reviews. The work described an approach to model online learning for handling concept drift in consumer reviews. Experimental findings illustrated that the approach had stable performance, thus making it applicable to practical electronic commerce sites with large numbers of reviews.

[Alvarez, M., Chen, Y., & Gupta, R. 2024] [32] performed a study on "Transformer-based Product Defect Detection From Large-scale E-commerce Reviews," suggesting a framework based on a transformer to detect product defects from a large number of electronic commerce customer feedback. In their study, fine-tuned models of BERT and RoBERTa were utilized to detect implicit expressions of defects in various product types. Experiments conducted on the Amazon and Alibaba reviews showed a significant improvement in recall and F1-measure of transformer models over conventional ML models and CNN models, especially when it comes to infrequent reviews related to defects.

In a study conducted by [Wang, S., Liu, J., & Park, Y. (2024] on (Multilingual Product Defect Detection Using Cross-lingual Transformer Models)[33], cross-lingual models of the transformer type were explored for their ability to detect product defects in a multi-lingual setting. The authors were trying to solve the problem of customer feedback being written in varying languages by using XLM-R and multi-lingual BERT models. Their method accomplished credible performance without the need for retraining in a specific language. This finding demonstrated the approach's appropriateness for international websites of electronic commerce.

Hernandez, L., and Kim, J. (2024) [34] carried out a study on (Explainable Deep Learning For Product Defect Detection Using Attention-based BiLSTM And SHAP) and analyzed the use of explainable deep learning techniques in product defect detection using the combination of attention-based BiLSTM models and SHAP explanations. This research focused on the interpretability of the model and demonstrated the use of attention weights and SHAP in pointing out the words and phrases that are indicative of defects. The experiment showed the effectiveness of the model in maintaining detection accuracy while improving interpretability in the field of industrial quality assurance tasks.

[Singh, R., Patel, A., & Mehta, S. 2024] [35] presented a paper on (Hybrid Topic Modeling And Deep Learning For Early Defect Signal Extraction). They carried out a study to explore a hybrid method combining topic modeling techniques with deep learning models for the purpose of early defect signal extraction. Latent topics related to defects in products were discovered by the use of neural topic models. After which, supervised classification techniques were applied by deep learning models. The study proves that hybrid models are effective for early warning systems by decreasing the number of false negatives.

[Zhang, D., Li, H., & Wang, H. 2023] [36] conducted a study on (A Hybrid Machine Learning Framework For Defect Detection In Product Reviews). They combined XGBoost and deep learning models for product defect detection from customer feedback. The authors' approach utilized gradient boosting for feature selection and deep learning for learning the semantic representation of text in customer reviews. The authors tested their model on electronic commerce datasets and found it to be the best regarding providing accurate and reliable results and F1 measure. The study demonstrated that integrating traditional machine learning with deep learning can effectively handle noisy and sparse textual data.

[Chen, Y., Xu, L., & Zhao, H. 2023] [37] conducted research about (Attention-based LSTM Networks For Implicit Product Defect Detection) employed attention-based LSTM networks to identify defect-related mentions in customer feedback data. It enabled the model to focus on important defect-indicative words and phrases within long reviews. Experimental results showed that the proposed approach outperformed traditional bag-of-words, TF-IDF, and standard LSTM models, particularly in detecting implicit defect descriptions.

[Fong, T. H. Y., Sarkani, S., & Fossaceca, J. 2023] [38] conducted research about (Temporal Modeling Of Online Customer Reviews For Early Defect Detection) and extended earlier recurrent neural network-based defect detection models by incorporating temporal dynamics of online customer reviews. Their study emphasized the importance of review timing and evolution in identifying emerging product defects. Results demonstrated that temporal modeling improved early defect detection accuracy, allowing manufacturers to respond proactively before defects escalated into large-scale recalls.

