

Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

Aspect-Based Deep Learning Model for Spam Review Detection

Shaik HafsaFarheen¹, Dr. Bushra Tahseen²

¹M.Tech Student, Dept.of CSE, Dr.K.V.Subba Reddy Institute of Technology, Kurnool, Andhra Pradesh

²Associate Professor, Dept.of CSE, Dr.K.V.Subba Reddy Institute of Technology, Kurnool, Andhra Pradesh

Abstract

The rapid growth of e-commerce platforms has led to an increasing number of online reviews that significantly influence customer purchasing decisions. However, the presence of deceptive or fake reviews undermines the reliability of these platforms and misleads consumers. This project proposes a system titled “Aspect-Based Deep Learning Model for Spam Review Detection,” which focuses on automatically identifying fraudulent reviews using machine learning and text analysis techniques. The proposed system processes customer review text by applying natural language processing (NLP) methods for feature extraction and classification. A TF-IDF (Term Frequency– Inverse Document Frequency) vectorization approach is used to convert textual reviews into numerical feature representations. These features are then utilized by a Logistic Regression classifier to distinguish between authentic and fake reviews. The model is trained on a labeled dataset of reviews and evaluated using metrics such as accuracy, confusion matrix, and ROC curve to measure performance.

To enhance usability, the system is implemented as a web-based application using the Flask framework, enabling users to register, log in, and submit product reviews for authenticity prediction. The system also includes an administrative interface that manages users, stores prediction history, and visualizes training statistics such as accuracy and loss across training epochs. Model training results and evaluation graphs are automatically generated and stored for analysis.

The proposed approach provides an efficient and scalable solution for detecting deceptive reviews in e-commerce environments. By combining machine learning-based text classification with a user-friendly web interface, the system helps improve the credibility of online reviews and assists consumers in making more informed purchasing decisions.

Keywords: review, vectorization, Model, Spam, TF-IDF

Introduction

With the rapid expansion of e-commerce platforms, online reviews have become one of the most influential factors affecting consumer purchasing decisions. Customers often rely on product reviews to evaluate product quality, reliability, and overall satisfaction before making purchases. However, the increasing popularity of online review systems has also led to the emergence of fake or deceptive reviews posted by individuals or organizations to manipulate product reputation and mislead potential buyers.

Fake reviews can significantly impact both consumers and businesses. Misleading positive reviews may promote low-quality products, while malicious negative reviews can damage the reputation of genuine products or sellers. Therefore, detecting and filtering such deceptive reviews has become a critical challenge for maintaining trust and credibility in e-commerce platforms.

To address this issue, machine learning and natural language processing techniques have been widely used to analyze review content and identify suspicious patterns. This project proposes a system titled Aspect-Based Deep Learning Model for Spam Review Detection. The system analyzes



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

review text using text preprocessing, feature extraction, and classification techniques to determine whether a review is authentic or fake.

The proposed system uses TF-IDF vectorization to transform textual data into numerical representations and applies a Logistic Regression classifier to categorize reviews into authentic or fake classes. The model is trained using a labeled dataset and evaluated through performance metrics such as accuracy, confusion matrix, and ROC curve. Additionally, the system is implemented as a web-based application using Flask, allowing users to register, log in, and check the authenticity of reviews in real time.

This approach aims to enhance the reliability of online review systems by providing an efficient and scalable mechanism for detecting deceptive reviews.

Problem Statement

Online shopping platforms rely heavily on user-generated reviews to help customers evaluate products. However, many reviews are intentionally manipulated by sellers, competitors, or automated bots to influence customer decisions. These fake reviews reduce the credibility of e-commerce platforms and lead to poor customer experiences.

Traditional manual moderation methods are inefficient and unable to handle the large volume of reviews generated daily. Therefore, there is a need for an automated system that can accurately detect and classify fake reviews using machine learning techniques. The main problem addressed in this project is:

To design and develop an intelligent system that can automatically analyze review text and accurately detect whether a review is genuine or fake, thereby improving the reliability and trustworthiness of e-commerce review platforms.

