Journal of Artificial Intelligence and Data Mining (JAIDM), x(x): xxx-xxx, xxxx

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

GAN-Based Anomaly Detection in Social Networks Text Data Using Lasso and Ridge Regression Models

Abolfazl Adressi and Amirhossein Amiri*

Department of Industrial Engineering, Faculty of Engineering, Shahed University, Tehran, Iran

Article Info

Article History:
Received 02 June 2025
Revised 09 September 2025
Accepted 19 December 2025

DOI:10.22044/jadm.2025.16297.2753

Keywords:
Social Networks, Anomaly Classification, Generative Adversarial Networks, Machine Learning, Lasso and Ridge Regression.

**Corresponding Author's: amiri@shahed.ac.ir (A. Amiri).*

Abstract

Identifying and classifying anomalies in textual data from social networks is challenging due to linguistic complexity and diverse user expressions. While deep learning and machine learning techniques offer promise in addressing this problem, their effectiveness is limited by insufficient data. This paper assesses the impact of Generative Adversarial Networks (GANs) on anomaly detection and classification, as well as their relevance for generating synthetic text data. Combining synthetic and real data enhances classification accuracy, particularly in settings with limited data. In this study, Lasso and Ridge regression techniques are employed for anomaly detection and classification. Experimental results demonstrate the superior performance of the proposed model in identifying and classifying anomalies in new datasets generated by GANs. By integrating statistical methods with generative techniques, the solution becomes more interpretable and scalable, making it better suited for advanced text analysis in fast-changing environments such as social media platforms.

1. Introduction

Social networks have become a necessary forum for consumers to share ideas and comments about several goods and services. Still, much user-generated content makes it quite tough to find anomalies including false information, illegal behavior, and inappropriate conduct. These deviations might compromise user confidence and experiences, mislead companies, and damage platform reputation. Often depending on set rules, conventional anomaly detection methods find it difficult to capture the complexity and dynamic properties of these aberrant patterns. Deep learning and machine learning approaches have become powerful tools for evaluating vast textual data to solve this challenge. Realistic synthetic text fit for training anomaly detection has shown promise from generative adversarial networks (GANs). Models consequently improve their ability to identify unusual trends. By revealing latent correlations in the data, therefore enhancing the capacity of the model to identify anomalies, using GAN-generated synthetic text in combination with regression techniques including Lasso and Ridge can increase classification accuracy. This paper investigates how synthetic data produced by GANs affects anomaly identification in the textual data analysis from social networks. We combine generative modeling with statistical regression to provide a stronger approach for identifying anomalies in user comments. Our results show that the precision of anomaly categorization is much enhanced by including GAN-generated synthetic data with statistical regression approaches. On digital platforms, this improvement helps to enable better decision-making and a more trustworthy and reliable online environment.

Commonly referred to as outliers, anomalies in data are events that depart from expected patterns, therefore signaling unusual behavior or even problems. While Aggarwal [1] notes the difficulties in spotting these abnormalities in big and sophisticated datasets, researchers like Chandola et al. [2] stress their relevance in fields including fraud detection. Particularly in dynamic surroundings like social networks, Hodge and Austin [3] define anomalies as outliers stemming from unusual events, systematic errors, or abnormal behavior. As Mao et al. [4] point out, these sites enable quick exchanges and user-generated content, so greatly help to distribute knowledge. Many studies, notably Tiwari et al. [5], see online review sections as social networks where consumer comments directly affect buying decisions and brand loyalty. Ahmed et al. [6] investigate how user interactions on social media and e-commerce sites impact consumer happiness and brand impression. Shan et al. [7] investigate the emotional effect of internet reviews and show how much sentiment-driven comments affect consumer decisions. Collectively, these studies demonstrate the significant role of anomaly detection in enhancing our understanding of user behavior, decision-making, and trust on digital platforms through advanced machine learning techniques. Wide-ranging applications of anomaly detection—a fundamental component of data analysis and machine learning—are cybersecurity, fraud detection, and system monitoring. Three main groups define most approaches for spotting anomalies: deep learning models, machine learning-based approaches, and conventional statistical methods. Traditional methods of anomaly detection and classification are usually based on statistical algorithms and simple machine learning models. These methods use historical data to detect abnormal data and deviant behaviors in datasets. For example, Chandola et al. [2] used traditional anomaly detection methods, including statistical algorithms such as mean and variance analysis and the k-nearest neighbors (KNN) method. Likewise, distance-based approaches and clustering techniques—like K-Means—have been used to spot anomalies employing data points outside of established groupings [3]. These techniques struggle with high-dimensional and noisy data even if they are good for organized and small-scale datasets. Breunig [8] presented the Local Outlier Factor (LOF), a density-based method that provides anomaly scores to data points depending on their local neighborhood density, providing a more dynamic method for outlier detection. Although conventional approaches are valuable, they sometimes fail to handle big and complicated datasets. For more accurate and scalable anomaly detection, advanced machine learning and deep learning techniques must thus be adopted.

Particularly in the management of complex and dynamic data, recent studies have shown the efficiency of machine learning methods in anomaly detection and categorization. Wette and Heinrichs [9] presented OML-AD, an online machine learning method for real-time anomaly identification in time-series data with better accuracy and efficiency than traditional methods. While Domingues et al. [10] investigated unsupervised methods like Density-based spatial clustering of applications with noise (DBSCAN) and Principal Components Analysis (PCA) for clustering anomalies in large-scale datasets, Manzoor et al. [11] proposed the xStream algorithm, a density-based method for detecting anomalies in high-dimensional streaming data to address the issues of unlabeled data. Apart from these developments in machine learning, Lasso and Ridge regression have become increasingly important instruments for anomaly identification especially in high-dimensional and multicollinear datasets. In sectors including pharmacovigilance and financial data analysis, Lasso regression—which uses feature selection by decreasing some coefficients to zero—has been notably applied to increase model efficiency and accuracy (Courtois et al. [12]). Applying an L2 penalty helps ridge regression to avoid overfitting and improve anomaly identification in fields including sensor networks and medical diagnostics. By constantly changing regularizing parameters and optimizing feature selection, adaptive variations of these techniques—such as adaptive Lasso and graphical Lasso—further improve anomaly classification. These methods' increasing relevance in spotting aberrant trends in big and complicated datasets is shown by their incorporation in cybersecurity, fraud detection, and industrial monitoring.