[Ahmad, W. 2022][8] presented a paper on (Improving Detection Accuracy for Fake Customer Product Reviews). He used WEKA open-source software to apply 5 algorithms (DNN, Naïve Bayes (NB), SMO(SVM), AdaBoost, Random Forest) to detecting fake reviews. This is taxonomy of 800 fake reviews and 800 real reviews of 20 hotels in the Chicago area. The results obtained showed the greater classification accuracy of (89.1%) obtained with (DNN) was statistically significant at the (.05) level. The algorithm with the accuracy scores closest to (DNN) was SMO (SVM) at (86.67%). Researcher provided by study and comparison with previous studies in this issue that (DNN) more powerful tool than (SVM) for detecting fake reviews with high accuracy.

[Hassan, S. U., Ahamed, J., & Ahmad, K. 2022][7] made a paper on (Analytics of machine learning-based algorithms for text classification). The authors used five algorithms: (Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Logistic Regression (LR), Multinomial Nave Bayes (MNB), and Random Forest (RF). Data were collected from (IMDB, Spam), with available online feedback totaling 50,000 records with two attributes: A sentiment and a review. SPAM is about the normal SMS messages with a two-level label: important and spam. This dataset contained, at the time of conducting the study, 50,572 records with two attributes, message and label. The findings revealed that the k-NN model outperformed others in the spam dataset with an accuracy of 98.5%. On the contrary, the LR model surpassed the others in the IMDB dataset, rating 85.8% accuracy. Researchers developed two models, the first one (K-NN) in the Spam dataset, and the second one (LR) in the IMDB dataset with high accuracy.

[Emil R. Kaburuana1, Yunita Sartika Sari b2, Ika Agustinaa3,2 2022][1] presented a paper on (Sentiment Analysis on Product Reviews from Shopee Marketplace using the Naïve Bayes Classifier). They combined different types of algorithms i.e., KNN, Naïve Bayes Classifier for analyzing sentiments of Shopee feedbacks and adding the Jaro Winkler Distance algorithm for word improvement. The reviews left by Shopee marketplace customers who had purchased home wear products were used to compile the data. utilizing the Shopee API and Google Collaboratory's Python programming language. The review and rating columns on the product review page were the sources of the data. Sentiment Analysis was performed using review data from Google Play to compare the accuracy of the Support Vector Machine and Decision Tree methods. The Support Vector Machine method achieves accuracy of (90.20%), while the Decision Tree method achieves an accuracy of 89.80%.The Naïve Bayes Classifier algorithm can classify reviews on discarded products into positive and negative sentiments with a reasonably high accuracy of (90.03%). Another algorithm could achieve the highest accuracy. Hence, it could be better if conducted as an application or system that automatically performs from data tracking to visualization. Consequently, the system is more beneficial for varying interested parties.

[Dridi, S. 2021][24] conducted research on (Supervised

Learning-A Systematic Literature Review). He proposed numerous supervised learning models (Naïve Bayes, XGBoost, Support Vector Machine, Decision Tree, Random Forest, Logistic Regression (LR), and K-Nearest Neighbor (KNN)). The paper included data from articles and proceedings dated to 2011, totaling (139) articles. "IEEE Xplore," "the Association for Computing Machinery (ACM)," and "Science Direct Elsevier" served as the three databases employed for getting these articles. Findings revealed that (NB) outperformed other algorithms. It made correct categorizations of (95%) of users' spam emails. The survey could be beneficial for scholars in using a supervised learning model or algorithm to employ for addressing problems and research areas.

[Anas, S. M., & Kumari, S. 2021, January][28] conducted research on (Opinion mining based fake product review monitoring and removal system). They developed supervised learning model for labelling the reviews using)Naïve Bayes and random forest methods. The dataset was collected from “amazon academic review”. The Yelp dataset contains information for 11,537 businesses. This dataset has (8,282) check-in sets, (43,873) users, (229,907) reviews for these businesses. (www.yelp.com/dataset).The results showed that The Random Forests (RF) model performed well compared to the Naïve Bayes (NB) algorithm in the fake review data analysis. Researchers developed a model of fake review detection and its elimination.