Objectives

The main objectives of this project are:

1. To analyze online review data and identify patterns that indicate fake or deceptive reviews.
2. To implement Natural Language Processing (NLP) techniques for preprocessing and feature extraction from textual reviews.
3. To develop a machine learning model capable of classifying reviews as authentic or fake.
4. To use TF-IDF vectorization for converting textual review data into numerical features suitable for classification.
5. To train and evaluate the model using performance metrics such as accuracy, confusion matrix, and ROC curve.
6. To develop a web-based application using Flask that allows users to submit reviews and obtain authenticity predictions.
7. To provide an administrative interface for managing users, viewing predictions, and monitoring training statistics.
8. To improve trust and reliability in e-commerce platforms by detecting and reducing fake reviews.

Existing System

In current e-commerce platforms, product reviews are mostly monitored using manual moderation or simple rule-based filtering methods. Many platforms allow users to freely post reviews without strict verification, which makes it easier for fake or deceptive reviews to appear in the system. Some



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

existing systems use basic keyword filtering or user reporting mechanisms to identify suspicious reviews.

A few platforms apply traditional machine learning approaches to detect fake reviews, but these methods often rely on limited features such as reviewer behavior or simple text patterns. Additionally, manual review processes require significant human effort and are not scalable for large volumes of reviews generated daily.

Due to the rapid growth of online shopping platforms and increasing numbers of user-generated reviews, these traditional approaches struggle to maintain accuracy and efficiency in identifying fake reviews.

Problems with Existing System

The existing systems for detecting fake reviews face several limitations: in real time.

Manual Moderation

Many platforms rely on manual review processes which are time-consuming and inefficient

Low Accuracy

Traditional rule-based methods cannot effectively detect complex patterns in deceptive reviews.

Scalability Issues

Handling large volumes of reviews generated daily becomes difficult using manual or simple automated systems.

Limited Feature Analysis

Existing systems often analyze only basic review features and ignore deeper linguistic patterns.

Delayed Detection

Fake reviews may remain visible for a long time before being identified and removed.

Lack of Automation

Many platforms do not have a fully automated intelligent system for detecting fake reviews

Proposed System

To overcome the limitations of existing systems, this project proposes a system titled “Aspect-Based Deep Learning Model for Spam Review Detection.”

The proposed system uses machine learning and natural language processing techniques to automatically detect fake reviews. The system processes review text and converts it into numerical features using TF-IDF (Term Frequency–Inverse Document Frequency) vectorization. These features are then used by a Logistic Regression classifier to classify reviews as authentic or fake.

The system is implemented as a web-based application using the Flask framework, allowing users to register, log in, and submit reviews for authenticity prediction. The trained model analyzes the review text and provides instant results.

The application also includes:

A database system to store user details and prediction history.

Model training and evaluation modules that generate accuracy graphs, confusion matrices, and ROC curves.

An admin dashboard to monitor system activity and review results.

This automated approach provides an efficient and scalable solution for identifying deceptive reviews in e-commerce platforms.

Advantages of Proposed System

The proposed system offers several advantages over traditional methods:

Automated Fake Review Detection



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

The system automatically analyzes reviews and predicts authenticity without manual intervention.

Improved Accuracy

Machine learning algorithms provide better detection accuracy compared to rule-based systems.

Efficient Text Analysis

TF-IDF feature extraction helps capture important textual patterns in reviews.

Real-Time Prediction

Users can instantly check whether a review is authentic or fake.

User-Friendly Web Interface

The Flask-based web application allows easy interaction for users and administrators.

Scalable System

The system can handle large numbers of reviews efficiently.

Performance Monitoring

Training accuracy, confusion matrix, and ROC curve graphs help evaluate the model's performance.

Data Storage and Management

The database maintains records of users, predictions, and training statistics for future reference.

Literature Survey

The rapid growth of e-commerce platforms has significantly increased the importance of online consumer reviews in influencing purchasing decisions. Customers rely heavily on product reviews to evaluate the quality, reliability, and performance of products before making a purchase. However, the increasing popularity of review systems has also resulted in the widespread problem of fake or deceptive reviews. These reviews are intentionally created to manipulate consumer opinions by promoting or demoting products unfairly. As a result, researchers have focused extensively on developing techniques to identify and filter fake reviews using machine learning, natural language processing, and deep learning approaches.