Deep learning has evolved into a potent tool for anomaly detection and classification by offering the ability to learn intricate patterns and automatically extract features from huge datasets. Investigating unsupervised deep learning approaches, Groenewald [13] focused on autoencoders as efficient tools for anomaly detection using reconstruction of normal data and deviation identification. Particularly in medical and industrial uses, Sanapati et al. [14] showed that Convolutional Neural Networks (CNNs) outperform others in identifying anomalies in images. In a similar research, Su et al. [15] used Recurrent Neural Networks (RNNs) on time-series data to illustrate how well they captured temporal dependencies and found erratic abnormalities. GANs have been developed to help further improvement in anomaly identification. Especially helpful in fraud detection, Perera et al. [16] presented OCGAN, a one-class GAN model meant to identify anomalies in datasets containing only one class of normal data. By essentially learning deep representations of normal patterns, Schlegl et al. [17] created f-AnoGAN, an enhanced GAN-based model for medical image analysis, which accelerates anomaly identification. With uses in quality control and time-series analysis, Akcay et al. [18] presented GANomaly, a semi-supervised method that learns normal data distributions and flags occurrences that cannot be adequately reconstructed as anomalies. These studies show how rapidly deep learning—especially GAN-based approaches—has improved anomaly identification in several disciplines including finance, healthcare, and industrial monitoring.

GANs are useful methods for augmenting machine learning and deep learning models, especially for data scarcity or imbalance. By improving model performance in several disciplines, including medicine, image processing, and time-series analysis, GANs create synthetic data that nearly mimics real-world data. By correcting dataset imbalances, Tanaka et al. [19] showed that GAN-generated data could improve machine learning models, particularly in fraud detection and cancer diagnosis. Chen et al. [20] likewise used CycleGAN and Wasserstein GAN (WGAN) to produce varied medical pictures, so enhancing disease diagnosis on MRI and CT scans. Examining GAN-based artificial picture creation for medical diagnostics, Singh et al. [21] found that models trained with Conditional GAN and CycleGAN data attained better accuracy in disease detection. Yahaya et al. [22] produced high-quality biological and brain-related data to support diagnosis and treatment models, extending the use of GANs to neuroscience and medicine. These studies show how well GANs produce synthetic data that enhances model performance, particularly in fields where real data is limited or uneven. Studies such as Naseri et al. [23] demonstrated that BERT-based semantic similarity analysis can serve as a foundation for detecting semantic deviation in user-generated content. Moreover, Yavari and Hasanpour [24] and Ahangari and Sebti [25] have highlighted that sentiment-based features—such as emotional polarity and retweet behavior—can effectively indicate abnormal shifts in user tendencies. Mohammadi Gohar et al. [28] presented a comprehensive survey on GANs for discrete data, with a particular focus on text-related applications. The study reviews GAN architectures, training strategies based on reinforcement learning, evaluation metrics, and highlights existing challenges and future research directions in discrete data generation.

Given the vast volume and unstructured nature of social networks text data, detecting conceptual and behavioral anomalies has emerged as a significant challenge in natural language processing. Combining generative and analytical approaches, such as GANs and regularized regression methods like Lasso and Ridge, provides a powerful framework for identifying abnormal patterns. Using deep learning methods such as GANs, enables modeling the normal distribution of linguistic behavior, thereby identifying outliers with high accuracy. Simultaneously, employing Lasso and Ridge regression supports effective feature selection and enhances model interpretability, making the combined approach well-suited for robust anomaly detection in social network environments.

Considering that textual data usually have complex and irregular structures that make their analysis and processing difficult, and also the diversity in languages, writing styles, and text formats makes it impossible for a fixed and simple approach to correctly identify anomalies, especially in situations where we are faced with a small volume of data, there is a need for hybrid approaches based on machine learning, especially methods that can assist in generating synthetic data suitable for existing datasets, as felt in the literature on the subject. The structure of the paper is as follows: In Section 2, the problem statement and proposed methodology are explained. Also, logical and sequential approach to find and classify anomalies in textual data is described. In Section 3, computational results are presented. The results of the Ridge and Lasso regression methods, as well as the combination of these methods with the GAN deep learning method, are presented in this section. Sensitivity analyses are presented in Section 4. Conclusion and suggestions for future research are given in the final section.

2. Problem Statement and Methodology

Using GANs in conjunction with Lasso and Ridge regression methods, this paper presents a model for anomaly detection and textual data classification derived from social media. This paper makes use of a dataset including Persian customer reviews from a well-known internet business. Focusing on building a basic model for anomaly detection and classification in textual data, the first stage of this project addresses problems including linguistic complexity, high dimensionality, and computational economy. Text's unstructured character and the variety of writing forms make a fundamental rule-based approach insufficient. A 10,000 samples subset is selected to handle these difficulties, and feature selection techniques—including Lasso and Ridge regression—are used in concert with the Bag of Words (BOW) approach to reduce dimensionality and increase model efficacy, hence lowering dimensionality. These techniques (Lasso and Ridge Regression) are chosen because they would improve feature selection and maximize computing economy. The model is evaluated with respect to accuracy, recall, and precision, therefore producing consistent anomaly detection. Then, we create synthetic data using the GAN deep learning technique to boost the volume of data associated to the current dataset. After that, we investigate the new dataset—which combines synthetic and actual data—using Lasso and Ridge regression techniques and investigate the outcomes with reference to the past indices. Comprising the following actions, this study uses a logical and sequential approach to find and classify anomalies in textual data. The complete algorithmic procedure of the proposed approach is outlined in Appendix A.

2.1. Data Collection and Preparation

This study uses a dataset of user reviews from an online retailer, randomly selecting 10,000 reviews for examination. A thorough preparation flow is used to guarantee the best model performance and enhance data quality. This technique includes word stemming using Python libraries, removing stop words, deleting punctuation marks and extraneous symbols, and Python libraries' word stemming by reducing words to their roots, stemming helps unify variants of the same term, hence lowering feature dimensionality and increasing model accuracy. For example, words like “buy”, “buyer”, and “they have bought” were all reduced to the root “buy”, ensuring consistency in text representation. These preprocessing steps enhance the dataset’s suitability for machine learning, increasing efficiency in anomaly detection and classification.

2.2. Feature Selection and Extraction

Textual data analysis depends much on feature selection and extraction, which helps machine learning models to quickly find patterns and lower data dimensionality. This work generates numerical representations from text using the BoW approach, capturing important information for anomaly identification. BoW guarantees that only the most pertinent information is used for analysis by transforming textual data into structured features, enabling efficient model training and enhancement of classification accuracy. A basic tool in natural language processing, the BoW approach views text as numerical vectors depending on word frequency while ignoring word order. Data preparation, vocabulary building, text to feature vector conversion, and feature matrix building comprise the method. BoW's simplicity, quick processing speed, and efficiency in many Natural language processing (NLP) applications help to explain why it is so popular. However, it has several limits, including the creation of high-dimensional vectors and a loss of contextual meaning. Notwithstanding these shortcomings, BoW is still a useful instrument for text analysis especially in cases when computing efficiency is given top importance.

