International Journal of Computer Science and Engineering Innovations
ISSN: 3107-9458 | Volume 2 Issue 1, 17-24, January-March 2026
DOI: https://doi.org/10.64137/31079458/IJCSEI-V2I1P103

Scholastic Research Publication

Received: 06/01/2026 Revised: 11/01/2026 Accepted: 22/01/2026 Published: 24/01/2026

Original Article

Adversarial Attacks and Robustness in SMS Spam Classification Models

SAMPATH KUMAR N
Senior Solution Delivery Lead, Deloitte USI Pvt Ltd, USA.

ABSTRACT: *Short Message Service (SMS) spam classification systems based on machine learning and deep learning have achieved high accuracy under standard evaluation settings; however, their vulnerability to adversarial attacks poses significant risks to real-world deployment. Attacks on SMS spam classifiers are realized as precision-level adversarial examples, where carefully crafted perturbations (e.g., character substitution/reordering, word disguising and synonym injection, benign token insertion) maintain human readability while drawing models to misclassify. In this paper, we study universal adversarial attacks on SMS spam clasification models and their adaptation to various feature representations and learning architectures: from classic machine learning models to deep neural networks and transformer-based solutions. And it introduces defense terms like adversarial training, data augmentation, ensemble learning, input sanitization and defensive distillation. Primary challenges such as trade-offs between robustness and accuracy, language diversity, and limited computation in mobile contexts are considered. The findings of the paper show possible future research challenges worth addressing for building robust SMS spam detection systems that are still performing effectively under adversarial conditions.*

KEYWORDS: *SMS spam classification, Adversarial attacks, Model robustness, Natural language processing, Adversarial training, Text perturbation, Spam detection security.*

1. INTRODUCTION

1.1. BACKGROUND OF SMS SPAM CLASSIFICATION

Short Message Service (SMS) spam classification is a critical application of Natural Language Processing (NLP) aimed at automatically distinguishing between legitimate (ham) and unsolicited or malicious (spam) text messages. The rapid growth of mobile communication and digital services has made SMS a common channel for advertising, phishing, fraud, and social engineering attacks. Due to the high volume and informal nature of SMS messages—characterized by short length, abbreviations, misspellings, and noisy text—manual filtering is impractical, necessitating automated classification approaches.

1.2. GROWING RELIANCE ON ML/DL MODELS FOR SPAM DETECTION

The dominance of machine learning (ML) and deep learning (DL) methods in SMS spam detection is because they can learn complex mappings from the data. Classic ML models, like Naïve Bayes, Support Vector Machines and Logistic Regression, operate with hand-crafted features; while modern DL models – like recurrent neural networks, convolutional neural networks and new transformer-based architectures – use distributed representations to obtain better classification accuracy. Such models are widely used in mobile devices and telecommunications, so they play critical roles in real-time anti-spam systems.

1.3. MOTIVATION FOR STUDYING ADVERSARIAL ATTACKS IN SMS-BASED SYSTEMS

ML- and DL-based SMS spam classifiers seem to be working well; however, they are susceptible to adversarial attacks. Attackers may intentionally modify message users' text through small perturbations (e.g., character-level scrambling, word replacement or token changing) to avoid detection but still retain semantic meaning for human understanding. In SMS-based models that are particularly sensitive to marginal textual changes, such vulnerabilities cast substantial doubt on the reliability, security and trustworthiness of the user in question. Consequently, being able to learn from (or about) adversarial threats is crucial for evaluating the real-world robustness of spam models.

1.4. OBJECTIVES AND SCOPE OF THE STUDY

In this paper, our aim is to investigate adversarial attack technique on SMS spam classification model and the effect on performance of the model. The work also intends to survey and compare current robustness improving techniques, such as adversarial training and defensive preprocessing approaches. Our work extends to both classical ML as well DL-based classifiers, specializing in text based adversarial attacks in SMS settings. This paper aims to further the cause of more robust and secure SMS spam detection systems by highlighting key problems and research directions.