Early research in fake review detection mainly relied on rule-based and statistical methods. These methods used predefined rules, keyword filtering, and basic linguistic patterns to identify suspicious reviews. Although these approaches were simple to implement, they lacked the ability to detect complex patterns in deceptive reviews. Fake reviews often mimic genuine writing styles, making them difficult to identify using simple rules. Therefore, these traditional methods were found to be insufficient when dealing with large-scale e-commerce platforms that generate thousands of reviews daily.

With the advancement of machine learning techniques, researchers began applying supervised learning algorithms to detect deceptive reviews. Algorithms such as Naïve Bayes, Support Vector Machines (SVM), and Logistic Regression were widely used for classification tasks. These models used textual features extracted from reviews, including word frequency, n-grams, and syntactic patterns. Feature extraction methods such as Term Frequency–Inverse Document Frequency (TF-IDF) helped convert textual data into numerical vectors suitable for machine learning algorithms. Although these approaches improved detection accuracy compared to rule-based methods, they still relied heavily on manually engineered features and often failed to capture deeper semantic relationships within the text.



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

As the complexity of deceptive review patterns increased, researchers started exploring Natural Language Processing (NLP) techniques to analyze linguistic structures and contextual meaning in review content. NLP methods enabled systems to perform tasks such as tokenization, stop-word removal, stemming, and sentiment analysis. These techniques helped identify emotional tone, writing style, and contextual cues that might indicate deceptive behavior. However, traditional NLP approaches still depended on handcrafted features and were limited in their ability to capture complex dependencies between words in long review texts.

Despite the advantages of deep learning, many early deep learning models focused only on the overall review text and did not consider the multiple aspects or attributes mentioned within reviews. In reality, product reviews often discuss different aspects such as price, quality, durability, delivery service, and customer experience. Each aspect may contain useful information for detecting whether a review is genuine or deceptive. Ignoring these aspects can lead to incomplete understanding of review content and reduce the effectiveness of detection systems.

To address this limitation, researchers introduced aspect-based sentiment analysis (ABSA) techniques. Aspect-based approaches aim to identify specific product attributes mentioned in reviews and analyze the sentiment associated with each aspect. By focusing on aspect-level information, these methods provide a more detailed understanding of review content. Aspect-based techniques can detect unusual sentiment patterns across different aspects, which may indicate deceptive behavior. For example, fake reviews may contain exaggerated praise for certain aspects while ignoring others.

The research paper titled “Advancing E-Commerce Authenticity: A Novel Fusion Approach Based on Deep Learning and Aspect Features for Detecting False Reviews” proposes an advanced framework that integrates deep learning models with aspect-based feature extraction. The proposed approach introduces an Aspect Fusion Network that combines multiple sources of information to improve the detection of deceptive reviews. The key idea behind this approach is to capture both the contextual information in review text and the importance of specific product aspects.

The proposed system utilizes a hierarchical attention mechanism to identify the most important words and aspects in a review. Attention mechanisms are widely used in deep learning models to assign different weights to words based on their importance in the classification task. In the context of fake review detection, the attention mechanism helps the model focus on words that are more likely to reveal deceptive patterns. By combining attention with aspect information, the system can better understand which aspects of a review contribute to its authenticity or deception.

Another important component of the proposed framework is the aspect dictionary, which contains predefined lists of aspect-related keywords. These dictionaries help the model identify relevant product attributes mentioned in reviews. The system then fuses these aspect features with deep learning representations to create a more comprehensive feature set. This fusion process allows the model to analyze both semantic and aspect-level information simultaneously.

The research also introduces a cardinality fusion model, which is designed to improve the stability and performance of the auxiliary learning components. In many deep learning systems, random initialization of weights can negatively affect the performance of auxiliary models. The cardinality fusion approach helps mitigate this problem by combining multiple feature representations and reducing the influence of random variations during training. This results in a more stable and robust model for detecting fake reviews.