To obtain structured numerical features, the textual comments were first converted into BoW vectors. These vectors were then input into Lasso and Ridge regression models, which reduced the feature space's dimensionality and performed feature selection. A count-based BoW representation was used instead of Term Frequency–Inverse Document Frequency (TF-IDF). After removing stop words, the raw frequency of each token was computed for each comment, and the resulting feature values were standardized using z-score normalization to ensure they were on a comparable scale. The final BoW matrix had dimensions of 10,000 × |V|, where |V| represents the union of tokens that appeared at least ten times across both classes. No explicit distance metric, such as Euclidean distance, was applied, as the Lasso and Ridge logistic regression models learned linear decision boundaries directly from the standardized feature space.

2.3. Lasso Regression Method

Particularly for feature selection and lowering of model complexity in high-dimensional data, Lasso regression is a frequently applied method in regression analysis and machine learning. Lasso reduces the sum of squared errors by using an L1 penalty applied to the regression coefficients, hence promoting sparsity and removing less important characteristics. Retaining just the most pertinent variables improves model interpretability and helps to reduce overfitting. Lasso regression is widely used in many disciplines, including finance, medicine, and text analysis, because of its capacity to raise computational accuracy and efficiency.

The Lasso regression equation is as follows:

$$\hat{\beta} = \arg\min_{\beta}\left(\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p}\left|\beta_j\right|\right) \quad (1)$$

where y_i is the dependent variable value for sample i, x_{ij} is the independent variable j value for sample i, β_0 is the intercept of the regression model, β_j is the coefficient of independent variable j, n is the number of samples, p is the number of independent variables (features) and λ is the penalty parameter that controls the amount of penalty applied.

Lasso regression changes the coefficients of the regression model such that some of them become zero. This function lets Lasso regression efficiently do feature selection, preserving only the significant and influential features in the model. Most of the coefficients approach 0 as the penalty parameter (λ) increases, producing a simpler model. With benefits in feature selection, dimensionality reduction, overfitting prevention, Lasso regression is a strong technique for anomaly identification. Applying an L1 penalty removes less important information, producing more interpretable and effective models. This approach is helpful especially when recognizing important trends is vital in complicated applications such as fraud detection, industrial system monitoring, healthcare diagnostics, and cybersecurity. Its capacity to autonomously choose important characteristics improves forecast accuracy while keeping model simplicity. Widely applied in early disease detection, failure prediction in industrial systems, and cyberattack identification, Lasso regression is a fundamental method for analyzing high-dimensional and complicated data.

Figure 1, show the comparison of Least Squares and Lasso Coefficients in Regression Model Optimization. This Figure depicts a constrained optimization scenario for regression coefficients. The black dot (•) represents the ordinary least squares (OLS) estimate, and the red ellipse contours show the sum of squared errors (SSE) level curves, corresponding to regions of equal error. The L1 penalty in Lasso creates a diamond-shaped constraint region which causes some of the coefficients to shrink exactly to zero. This feature selection power simplifies the model and makes it more interpretable by retaining the most significant variables, and this is particularly beneficial for high-dimensional text data.

2.4. Ridge Regression Method

By adding an L2 penalty to the goal function, Ridge regression—a regularization method— helps to minimize overfitting by reducing model coefficients without setting them exactly to zero.

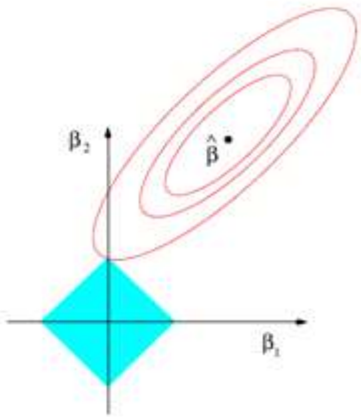

Figure 1. Comparison of Least Squares and Lasso Coefficients in Regression Model Optimization [26].

Particularly with high-dimensional datasets, Ridge regression improves generalization by balancing the trade-off between model complexity and prediction accuracy (Figure 2). In this figure, the red ellipse contours show the SSE level curves, corresponding to regions of equal error. The blue circle represents the L2 constraint region imposed by Ridge regression. In Ridge regression, the optimal coefficients are found at the first point where an SSE ellipse touches the L2 constraint circle. This intersection shrinks the coefficients toward zero but does not set any of them exactly to zero.

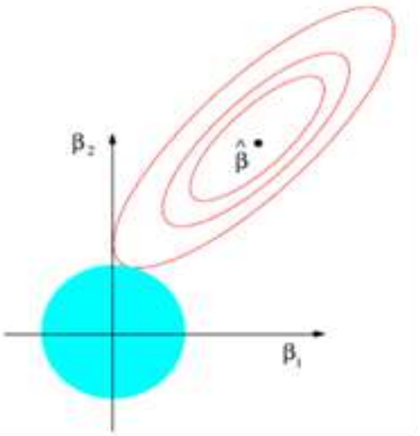

Figure 2. Estimation of Ridge Regression Coefficients Compared to Ordinary Least Squares (OLS) Method [26].

As a result, Ridge enhances model stability and reduces overfitting while retaining all features in the model. Its application involves data preparation, choosing a suitable penalty value, model training, and performance assessment. This approach is often employed in applications requiring stable and interpretable models, making it a useful tool for regression analysis in challenging data settings.

Ridge regression is an appropriate model for high-dimensional data since it provides important benefits including lowered overfitting, enhanced model stability, and the possibility to use all features. Its primary drawback, meanwhile, is that all characteristics are retained, which can lower interpretability relative to techniques like Lasso regression model. Widely used in disciplines such as economics, biology, engineering, and social sciences where forecast accuracy and model stability are crucial, Ridge regression preserves all features by shrinking coefficients, unlike Lasso, which achieves feature selection by lowering some coefficients to zero. This makes it more appropriate in situations where all variables help to forecast. The Ridge regression equation is as follows:

$$\hat{\beta} = \arg\min_{\beta}(\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p}\beta_j^{\,2}) \quad (2)$$

where y_i is the dependent variable value for sample i, x_{ij} is the independent variable j value for sample i, β_0 is the intercept of the regression model, β_j is the coefficient of independent variable j, n is the number of samples, p is the number of independent variables (features) and λ is the penalty parameter that controls the amount of penalty applied.

Regularity methods like Ridge and Lasso regression help to lower model complexity and stop overfitting. Perfect for models where all features help to predict, Ridge utilizes an L2 penalty to decrease all coefficients without deleting any. For choosing important features, Lasso is recommended; for maintaining all predictors, Ridge is more suited. Against techniques including PCA, KNN, and RF, both approaches increase model stability and generalizability. Their strong and quick models make them great methods for predictive analytics and anomaly identification in many fields. The advantages and disadvantages of Ridge and Lasso regression models in comparison with PCA, KNN, and RF are summarized in Table 1.