2. OVERVIEW OF SMS SPAM CLASSIFICATION MODELS

2.1. TRADITIONAL MACHINE LEARNING APPROACHES

Conventional machine learning techniques are popularly employed in SMS spam filtering because of their simplicity, effectiveness and interpretability. Some of the frequently used algorithms include NB, SVM and LR. Naïve Bayes classifiers make use of the probabilistic assumptions about feature independence and work well on high dimensional sparse text representations like Bag-of-Words and TF-IDF. You cannot go wrong with support vector machines in text classification due to their large feature space nature and their ability to draw optimal decision boundaries. Logistic Regression furnishes a simple, linear and interpretable model that is competitive when working in combination with well-conditioned features. Though these methods are efficient and possible for the realtime usage, due to their dependence on local features they could be vulnerable to the crafted adversarial noises.

2.2. DEEP LEARNING MODELS

Deep learning models have become popular in SMS spam detection because they can automatically learn hierarchical and semantic characteristics from the raw text. CNNs also mitigate to capture a local n-gram pattern and are capable at spotting indicative spam phrases. RNNs (and specifically LSTM networks) are used to capture sequential dependencies and contextual information in SMS messages. More recently, transformer architectures like BERT and its cousins achieved the best performance by using self-attention to model global context and subtle language patterns. While DL models have generally been found to outperform the legacy approaches, they are more complex, data-hungry and computationally expensive and still susceptible to adversarial text manipulations.

2.3. FEATURE REPRESENTATIONS

Feature representation has an influence on the performance of SMS spam classification models. Bag-of-Words (BoW) and Term Frequency–Inverse Document Frequency (TF-IDF) are traditional representations that capture word occurrence and importance, while neglecting word order as well as semantics. Word embeddings, e.g., Word2Vec and GloVe, yield dense vector space representations encoding semantic information between words that support better generalization. Contextual embeddings created by transformer-based models produce dynamic word representations depending on the surrounding context, which carries richer semantic and syntactic information. Although fine-grained embeddings may improve classification, they can also add new venues of attack, since adversarial perturbations could leverage the model's sensitivity to minor context drift.

3. ADVERSARIAL ATTACKS IN NATURAL LANGUAGE PROCESSING

3.1. DEFINITION AND TAXONOMY OF ADVERSARIAL ATTACKS

Adversarial examples in the context of natural language processing, referred to as adversarial attacks, are manipulated input text that is intentionally perturbed to cause a machine learning model to make wrong predictions while for humans there would be no difference. These are usually small and targeted perturbations that make leverage from model weaknesses. Adversarial attacks in NLP come with different levels of manipulation, such as character-level changes (e.g., misspellings, character substitutions), word-level changes (e.g., grammatical errors and semantic modifications), and sentence-level or semantics-based modifications (e.g., paraphrasing, syntactic restructuring). This taxonomy is useful for understanding how various attack strategies impact model behaviour and informing the design of suitable defence mechanisms.

3.2. DIFFERENCES BETWEEN ADVERSARIAL ATTACKS IN TEXT VS. IMAGES

Adversarial attacks in text differ fundamentally from those in image domains. In computer vision, adversarial perturbations often involve adding imperceptible pixel-level noise that does not alter human perception. In contrast, text data is discrete and structured, making arbitrary perturbations easily noticeable or semantically destructive. NLP adversarial attacks must therefore maintain grammatical correctness and semantic coherence to remain effective. Additionally, small textual changes—such as replacing a single word or altering spelling—can disproportionately impact model predictions, especially in short texts like SMS messages. These constraints make text-based adversarial attacks more challenging to design but also harder to defend against.

3.3. THREAT MODELS

Threat models specify the extent of attacker's knowledge and access to target system. In white-box attacks, the adversary has full information of model structure, weights as well as training data which makes these attacks highly effective and tailored. Black-box attacks take no internal knowledge about the model; adversaries utilize model queries and observations on the output to generate adversarial examples with transferability using substitute models. When there is only limited information, e.g., feature types or training regime, gray-box attacks can attack the intermediate scenarios between these two extremes. It is important to understand these threat models in order to assess the resilience of SMS spam classification systems against real-world attacks.