Experimental evaluations conducted in the study demonstrate that the proposed fusion model significantly improves detection accuracy compared to traditional machine learning and basic deep learning approaches. The integration of aspect-based features with deep learning representations

Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

enables the model to capture complex patterns in review content more effectively. As a result, the system achieves better performance in identifying deceptive reviews across multiple datasets. The study also highlights the importance of combining multiple analytical perspectives when dealing with fake review detection. Purely textual analysis may overlook important contextual signals, while aspect-based analysis alone may fail to capture deeper semantic relationships. By integrating both approaches, the proposed fusion model creates a more comprehensive understanding of review data. This hybrid approach improves the system's ability to distinguish between genuine user experiences and artificially generated content. Furthermore, the research emphasizes the practical importance of developing automated fake review detection systems for real-world e-commerce applications. Online

marketplaces such as Amazon, eBay, and other retail platforms rely heavily on customer reviews to maintain transparency and trust among users. Effective detection systems can help platforms maintain credibility by filtering out misleading reviews and protecting consumers from fraudulent information. Although the proposed approach demonstrates promising results, the study also acknowledges certain challenges and future research opportunities. Fake review generation techniques continue to evolve, making it necessary to develop more advanced detection models. Future research may explore the use of transformer-based models, large language models, and multimodal data analysis to further improve detection accuracy. Additionally, incorporating behavioral data such as reviewer activity patterns may enhance the ability to identify suspicious users.

In conclusion, the literature indicates that fake review detection has evolved from simple rule-based systems to advanced deep learning frameworks that integrate multiple analytical techniques. The fusion of deep learning and aspect-based features represents a promising direction for improving the accuracy and robustness of detection systems. The proposed aspect fusion network demonstrates how combining hierarchical attention mechanisms, aspect dictionaries, and fusion models can effectively identify deceptive reviews in e-commerce platforms. These advancements contribute to building more trustworthy online marketplaces and improving the overall reliability of consumer review.

System Architecture

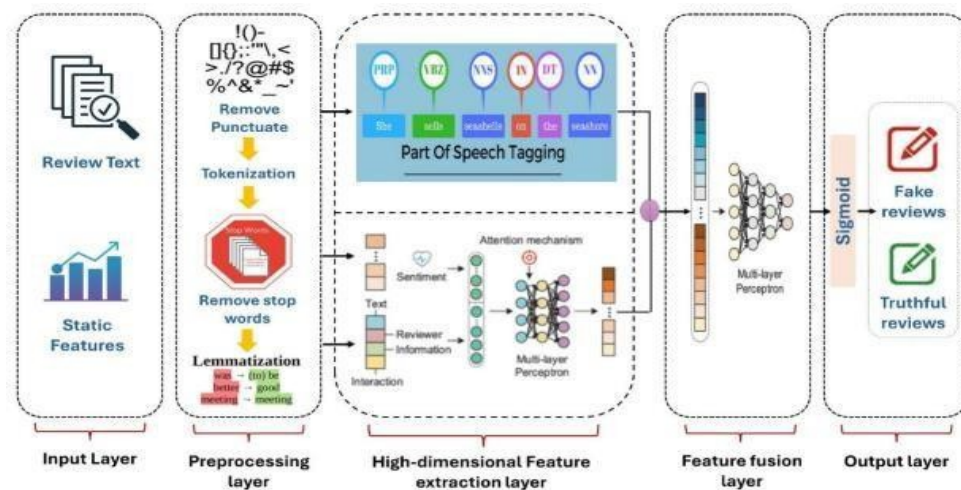


Figure: System Architecture



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

The model is divided into five sequential layers, as illustrated in the provided diagram:

1. Input Layer

Review Text: The raw unstructured textual content submitted by users.

Static Features: Numerical or categorical data such as user ratings and metadata.

2. Preprocessing Layer

This stage refines the raw data for analysis through four main steps: **Remove Punctuation:** Eliminating symbols to focus on word content. **Tokenization:** Splitting sentences into individual words or tokens.

Remove Stop Words: Filtering out common language words (e.g., "the", "and") that do not carry significant sentiment.

Lemmatization: Reducing words to their base or dictionary form (e.g., "better" to "good").