The designed models (based Lasso and Ridge Regression) are evaluated using key performance metrics, including accuracy, recall, and precision, to assess their effectiveness in anomaly detection. Initial results demonstrated promising efficiency, highlighting the model's capability in identifying anomalies. During this phase, strengths and limitations were analyzed, allowing for targeted optimizations to enhance performance and reliability.

Table 1. Comparison of Lasso and Ridge Methods with Other Similar Approaches in the Literature.

RF	KNN	PCA	Ridge	Lasso	Criterion
Decision Tree	Distance-based	Dimensionality Reduction	Regression	Regression	Model Type
Bagging	No	No	Penalty	Penalty	Regularization
Yes	No	No	No	Yes	Feature Selection
Yes	No	Yes	Yes	Partial	Resistance to Multicollinearity
Yes	No	No	Yes	Yes	Prevention of Overfitting
Low	High	Medium	Medium	High	Simplicity and Interpretability
High	Low	High	Medium	Medium	Computational Complexity
Yes	No	Yes	Yes	Yes	Performance in High Dimensions
Yes	Yes	No	Medium	Medium	Anomaly Detection in Imbalanced Data
Yes	Yes	Low	Medium	Medium	Flexibility
Anomaly detection, feature importance	Distance-based anomaly detection	Dimensionality reduction and Noise removal	Dimensionality ,reduction Resistance to multicollinearity	Feature selectio, dimensionality reduction	Main Applications

2.5. Text Data Generation Using GAN

In this section, we first explain the structure of the GAN network, and then describe the process of text generation.

2.5.1. Structure of GAN

GANs are machine learning techniques that model complex data distribution and produce authentic synthetic samples. They consist of two adversarial neural networks: a generator that generates new data and a discriminator that distinguishes between authentic and fake samples. GANs have various applications, including picture synthesis, synthetic data augmentation for machine learning models, and creative material in music and prose. More advanced variants like WGANs with gradient penalty have been suggested to deal with instability and mode collapse in traditional GAN training. WGANs have been successfully used for not just photo-realistic image creation but also for modeling network traffic and arrival rates, as shown by Abhari for deep learning-based video streaming in edge networks [27]. These examples show WGAN’s ability to closely match complex real-world distributions, making it a strong option for creating realistic synthetic text in future studies. However, training GANs is challenging due to their adversarial characteristics and parameter sensitivity. Despite this, GANs offer enhanced diversity and realism, making them a crucial tool in artificial intelligence applications.

2.5.2. Structure of the Generator Network

The generator network is an essential element of a GAN, tasked with producing synthetic data that closely mimics authentic data. A random input vector, usually sourced from a normal or uniform distribution, is processed via several layers of a deep neural network (DNN). These layers, typically comprising fully connected or transposed convolutional architectures, transform the input into organized outputs such as images or text. Activation methods such as rectified linear unit (ReLU) or Leaky ReLU add complexity to hidden layers, but functions like tanh or sigmoid in the output layer guarantee that the synthetic data remains within an acceptable range. The generator aims to create data that is indistinguishable from authentic data, so misleading the discriminator. Training occurs via backpropagation, wherein the generator incrementally enhances its performance by reducing its loss function in response to feedback from the discriminator. This iterative adversarial approach improves the quality and realism of the synthetic data.

This model's input is a random noise vector processed through various layers to produce an output in the form of a sequence of words. Its goal is generally to generate new comments that reflect actual comments.

2.5.3. Structure of the Discriminator Network

The discriminator in a GAN is essential for differentiating authentic data from synthetic data produced by the generator. It functions as a deep neural network that analyzes input from both the original dataset and the generator, classifying each instance as authentic or counterfeit. The discriminator, typically composed of fully connected or convolutional layers, extracts essential features from the input data. Activation functions such as ReLU or Leaky ReLU facilitate feature learning, while sigmoid or softmax in the final layer produce a binary classification output. The discriminator is trained through supervised learning, using actual data labeled as positive and synthetic data labeled as negative. By optimizing its loss function, the discriminator continually improves its ability to identify counterfeit data, compelling the generator to produce increasingly authentic samples, thereby enhancing the overall quality of the synthetic data. This model helps distinguish generated comments from actual ones, producing a result of 0 or 1 that indicates whether the supplied data is real or synthetic.

Training the GAN model comes next after these phases. Here, the generator model is trained to try to fool the discriminator model while the discriminator model is developed using actual and synthetic data. Following training, the model can provide fresh comments that fall into two categories: recommend and not recommend.

2.5.4. GAN Architecture and Hyperparameters

The proposed GAN uses noise vectors (dim = 100) as input for the generator, which follows the architecture Dense (64) → Dense (128) → Dense (max_sequence_len × vocab_size) → Reshape → LSTM (256) → Dense (vocab_size, softmax). The discriminator flattens sequences and employs Dense (512) → Dense (256) → Dense (128) → Dense (1, sigmoid). LeakyReLU ($\alpha = 0.2$), Dropout (0.3), and the Adam optimizer (learning rate = 2×10^{-4}, $\beta_1 = 0.5$) are used, along with binary cross-entropy loss, a batch size of 32, and label smoothing (0.9 for real samples). Text is tokenized and padded without embedding layers; the generator outputs probability distributions directly. Sampling uses a temperature of 0.8 with repetition limits. Early stopping with a patience of 5 controls training.

3. Computational Results

This part assesses the efficacy of the anomaly detection and classification model utilizing the original dataset. A random sample of 10,000 customer comments was chosen to examine the distribution of satisfaction and displeasure levels. Figure 3 depicts this distribution, offering insights into the predominance of positive and negative feedback, which is crucial for evaluating the model's efficacy in detecting anomalies in customer reviews.

Figure 3. Comparison of the Number of Comments Based on Labels (Original dataset).

The efficiency of the Lasso and Ridge regression models in anomaly detection is investigated in this work. The graph shows that both models can effectively identify regular (recommended) from anomalous (not recommended) signals. Emphasizing their dependability for anomaly identification in textual data, the accuracy rates of 86.50% for Lasso and 86.23% for Ridge expose their considerable capability to detect anomalous signals (Figure 4).

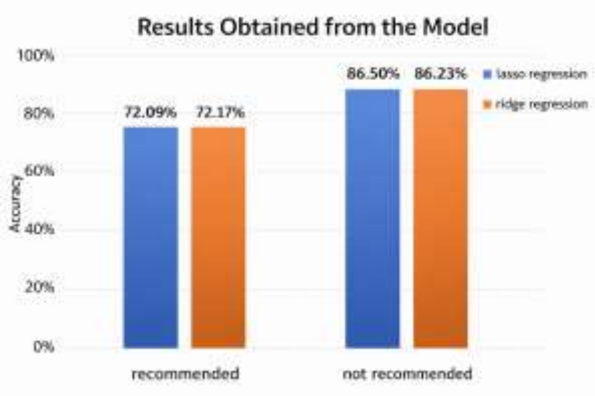

Figure 4. Comparison of the Results Obtained from Lasso and Ridge Regression (Original dataset).