4. TYPES OF ADVERSARIAL ATTACKS ON SMS SPAM CLASSIFIERS

4.1. CHARACTER-LEVEL ATTACKS

Character-level perturbations are one of the adversarial attack types frequently used in SMS spam classification tasks because text messages are typically short and informal. These attacks include misspellings, addition or deletion of the characters, and use of homoglyphs that are visually confusable characters in more than one script (e.g., substituting “o” with “0”). These perturbations typically have negligible effect on human readability, but can severely disturb word segmentation and feature extraction, especially when building models based on exact word matching or sub-word tokenization. Character-level attacks are cheaper in terms of computation and used by spammers as they bypass many detection systems.

4.2. WORD-LEVEL ATTACKS

Word level attacks perturb the tokens in the lexical level trying to keep the same mean- ing of the original message. Common approaches include (i) synonym substitution, replacing words with their synonyms and (ii) word reordering, rearranging words in sentence to satisfy the expected context. Such attacks could take advantage of the sensitivity of models to certain, spam-related keywords. Word level attacks are generally more difficult to implement than character level attack, but can significantly reduce the word distribution or static embedding-based models.

4.3. SEMANTIC-PRESERVING ATTACKS IN SHORT TEXTS

Semantic-preserving attacks are trying to keep the general sense and semantic of the original message, while making it classified wrong. This appears to be even a more difficult task in short text domains like SMS where there is weak contextual redundancy. Methods such as paraphrasing, phrase replacement or adding neutral or harmless tokens are employed to modify model perception without human users suspecting anything. Even minor semantic-preserving modifications can have a profound impact on model predictions in SMS classification, which accentuates the fragile nature of existing models for small-scale text.

4.4. OBFUSCATION TECHNIQUES COMMONLY USED BY SPAMMERS

In practice, spammers adopt many obfuscation techniques that are inline to adversarial attack strategies. Such as using symbols or numbers in place of characters, spam keywords being broken up with special characters, adding a word for no other seasonal than to fill the page with it, weird spacing or punctuation. The obfuscation can be automated, and is intended to defeat both rule-based and learning-based filters. "The purpose of understanding these practical strategies is so that we can build strong SMS spam classifiers and shield them from academic adversarial attacks and real-world spam evasion tactics."

5. IMPACT OF ADVERSARIAL ATTACKS ON SMS SPAM DETECTION

5.1. DEGRADATION OF CLASSIFICATION ACCURACY AND ROBUSTNESS

Adversarial attacks can perform challenges for SMS spam classification models. Spam can be misclassified as legitimate messages, and vice versa, even when minor perturbation is induced by character- or word-level changes, for both conventional machine learning based system and also deep learning models. This loss in accuracy indicates that the models are very sensitive to slight differences in input text, and particularly for short messages, where each token has significant influence. In practice, loss of robustness is particularly worrisome as attackers can leverage it to circumvent automated filters and get payload delivered to users.

5.2. EFFECTS ON PRECISION, RECALL, AND FALSE NEGATIVE RATES

Adversarial perturbations degrade the performance of critical metrics. But also, there is the negative effect to recall because spam messages are not caught which raises false negatives. High false negative rates are particularly risky for SMS systems, since missed spam can contain phishing, fraud or malware URLs. It has been demonstrated that some attacks can dramatically lower the recall by 20–40% or even higher given different model architectures and input perturbation. Balancing precision and recall under adversarial conditions is still one of the main challenges for SMS spam filtering.