[Fong, T. H. Y., Sarkani, S., & Fossaceca, J. [2021][6] carried out a study on (Auto Defect Detection Using Customer Reviews for Product Recall Insurance Analysis). The researchers conducted research and developed a new model integrated into LDA and RNN in order to give the insurers and the manufacturers an early insight into the defects of products. The main source of data in this study is (OCRs) and associated metadata from the Amazom.com marketplace. This dataset contains customer text reviews, product information, review dates, star ratings, and other relevant information. The furniture section subset of this dataset was used in this study. Results revealed that both insurers and policyholders are able to detect early manifestations of possible defects and improvement opportunities when applying the new model on (8) of the bestselling Amazon home furnishing products. This new predictive model could help beneficiaries with an early risk analysis of product defects and their detection mechanisms.

[Williams,H. 2021][7] conducted a study to analyze customer reviews (Predicting Product Recall by Using Machine Learning to Analyze Customer Reviews) the study used Convolutional Neural Networks (CNN) as a solution to identify potential recall products by examining information contained in online customer reviews (OCRs) of consumer products. Two different datasets were obtained from Walmart’s website: Furniture, which comprised 5,637 records, and power equipment, which comprised 4,532 records. The results suggest that CNNs are a reliable solution to assist product managers in determining possible product recalls. This is the first study to develop a method for expecting possible consumer product recalls.

[Fong, T. H. Y. 2020][8] presented research on (Identifying Product Defects by Applying a Predictive Model to Customer Reviews). The used models were (RNN) for analyzing the sentiments observed in the feedback of customers in order to identify negative opinions that highlighted product defects, and then the LDA model was applied to summarize the major defect insight words from the feedback. The principal source of the data collected was the (OCRs) and their associated metadata from the (Amazon.com) marketplace. The data contained customer text feedback, product information, review dates, and star ratings. The furniture section subset of this dataset was included. The findings illustrated that engineering teams could identify early manifestations of possible defects and enhancement opportunities when employing the new model on (eight) of the bestselling Amazon home furnishing products. The study investigated the ultimate way to give engineers an early insight into product defects.

[Yang, L., Li, Y., Wang, J., & Sherratt, R. S. 2020][18] conducted research on (sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning) and proposed a new model (SLCABG) for sentiment analysis on product reviews, constructed via a sentiment dictionary, the BERT model, the CNN model, the BiGRU model, and an attention mechanism. They conducted research on the book review dataset from the real electronic commerce platform in order to determine how effective the model is. By analyzing the findings, it was revealed that the model outperformed in terms of classification other sentiment analysis models. In comparison with other models, SLCABG could enhance the sentiment features of the input text and integrate the text context features and main features to improve the quality of classification of the sentiment analysis model.

[Zhang, X., Qiao, Z., Ahuja, A., Fan, W., Fox, E. A., & Reddy, C. K. 2019, May][9] presented a paper on (discovering product defects and solutions from online user generated contents). The researchers conducted research about the discovery of product defects and solutions from online user generated content. They proposed the LDA based latent Product Defect Allocation (PDLDA) model, a probabilistic model that could identify domain-specific knowledge about product problems issues, making use of the interrelated 3D topics: Component, symptom, and resolution. The dataset was collected from the complaint database of (NHTSA3), the problem discussion threads in the official forum of Apple, and the disease discussion threads in the Patient.info forum. In the (NHTSA) and Apple datasets, we aim to detect product defects and resolutions. In the Patient.info dataset, human illnesses can be seen as similar to product defects. Researchers created a (PDLDA) model are shown to efficiently determine the defects of products and to excel other current methods. PDLDA outperformed others in terms of the existing aspect summarization models. It cooperatively identified major entities of product defects as an interdependent 3D theme. Furthermore, it reduces the dependence on word pairs that are current models require.

[Rakesh, V., Ding, W., Ahuja, A., Rao, N., Sun, Y., & Reddy, C. K. (2018, April] [14] presented a paper on (A sparse topic model for extracting aspect-specific summaries from online reviews). The researchers proposed a model called APSUM, which is a generative aspect summarization model that can provide fine-grained summaries of web-based feedback. The dataset was from three different domains. (a) Restaurant reviews from SEMVAL, b) Movie reviews from IMDB, and (c) Product reviews from Amazon. APSUM outperformed others, as it produced better accuracy scores. In the case of making comparisons between TTM and APSUM, APSUM scored a 6-7% gain over top-5 and top10 words. However, it declined to just 3% in case of including the top20 words. In the case of making comparisons between the suggested model, on one hand, and MG-LDA and M-ASUM, on the other, it was noticed that performance increased up to 15%. Predictably, the best gain attained by APSUM outperformed the standard LDA model, rating 37% enhancement compared to the top 10 ranked words. In sum, APSUM excelled others by giving more accuracy scores.