The model's accuracy in identifying normal reviews is 72.09% for Lasso and 72.17% for Ridge, slightly lower than their performance in detecting anomalies but still indicative of strong effectiveness. The similarity in accuracy suggests that both methods perform equally well in anomaly detection. To evaluate the effectiveness of our Lasso and Ridge regression models in identifying anomalies (not recommended reviews), we calculated additional key classification metrics: precision, recall, and F1-score (Table 2). Notably, Ridge regression exhibits a slightly higher precision (86.57%) than Lasso (85.29%), indicating it makes fewer false positive classifications (i.e., it misclassifies fewer normal reviews as anomalies). Ridge also surpasses Lasso in recall (84.06% vs. 82.86%), meaning it detects more of the actual anomalies. The F1-score, which balances precision and recall, is also higher for Ridge (85.29%) compared to Lasso (84.06%), suggesting that Ridge regression provides a more reliable anomaly classification. These results demonstrate that Ridge regression offers superior anomaly detection performance, with a more balanced precision-recall trade-off, making it the preferred choice for detecting anomalies in textual data. These metrics offer valuable insights into the models' ability to accurately identify and classify anomalous reviews within the dataset. The findings underscore the potential of regression-based approaches for complex anomaly detection tasks, particularly in social media analysis and user feedback monitoring. Their high accuracy emphasizes their value in enhancing data analysis processes, fraud detection, and content review on online platforms.

Table 2. Lasso and Ridge Regression Comparison in Anomaly Detection.

F1-Score	Recall	Precision	Model
84.06%	82.86%	85.29%	**Lasso Regression**
85.29%	84.06%	86.57%	**Ridge Regression**

3.1. Performance of the Generator and Discriminator Models in the Training Process

Figures 5 and 6 show the trends in Generator Loss and Discriminator Loss throughout several epochs.

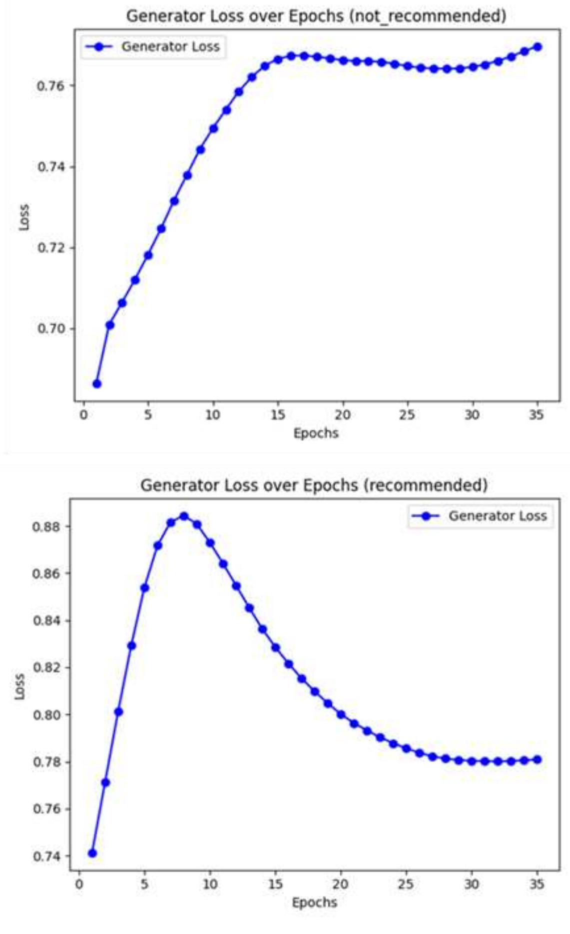

Figure 5. Comparison of the Trend in Generator Losses Across Different Epochs.

One set relates to the "recommended" range; the other relates to the "not recommended" range. These graphs illustrate the development of the generator and discriminator in their capacity to generate and classify real data, clarifying the training dynamics of the GAN model.

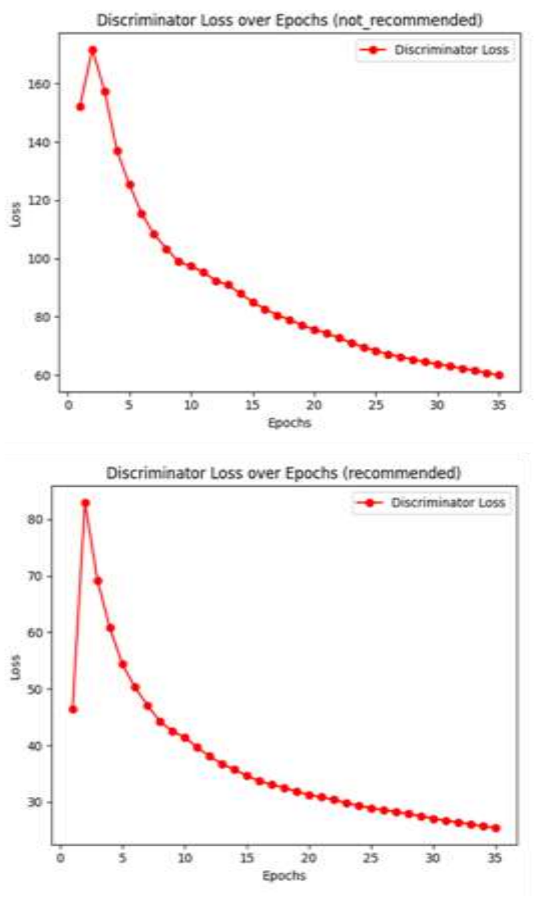

Figure 6. Comparison of the Trend in Discriminator Losses Across Different Epochs.

Figures 5 and 6 show Generator and Discriminator Loss throughout several training cycles. The Generator Loss first rises, peaks about epoch 10, then progressively declines to indicate its improvement in generating realistic data over time. On the other hand, the discriminator loss falls quickly and stabilizes close to epoch 30, suggesting better ability to distinguish genuine from fake data. The Generator Loss shows a similar path in the "not recommended" category, peaking around epoch 15 but slowing down at a slower rate, implying more difficulties producing meaningful data. Reflecting difficulties separating actual from synthetic data within this category, the Discriminator Loss likewise reduces but starts at a greater value. Overall, both models show improvement with more training epochs; nevertheless, the production and classification of synthetic data in the "not recommended" category seem more complex based on the different rates of Loss reduction.