3. High-Dimensional Feature Extraction Layer

The system extracts two parallel sets of features:

Part-Of-Speech (POS) Tagging: Categorizing words (Nouns, Verbs, Adjectives) to understand the grammatical structure.

Deep Feature Extraction: Utilizing an Attention Mechanism and a Multi-layer Perceptron (MLP) to process sentiments, reviewer information, and textual interactions. This focuses the model on the most informative parts of the review.

4. Feature Fusion Layer

This layer acts as the "Fusion Network," where deep-learned patterns are synergized with linguistic aspect features.

It utilizes another MLP to combine these disparate data streams into a single, robust feature vector for classification.

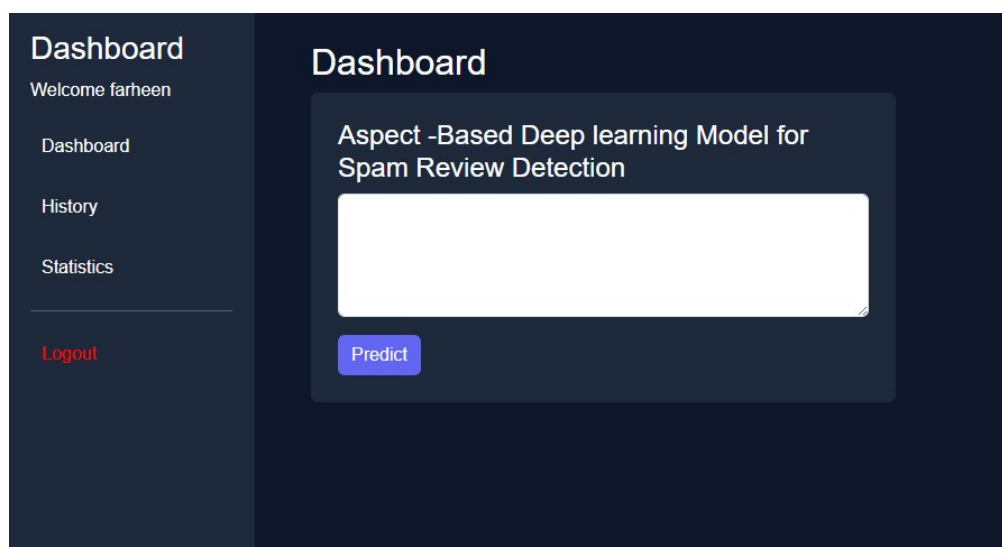
5. Output Layer

Sigmoid Function: A mathematical activation function used to map the final features to a probability score.

Classification: The system makes a final decision, categorizing the review as either Fake or Truthful.

Results

Pridiction Review:





Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

The Review Prediction page allows users to check the authenticity of a product review. Users can enter the review text in the provided input box and click the Predict button. The system analyzes the text using the trained machine learning model to determine whether the review is genuine or fake. A Back to Dashboard option is also available for easy navigation.

Prediction Result:

This screen allows users to analyze whether a product review is genuine or fake using the trained Hybrid Deep–Aspect Fusion Network model. Users can enter the review text in the input field and click the Predict button to perform the analysis. The system processes the text and displays the prediction result indicating whether the review is Fake or Genuine. A Back to Dashboard link is provided to return to the main interface.

Conclusion

The rapid growth of e-commerce platforms has significantly increased the importance of online product reviews in influencing customer purchasing decisions. However, the presence of fake or deceptive reviews poses a serious challenge to maintaining the reliability and credibility of online review systems. Detecting such misleading reviews has therefore become an essential task for ensuring transparency and trust in e-commerce environments.

This project, titled “Hybrid Deep–Aspect Fusion Network for Robust Fake Review Detection in E-Commerce,” focuses on developing an automated system capable of identifying fake reviews using machine learning and natural language processing techniques. The system processes review text through several stages including preprocessing, feature extraction, model training, and prediction. The use of TF-IDF vectorization helps convert textual reviews into meaningful numerical representations, while the Logistic Regression classifier effectively distinguishes between genuine and fake reviews.