[Qiao, Z., Zhang, X., Zhou, M., Wang, G. A., & Fan, W. 2017][10] presented a study about (A domain oriented LDA model for mining product defects from online customer reviews). They proposed a new unsupervised Bayesian inference model, known as the latent product defect model (integrated with the Gibbs sampling algorithm). Data were collected from the NHTSA open database, with vehicle complaints totaling 1.13 million. Gibbs also contains negative feedback about varying vehicle models. The researchers introduced a new LDA model, called (LPDM), in order to define and get integral information about the defects of products. Their findings revealed that the suggested model excelled the standard LDA model and had the ability to obtain further valuable information.

[Zhang, X., Qiao, Z., Tang, L., Fan, W. P., Fox, E. A., & Wang, G. A. 2016] [11] conducted a study on (Identifying Product Defects from User Complaints: A Probabilistic Defect Model). They used the K-Means clustering algorithm with the aim of identifying various defect types. They applied the same set of words and phrases employed for PDM as features in K-Means clustering to fairly compare the data in hand. Data were collected from the complaint database of NHTSA, which has 1.13 million records in total, presenting complaints on various vehicle models. Create a novel model, called (PDM) to identify different types of defects and experimental results show that the proposed model outperforms the K-Means model in terms of accuracy, recall, and F-Measure in most cases. But The study did not report accuracy values, as the evaluation focused on unsupervised clustering tasks where accuracy is not directly applicable. Instead, the model's performance was assessed using precision, recall, and F-measure. For the Chevrolet dataset, the obtained results were: Precision = 87.80%, Recall = 86.90%, and F-measure = 84.53%.

5. CHALLENGES AND RESEARCH GAPS

Although machine learning techniques have demonstrated promising results in detecting product defects from online customer reviews, several challenges remain unresolved.

- **Noisy and Informal Text**

Customer reviews often contain grammatical errors, slang, abbreviations, and ambiguous expressions, which negatively impact model performance [12]. Handling such noise remains a major challenge, especially for traditional machine learning approaches.

- **Implicit Defect Mentions**

Many customers describe defects indirectly rather than explicitly stating the problem. Capturing such implicit defect indicators requires models capable of deep semantic understanding and contextual reasoning [9].

- **Class Imbalance**

Defect-related reviews typically represent a small fraction of the total review volume, leading to highly imbalanced datasets. This imbalance can bias classifiers toward the majority class and reduce defect recall [14].

- **Domain Dependency**

Models trained on reviews from a specific product category often fail to generalize to other domains because of different vocabulary and defect patterns [15].

- **Scalability and Real-Time Constraints**

E-commerce platforms generate massive volumes of reviews daily. Ensuring real-time or near real-time defect detection while maintaining accuracy is still an open research problem [6].

- **Lack of Interpretability**

Deep learning and transformer-based models often operate as black boxes, making it difficult for manufacturers to understand why a defect is detected. This limits trust and adoption in industrial environments [20].

These challenges highlight the need for enhanced machine learning frameworks that integrate robust text representations, imbalance-aware learning strategies, and explainable AI techniques.

6. CONCLUSION & FUTURE WORK

This paper presents an analysis of the current status of the use of ML techniques for the detection of defects within products, as identified from the analysis of online reviews posted by consumers. In the current scenario, with the tremendous growth of electronic commerce sites and the ever-increasing number of user-generated content, the automatic detection of defects has become a key part of the overall quality control process for all current and future products [5], [13].

There has been an apparent evolution in the existing literature with respect to the methodologies adopted for the detection of defects. Early studies were primarily focused on the use of probabilistic topic models and conventional supervised learning models, including Naïve Bayes, Support Vector Machines, and Logistic Regression, in combination with carefully designed text features [10], [11], [16]. Although these models provided efficient and interpretable advantages, they often faced challenges in handling the implicit representation of defects and complex linguistic structures. Later studies incorporated the use of deep learning models, such as Convolutional Neural Networks and LSTM Networks, that significantly improved the model's comprehension of the context and detection capability [18], [19]. Recently, transformer models have emerged with the best results using the self-attention mechanism to handle long-range dependencies in customer feedback [19], [20].