3.2. Performance of the Anomaly Detection and Classification Model in the Combination of Real and Synthetic Data

To increase anomaly detection effectiveness and expand the dataset, this work used a data augmentation technique combining 30% synthetic data with 70% real data. Based on past studies in machine learning and data augmentation, this ratio was chosen to keep data quality while including variability using GAN-generated samples. The balance of synthetic and actual data reduces overfitting and improves the accuracy of the model for exactly identifying anomalies. The results confirm that under several conditions this approach greatly increases model robustness and accuracy. As Figure 7 shows, the data combination with a 20-80 ratio (20% synthetic data with 80% real data) likewise offers performance nearly similar to the situation whereby just real data is used. This slightly supports the theory that the effectiveness of the model in identifying and categorizing anomalies can be improved by combining data instead of original data.

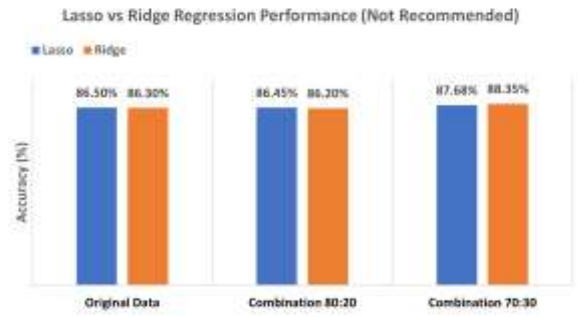

Figure 7. Comparison of Real vs Synthetic Data Accuracy Across Different Ratios.

This chart compares the effectiveness of Lasso and Ridge regression in detecting anomalies with simply actual data vs a hybrid of real and GAN-generated synthetic data. For "recommended" data, both models initially have roughly 72% accuracy; for "non-recommended" data, they have 86%. Still, including fake data significantly improves performance. Accuracy improves to 75% for "recommended" data and exceeds 86% for "non-recommended" data with an 80/20 ratio. With a 70/30 ratio, the ideal accuracy is reached; for Lasso, it is 87.68%; for Ridge, it is 88.35%. These results highlight the effectiveness of synthetic data in improving model performance and show the need of GAN-based data augmentation in enhancing anomaly detection precision and scalability. To assess the contribution of GAN-based augmentation to the overall performance, the Lasso and Ridge classifiers were retrained on the combined real-and-synthetic dataset. This led to consistent improvements in predictive accuracy—from 86.5% to 87.7% for Lasso and from 86.2% to 88.3% for Ridge—together with higher precision, recall, and F1-scores, particularly for the not-recommended class. To further examine the authenticity of the synthetic comments, a humanevaluation study was conducted in which three independent evaluators were asked to distinguish between real and synthetic comments in a balanced sample of 80 instances. Strikingly, more than 70% of the synthetic comments were perceived as real, indicating that the generated text was both fluent and credible. These findings underscore that augmenting the training set with high-quality synthetic data not only strengthens the robustness of the anomaly-detection models but also enhances their practical applicability to real-world social-network text analysis.

3.3. Comparison of the Proposed Model with Conventional Classifiers

To provide a fair benchmark, we compared our proposed method against several conventional classifiers: Multinomial Naïve Bayes, Linear Support Vector Machine (SVM), RF, and a Simple Neural Network (SNN).

These models were evaluated under a GAN-augmented scenario (with 30% GAN-generated synthetic reviews).

Because certain linear models such as Naïve Bayes, SVM, and Logistic Regression typically perform better with term weighting rather than raw counts, we adopted a TF-IDF feature representation instead of a bag-of-words approach for this comparison.

Table 3 reports the derived Precision, Recall, F1-score, as well as the class-specific Accuracy for the “not recommended” class. This table shows that the Logistic Regression models proposed in this study – the Lasso (L1) and Ridge (L2)– deliver the most consistent results among all the evaluated classifiers.

The Ridge model achieved the highest overall accuracy (90.38%), while Lasso also performed strongly with 88.59%, placing both near the top of the table.

For comparison, Naïve Bayes (90.02%) and Linear SVM (89.88%) achieve similar but slightly lower accuracies, and Random Forest lags behind at 80.69%, mainly because of its weaker recall on the “not recommended” class. These findings highlight that regularized logistic regression approaches (Lasso and Ridge) are better suited than the other conventional classifiers for reliably detecting the “not recommended” reviews.

Table 3. Performance (%) of conventional classifiers respect to the proposed models.

Model	Accuracy	Precision	Recall	F1
Lasso	88.59	82.34	85.57	83.92
Ridge	90.38	85.31	87.42	86.35
Naïve Bayes	90.02	93.25	76.91	84.29
Linear SVM	89.88	86.13	84.54	85.33
RF	80.69	92.52	48.45	63.60
SNN	89.59	86.48	83.09	84.75

4. Sensitivity Analysis

4.1. Sensitivity to Regularization Strength

The regularization parameter (C) governs the balance between model complexity and overfitting. Elevated values of C diminish regularization, whilst diminished values enhance sparsity (Lasso) or decrease coefficient values (Ridge).

Lasso exhibits optimal performance at C ≈ 10, suggesting that a moderate degree of sparsity enhances detection.

Ridge achieves optimal accuracy at C ≈ 20, indicating that keeping all features with mild decreasing is desirable.

Elevated C values (>40) result in overfitting, whereas significantly low values (C < 1) decrease performance due to excessive regularization (Table 4).

Table 4. Sensitivity to Regularization Strength for Lasso and Ridge Regression.

Regularization Type	Optimal C Value	Effect of Increasing C	Effect of Decreasing C
Lasso	≈ 10	Reduces sparsity, may overfit (>40)	Increase sparsity, may underfit (<1)
Ridge	≈ 20	May overfit (>40)	Excessive regularization, lower performance (<1)

4.2. Sensitivity to Train-Test Split Ratio

The dataset was evaluated using several train-test splits to investigate how data availability influences model performance.

With an 80/20 or 90/10 split, the best accuracy is attained; hence, a bigger training set suggests better model performance.

When the training set is too small—that is, 60/40 split—that means the model finds it difficult to generalize from little data and causes recall reductions.

In Lasso specifically, larger test sets (30–40%) raise overfitting risk (Table 5).

Table 5. Sensitivity to Train-Test Split Ratio.

Train-Test Split Ratio	Effect on Model Performance
80/20 or 90/10	Best accuracy, better generalization
60/40	Difficult to generalize, recall reduction, higher risk of overfitting
Larger Test Sets (30-40%)	Increased overfitting risk

4.3. Sensitivity to Synthetic Data Ratio (GAN Augmentation)

Various ratios of actual to synthetic data were investigated to evaluate GAN-generated synthetic data. With 30% synthetic data, the best performance occurs where Ridge gets 89.0% accuracy. Beyond 40% synthetic data, performance suffers somewhat; this suggests that too much synthetic data could cause noise. Between 20 and 30%, the ideal synthetic-real data ratio increases recall while keeping stability (Table 6).