A web-based application was developed using the Flask framework, allowing users to register, log in, and submit reviews for authenticity verification. The trained model analyzes the review text and provides predictions in real time. Additionally, the system generates performance evaluation metrics such as accuracy graphs, confusion matrices, and ROC curves to assess the effectiveness of the model.

The experimental results demonstrate that the proposed system is capable of detecting deceptive reviews with good accuracy. By automating the review verification process, the system reduces manual effort and improves the reliability of online review platforms. Overall, the proposed solution provides an efficient and scalable approach for detecting fake reviews and contributes to enhancing trust in e-commerce ecosystems.



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005
p-ISSN: 3117-4418

Volume: 02 Issue: 04, April, 2026
e-ISSN: 3117-4426

Future Work

Although the proposed system demonstrates promising results in detecting fake reviews, there are several opportunities for further improvements and enhancements. One potential direction for future work is the integration of advanced deep learning models such as Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Transformer-based architectures. These models can capture deeper semantic relationships and contextual dependencies in review text, which may further improve detection accuracy.

Another improvement could involve the incorporation of aspect-based sentiment analysis to analyze specific product features mentioned in reviews, such as price, quality, and delivery. This approach can provide more detailed insights into review authenticity by examining how users express opinions about different product aspects. Future systems may also integrate user behavioral analysis, including reviewer activity patterns, posting frequency, and rating behavior. Combining textual analysis with behavioral data can enhance the ability to detect suspicious review patterns. Additionally, the system can be extended to support multilingual review detection, allowing the detection of fake reviews written in different languages. This feature would make the system more applicable to global e-commerce platforms. Another possible improvement is the deployment of the model in large-scale cloud environments, enabling real-time analysis of massive volumes of reviews generated across multiple online marketplaces. Overall, future research can focus on combining advanced machine learning techniques, behavioral analysis, and large-scale deployment to build more intelligent and robust fake review detection systems.

References

1. Jindal, N., & Liu, B., "Opinion Spam and Analysis," Proceedings of the International Conference on Web Search and Data Mining, 2008.
2. Ott, M., Choi, Y., Cardie, C., & Hancock, J., "Finding Deceptive Opinion Spam by Any Stretch of the Imagination," Proceedings of the Annual Meeting of the Association for Computational Linguistics, 2011.
3. Mukherjee, A., Liu, B., & Glance, N., "Spotting Fake Reviewer Groups in Consumer Reviews," Proceedings of the World Wide Web Conference, 2012.
4. Li, F., Huang, M., Yang, Y., & Zhu, X., "Learning to Identify Review Spam," Proceedings of the International Joint Conference on Artificial Intelligence, 2011.
5. Rayana, S., & Akoglu, L., "Collective Opinion Spam Detection Using Active Inference," Proceedings of the IEEE International Conference on Data Mining, 2015.
6. Ren, Y., Ji, D., & Zhang, H., "Positive Unlabeled Learning for Deceptive Review Detection," Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2014.
7. Feng, S., Banerjee, R., & Choi, Y., "Syntactic Stylometry for Deception Detection," Proceedings of the Association for Computational Linguistics, 2012.
8. Li, J., Ott, M., Cardie, C., & Hovy, E., "Towards a General Rule for Identifying Deceptive Opinion Spam," Proceedings of the Annual Meeting of the Association for Computational Linguistics, 2014.
9. Wang, G., Xie, S., Liu, B., & Yu, P., "Review Graph Based Online Store Review Spam Detection," Proceedings of the IEEE International Conference on Data Mining, 2011.
10. Zhang, Y., Lauw, H. W., & Lim, E. P., "Spam Detection in Online Social Networks," Proceedings of the International Conference on Web Intelligence, 2010.
11. Kim, Y., "Convolutional Neural Networks for Sentence Classification," Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2014.



Research Digest on Engineering Management and Social Innovations

Available on: <http://rdems.in>

DOI: 10.46647/rdems0204005

Volume: 02 Issue: 04, April, 2026

p-ISSN: 3117-4418

e-ISSN: 3117-4426

12. Tang, D., Qin, B., & Liu, T., “Aspect-Level Sentiment Classification with Deep Memory Network,” Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2016.