However, despite such achievements, many problems remain unsolved. For instance, the problems of class imbalance, noisy and informal language, domain dependence, and lack of interpretability remain a challenge for the robustness and generalizability of existing defect detection models [12], [14], [21]. Moreover, the costly computational efforts involved in deep learning models and transformer models also pose a challenge for scalability and real-time processing [23].

On the basis of the results obtained from the survey, some potential trends for future investigations that have arisen include: Firstly, the expansion of the scope of the dataset to include varying products and electronic commerce websites may improve the generalizability of the results. Secondly, incorporating highly developed deep learning models, including attention-based BiLSTM and transformer models, can improve implicit and rare expressions of defects. Thirdly, incorporating explainable AI may improve explainability and facilitate the acceptance of automated systems for defect detection in industry [21], [22].

In addition, further pieces of research are needed on multi-lingual and cross-lingual methods for the detection of defects that will help electronic commerce sites around the world. Real-time monitoring systems that are able to process streaming reviews and handle the problem of concept drift are important directions for future studies [6], [7]. Ultimately, better integration of models for the detection of defects and business decision-support systems may help with proactive quality control and quick corrective measures that improve customer satisfaction.

To conclude, ML techniques can be considered very effective for obtaining defect-related knowledge from customer feedback. Although a lot has already been accomplished in this domain, further studies and analyses are required to deal with the current limits and develop an interpretable and reliable defect detection system for product quality management.

7. REFERENCES

[1] Kaburuana, E. R., Sari, Y. S., & Agustinaa, I. (2022). Sentiment Analysis on Product Reviews from Shopee Marketplace using the Naïve Bayes Classifier.

[2] Vasić, N., Kilibarda, M., & Kaurin, T. (2019). The influence of online shopping determinants on customer satisfaction in the Serbian market. Journal of theoretical and applied electronic commerce research, 14(2), 70-89.

[3] Jayadianti, H., Kaswidjanti, W., Utomo, A. T., & Saifullah, S. (2022). Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN. ILKOM Jurnal Ilmiah, 14(3).

[4] Rangsang, R. A., & Millayani, H. (2021). The Effect Of Online Consumer Review On Customer Purchase Decision Process In The E-commerce Site Blibli. com. eProceedings of Management, 8(6).

[5] Ahmad, W. (2022). Improving Detection Accuracy for Fake Customer Product Reviews (Doctoral dissertation, The George Washington University).

[6] Fong, T. H. Y., Sarkani, S., & Fossaceca, J. (2021). Auto Defect Detection Using Customer Reviews for Product Recall Insurance Analysis. Frontiers in Applied Mathematics and Statistics, 38.

[7] Williams, H. (2021). Predicting Product Recall by Using Machine Learning to Analyze Customer Reviews (Doctoral dissertation, The George Washington University).

[8] Fong, T. H. Y. (2020). Identifying Product Defects by Applying a Predictive Model to Customer Reviews (Doctoral dissertation, The George Washington University).

[9] Zhang, X., Qiao, Z., Ahuja, A., Fan, W., Fox, E. A., & Reddy, C. K. (2019, May). Discovering product defects and solutions from online user generated contents. In The World Wide Web Conference (pp. 3441-3447).

[10] Qiao, Z., Zhang, X., Zhou, M., Wang, G. A., & Fan, W. (2017). A domain oriented LDA model for mining product defects from online customer reviews.

[11] Zhang, X., Qiao, Z., Tang, L., Fan, W. P., Fox, E. A., & Wang, G. A. (2016). Identifying product defects from user complaints: A probabilistic defect model. Department of Computer Science, Virginia Polytechnic Institute & State University.

[12] Ma, Y., & Li, F. (2012, October). Detecting review spam: Challenges and opportunities. In 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom) (pp. 651-654). IEEE.

[13] Zheng, L., He, Z., & He, S. (2020). A novel probabilistic graphic model to detect product defects from social media data. Decision Support Systems, 137, 113369.