Table 6. Sensitivity to Synthetic Data Ratio (GAN Augmentation).

Synthetic Data Ratio	Accuracy (Ridge)	Observation
20-30%	≈89%	Optimal balance; increases recall while maintaining stability
30%	89%	Best Performance
>40%	Decreases	Performance declines due to noise

4.4. Sensitivity to Feature Scaling

Unscaled data produces low accuracy since feature magnitudes control model decisions. Standardizing improves recall and helps to maintain weights, thereby optimizing performance.

4.5. Sensitivity to GAN Model

We examined models trained with real data only vs. real + generated data using GAN. GAN augmentation enhances model performance across all measures. Recall increases dramatically, thus fewer anomalies are missed. GAN-generated data boosts overall model robustness. The choice of optimizer impacts both GAN training stability and final model accuracy. Adam is the best optimizer, providing the steadiest training and highest-quality synthetic data. Stochastic gradient descent (SGD) performs poorly, with slower convergence and worse anomaly detection accuracy. RMSprop is a middle-ground solution but slightly less stable than Adam.

5. Conclusion and Recommendation for Future Research

Anomaly detection remains a crucial area in data mining and machine learning due to data complexity and variability. While real data is critical for training models, data volume and quality constraints frequently dictate alternate approaches. This study used GANs to generate synthetic textual data and merged it with real data, thereby improving the accuracy of anomaly detection algorithms. A fundamental feature of this research was the use of GANs to generate high-quality synthetic textual data, filling a gap in previous studies that mostly focus on non-textual data, such as photos or sounds. By producing realistic synthetic text, the model enhanced its capacity to detect complicated anomalous patterns, especially in social network data where language diversity and complexity are key challenges. Another novelty of this study was the integration of actual and synthetic textual data, which showed that this combination improves the accuracy of Lasso and Ridge regression models in anomaly identification, particularly when real data is restricted. The findings showed that synthetic textual data improves model performance for detecting anomalies in social network texts. Future studies could look into WGANs with gradient penalty to improve text generation stability and quality. WGANs have been commonly used for image synthesis and, as shown by Abhari [27], for creating artificial network traffic and frame arrival-rate distributions to enhance simulation accuracy in edge/fog environments. Likewise, using WGANs for generating synthetic text could improve limited or uneven textual datasets, boost anomaly detection performance, and lower the cost and effort of gathering real-world social network text. Furthermore, applying Transfer Learning techniques may speed up training while minimizing data requirements. Optimizing computational efficiency and lowering training time are also important areas for advancement. Furthermore, using models like GPT-4 for synthetic text generation may improve linguistic accuracy and structural coherence, resulting in more precise anomaly detection in textual data.

References

[1] C. C. Aggarwal, "Outlier ensembles: position paper," *ACM SIGKDD Explorations Newsletter*, vol. 14, no. 2, pp. 49–58, 2013.

[2] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[3] V. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, pp. 85–126, 2004.

[4] J. Mao, Y. Yang, and T. Zhang, "Empirical analysis of attribute inference techniques in online social network," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 881–893, 2020.

[5] D. Tiwari et al., "A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques," *Artificial Intelligence Review*, vol. 56, no. 11, pp. 13407–13461, 2023.

[6] C. Ahmed, A. ElKorany, and E. ElSayed, "Prediction of customer's perception in social networks by integrating sentiment analysis and machine learning," *Journal of Intelligent Information Systems*, vol. 60, no. 3, pp. 829–851, 2023.

[7] S. Shan, Y. Yang, and C. Li, "Which Receives More Attention, Online Review Sentiment or Online Review Rating? Spillover Effect Analysis from JD.com," *Behavioral Sciences*, vol. 14, no. 9, p. 823, 2024.

[8] M. M. Breunig, H. P. Kriegel, R. T. Ng, and J. Sander, "LOF: identifying density-based local outliers," *in Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*. *New York: ACM*, 2000, pp. 93–104.

[9] S. Wette and F. Heinrichs, "OML-AD: Online Machine Learning for Anomaly Detection in Time Series Data," *arXiv preprint, arXiv:2409.09742, 2024. [Online]*. Available: https://arxiv.org/abs/2409.09742. [Accessed: Jun. 1, 2024].

[10] R. Domingues, M. Filippone, P. Michiardi, and J. Zouaoui, "A comparative evaluation of outlier detection algorithms: Experiments and analyses," *Pattern Recognition*, vol. 74, pp. 406–421, 2018.

[11] E. Manzoor, H. Lamba, and L. Akoglu, "xstream: Outlier detection in feature-evolving data streams," *in Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM*, 2018, pp. 1963–1972.

[12] É. Courtois, P. Tubert-Bitter, and I. Ahmed, "New adaptive lasso approaches for variable selection in automated pharmacovigilance signal detection," *BMC Medical Research Methodology*, vol. 21, pp. 1–17, 2021.

[13] E. S. Groenewald et al., "Real-Time Anomaly Detection in Industrial Systems Using Stream Processing and Online Machine Learning," *in International Conference on Deep Learning and Visual Artificial Intelligence. Singapore*: Springer, 2024, pp. 513–529.

[14] D. P. Senapati et al., "DG-GAN: A Deep Neural Network for Real-World Anomaly Detection in Surveillance Videos," *in International Conference on Advanced Network Technologies and Intelligent Computing. Cham*: Springer, 2023, pp. 93–106.

[15] Y. Su et al., "Robust anomaly detection for multivariate time series through stochastic recurrent neural network," *in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM*, 2019, pp. 2828–2837.

[16] P. Perera, R. Nallapati, and B. Xiang, "Ocgan: One-class novelty detection using gans with constrained latent representations," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE*, 2019, pp. 2898–2906.

[17] T. Schlegl et al., "f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks," *Medical Image Analysis*, vol. 54, pp. 30–44, 2019.

[18] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "Ganomaly: Semi-supervised anomaly detection via adversarial training," *in Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Revised Selected Papers, Part III. Cham: Springer*, 2019, pp. 622–637.

[19] F. H. K. D. S. Tanaka and C. Aranha, "Data augmentation using GANs," *arXiv preprint, arXiv:1904.09135, 2019. [Online]*. Available: https://arxiv.org/abs/1904.09135. [Accessed: Oct. 10, 2023].

[20] S. Chen, M. Peng, H. Xiong, and S. Wu, "An anomaly detection method based on Lasso," *Cluster Computing*, vol. 22, pp. 5407–5419, 2019.