[14] Liu, Y., Jiang, C., & Zhao, H. (2018). Using contextual features and multi-view ensemble learning in product defect identification from online discussion forums. Decision Support Systems, 105, 1-12.

[15] Abrahams, A. S., Jiao, J., Fan, W., Wang, G. A., & Zhang, Z. (2013). What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings. Decision Support Systems, 55(4), 871-882.

[16] Zhang, X., Niu, S., Zhang, D., Wang, G. A., & Fan, W. (2015). Predicting vehicle recalls with user-generated contents: A text mining approach. In Intelligence and Security Informatics: Pacific Asia Workshop, PAISI 2015, Ho Chi Minh City, Vietnam, May 19, 2015. Proceedings (pp. 41-50). Springer International Publishing.

[17] Pedregosa, F., Varoquaux, G., Gramfort, G., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

[18] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521, 436–444.

[19] Torres, P., & Singh, K. (2025). Automated mining of consumer safety signals using LSTM and topic modeling. IEEE Internet of Things Journal, 12(8), 16744–16756.

[20] Zhang, D., Li, H., & Wang, H. (2023). A hybrid machine learning approach for defect detection in product reviews. Expert Systems with Applications, 225, 120109.

[21] Abrahams, A. S., Jiao, J., Fan, W., Wang, G. A., & Zhang, Z. (2013). What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings. Decision Support Systems, 55(4), 871-882.

[22] Zhang, X., Niu, S., Zhang, D., Wang, G. A., & Fan, W. (2015). Predicting vehicle recalls with user-generated contents: A text mining approach. In Intelligence and Security Informatics: Pacific Asia Workshop, PAISI 2015, Ho Chi Minh City, Vietnam, May 19, 2015. Proceedings (pp. 41-50). Springer International Publishing.

[23] Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet], 9, 381-386.

[24] Dridi, S. (2021). Supervised Learning-A Systematic Literature Review

[25] Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160(1), 3-24.

[26] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

[27] Anas, S. M., & Kumari, S. (2021, January). Opinion mining based fake product review monitoring and removal system. In 2021 6th International Conference on Inventive Computation Technologies (ICICT) (pp. 985-988). IEEE.

[28] McCallum, Andrew. "Graphical Models, Lecture2: Bayesian Network Represention" (PDF). Retrieved 22 October 2019.

[29] Hammad, A. S., Tajammul, A., Dergaa, I., & Al-Asmakh, M. (2025). Machine learning applications in the analysis of sedentary behavior and associated health risks. Frontiers in Artificial Intelligence, 8, 1538807.

[30] Li, H., Zhang, D., & Wang, H. (2025). Transformer-based defect detection from online customer reviews using BERT. Expert Systems with Applications.

[31] Zhao, Q., Sun, M., & Zhou, Y. (2024). Real-time defect detection from streaming customer reviews using online learning. IEEE Transactions on Industrial Informatics.

[32] Alvarez, M., Chen, Y., & Gupta, R. (2024). Transformer-based product defect detection from large-scale e-commerce reviews. Information Processing & Management, 61(1).

[33] Wang, S., Liu, J., & Park, Y. (2024). Multilingual product defect detection using cross-lingual transformer models. Knowledge-Based Systems, 284.

[34] Hernandez, L., & Kim, J. (2024). Explainable deep learning for product defect detection using attention-based BiLSTM and SHAP. Applied Soft Computing, 145.

[35] Singh, R., Patel, A., & Mehta, S. (2024). Hybrid topic modeling and deep learning for early defect signal extraction. Pattern Recognition Letters, 176.

[36] Zhang, D., Li, H., & Wang, H. (2023). A hybrid machine learning framework for defect detection in product reviews. Expert Systems with Applications, 225, 120109.

[37] Chen, Y., Xu, L., & Zhao, H. (2023). Attention-based LSTM networks for implicit product defect detection. Neurocomputing, 512, 35–47.

[38] Fong, T. H. Y., Sarkani, S., & Fossaceca, J. (2023). Temporal modeling of online customer reviews for early defect detection. Decision Analytics Journal, 7.