[21] N. K. Singh and K. Raza, "Medical image generation using generative adversarial networks: A review," *in Health Informatics: A Computational Perspective in Healthcare. Singapore: Springer*, 2019, pp. 77–96.

[22] M. S. Meor Yahaya and J. Teo, "Data augmentation using generative adversarial networks for images and biomarkers in medicine and neuroscience," *Frontiers in Applied Mathematics and Statistics*, vol. 9, p. 1162760, 2023.

[23] J. Naseri, H. Hasanpour, and A. Ghanbari Sorkhi, "Accelerating Legislation Processes through Semantic Similarity Analysis with BERT-based Deep Learning," *International Journal of Engineering, Transactions C: Aspects*, vol. 37, no. 6, pp. 1050–1058, 2024.

[24] A. Yavari and H. Hasanpour, "Election Prediction Based on Messages Feature Analysis in Twitter Social Network," *International Journal of Engineering, Transactions C: Aspects*, vol. 36, no. 6, pp. 1179–1184, 2023.

[25] M. Ahangari and A. Sebti, "A Hybrid Approach to Sentiment Analysis of Iranian Stock Market User's Opinions," *International Journal of Engineering, Transactions C: Aspects*, vol. 36, no. 3, pp. 573–584, 2023.

[26] Hastie, T., Tibshirani, R., & Friedman, J., *The elements of statistical learning*. Springer, 2009.

[27] A. Abhari, "Towards New Simulation Models for DL-Based Video Streaming in Edge Networks, " *Proceedings of the 2024 Winter Simulation Conference (WSC)*. Los Alamitos, CA: IEEE, pp. 656–665, 2024.

[28] A. Mohammadi Gohar, K. Rahbar, B. Minaei-Bidgoli, and Z. Beheshtifard, "Study on Generative Adversarial Network in Discrete Data: A Survey," *Journal of Artificial Intelligence and Data Mining*, vol. 12, no. 4, pp. 567–581, 2024.

Appendix A: Pseudocode of the Proposed GAN-Based Anomaly Detection Algorithm

Step 1: Load Required Libraries

Import necessary libraries

pandas and numpy for data manipulation **-**

- keras for GAN model

sklearn for classification models -

- matplotlib for visualization

Step 2: Load and preprocess data

Read CSV file (nk.csv) into DataFrame. Extract 'comment' column as text data. Extract 'recommend' column as labels (recommended / not_recommended).

Split data into two DataFrames

df١ → recommended comments -

df٢ → not recommended comments -

Clean text data

.Remove non-Persian characters using regex -

.Tokenize and convert to sequences -

- Pad sequences to maximum length.

Step 3: Define and Train GAN Model

Define Generator

Input: random noise (١٠٠-dimensional) -

Dense layers with LeakyReLU and Dropout -

LSTM layers for sequence generation -

Output: softmax activation (word - probabilities)

Define Discriminator

Input: real or fake sequences -

- Dense layers with LeakyReLU and Dropout

Output: sigmoid activation (binary - classification)

Compile GAN

- Discriminator is frozen while training GAN .

- Generator output passes through Discriminator .

Train GAN

.Generate fake samples from noise -

Train the discriminator with real and fake - .samples

- Train Generator using GAN model .
- Track Discriminator and Generator losses .

- Plot the loss curves.

Step 4: Generate Synthetic Comments

Use the trained generator to create 1,500 new comments for each class. Apply temperature-based sampling to select words. Save the generated comments as CSV files.

Step 5: Merge and Clean Data

Load original dataset and generated dataset. Remove stop words using a predefined stop-word list. Count word frequency in each category. Select words appearing at least 10 times.

Step 6: Feature Extraction

Create a vocabulary from the unique words in both classes. Convert each comment into a feature vector (word counts). Normalize the feature vectors using standard scaling. Encode labels with a label encoder.

Step 7: Train and Evaluate Classification Models

Split the data into training (80%) and testing (20%).

Train two classifiers:

.Logistic regression with L١ (Lasso) penalty -

.Logistic regression with L٢ (Ridge) penalty -

Evaluate model

.Compute the confusion matrix -

Calculate accuracy for both recommended - .and not recommended comments

- Print and compare the results.

رویکردی مبتنی بر GAN برای شناسایی ناهنجاری در داده‌های متنی شبکه‌های اجتماعی با استفاده از مدل‌های رگرسیون لاسو و ریج

ابوالفضل آدرسی و امیرحسین امیری*

گروه مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه شاهد، تهران، ایران .

ارسال ۲۰۲۵/۰۶/۰۲؛ بازنگری ۲۰۲۵/۰۹/۰۹؛ پذیرش ۲۰۲۵/۱۲/۱۹

چکیده:

شناسایی و طبقه‌بندی ناهنجاری‌ها در داده‌های متنی استخراج‌شده از شبکه‌های اجتماعی به دلیل پیچیدگی‌های زبانی و تنوع شیوه‌های بیان کاربران، چالشی اساسی محسوب می شود. اگرچه روش‌های یادگیری عمیق و یادگیری ماشین ظرفیت بالایی برای مقابله با این چالش دارند، اما کارایی آن‌ها اغلب به‌دلیل کمبود داده‌های آموزشی محدود می‌شود. این مقاله تأثیر شبکه‌های مولد تخاصمی (GAN) را بر فرآیند شناسایی و طبقه‌بندی ناهنجاری‌ها بررسی کرده و همچنین نقش آن‌ها را در تولید داده‌های متنی مصنوعی مورد ارزیابی قرار می‌دهد. ترکیب داده‌های مصنوعی تولیدشده با داده‌های واقعی منجر به بهبود قابل‌توجه صحت طبقه‌بندی، به‌ویژه در شرایط کمبود داده، می‌شود. در این پژوهش، از روش‌های رگرسیون لاسو (Lasso) و ریج (Ridge) برای شناسایی و طبقه‌بندی ناهنجاری‌ها استفاده شده است. نتایج تجربی نشان می‌دهد که مدل پیشنهادی در شناسایی و طبقه‌بندی ناهنجاری‌ها در مجموعه‌داده‌های جدید تولیدشده توسط روش GAN عملکرد بهتری دارد. با تلفیق روش‌های آماری و تکنیک‌های مولد، راهکار ارائه‌شده از تفسیرپذیری و مقیاس‌پذیری بالاتری برخوردار بوده و برای تحلیل پیشرفته متون در محیط‌های پویایی همچون شبکه‌های اجتماعی، مناسب‌تر است.

کلمات کلیدی: شبکه‌های اجتماعی، طبقه‌بندی ناهنجاری، شبکه‌های مولد تخاصمی، یادگیری ماشین، رگرسیون لاسو و ریج.