5.3. CASE STUDIES AND EMPIRICAL EXAMPLES FROM PRIOR RESEARCH

It has been shown in practice that SMS spam classifiers are sensitive to adversarial examples. For instance, studies based on character-level misspellings or homoglyph substitutions demonstrate that Naïve Bayes and SVM models may misclassify up to 30% of spam messages. The deep learning models such as LSTM and transformer-based classifiers, although more accurate when clean data is available, have been reported to fail under word-level targeted or semantic-preserving poisoning attacks. The findings of case studies underscore that models trained without regard to adversarial robustness are especially susceptible to evasion, further underlining the importance of mitigating measures like adversarial training, data augmentation and input sanitization. These results indicate that the robustness of SMS spam detection systems need to not only be examined on clean data-masses, but also in adversarial cases for real-life reliability.

6. ROBUSTNESS EVALUATION OF SMS SPAM CLASSIFICATION MODELS

6.1. METRICS FOR ROBUSTNESS ASSESSMENT

Evaluating the robustness of SMS spam classifiers requires metrics that go beyond traditional accuracy. Key metrics include:

- **Robust Accuracy:** Measures model accuracy on adversarial perturbed inputs rather than clean data, reflecting the system’s ability to maintain correct predictions under attacks.

- **Attack Success Rate (ASR):** The proportion of adversarial examples that successfully cause misclassification, indicating the vulnerability of a model to specific attack strategies.
- **False Negative Rate (FNR) under Attack:** Tracks undetected spam messages, which is critical for assessing real-world risks.
- **Robust F1-Score and Precision/Recall:** Evaluated on adversarial test sets to provide a balanced view of performance under perturbations.
- **Perturbation Metrics:** Quantify the magnitude of input modifications (e.g., number of character/word changes) required to mislead the model, helping to compare model resilience across different attack strengths.

6.2. BENCHMARK DATASETS AND ADVERSARIAL TEST SETS

Standard SMS datasets such as **SMS Spam Collection**, **NUS SMS Corpus**, and **LingSpam** are commonly used for evaluating baseline classification performance. For robustness testing, adversarial test sets are generated by applying attack strategies to these datasets, including:

- Character-level perturbations (misspellings, homoglyphs)
- Word-level substitutions and reordering
- Semantic-preserving paraphrasing and token insertion
- Obfuscation patterns mimicking real-world spam tactics

These adversarial datasets allow for systematic evaluation of a model's vulnerability and help benchmark different defense mechanisms.

6.3. EXPERIMENTAL SETUPS FOR EVALUATING ADVERSARIAL RESILIENCE

Experimental evaluation typically involves the following steps:

- **Baseline Training:** Train the SMS spam classifier on clean training data using selected ML or DL architectures.
- **Adversarial Example Generation:** Apply one or more attack methods to the test set to create adversarial samples.
- **Performance Measurement:** Evaluate model performance on both clean and adversarial test sets using robustness metrics.
- **Comparative Analysis:** Compare different models or defense strategies under identical adversarial conditions to identify strengths and weaknesses.
- **Ablation Studies (Optional):** Assess the impact of specific features, embeddings, or preprocessing techniques on robustness to understand key vulnerabilities.

This evaluation framework provides a comprehensive view of how SMS spam classifiers behave under adversarial conditions, guiding the design of more resilient systems.

7. DEFENSE MECHANISMS AND ROBUSTNESS ENHANCEMENT TECHNIQUES

7.1. ADVERSARIAL TRAINING APPROACHES

Adversarial training is one of the most popular methods to enhance model robustness. It consists in enriching the training data set with samples of adversarially perturbed instances to teach the model how to reach the right prediction for clean and infected inputs. In the context of SMS spam filtering, this might be character-level, word-level and preserving semantics perturbations. Adversarial training achieves this by injecting potential attack patterns into the model during training to mitigate the mis-classification rates under adversarial scenarios. This are extended to online adversarial training which generates the adversarial examples on-the-fly during model training as well as multi-attack adversarial training that combines multiple attack strategies for enhanced overall robustness.

7.2. DATA AUGMENTATION AND NOISE INJECTION

Data augmentation methods artificially increase the size of training dataset and to obtain better generalization and robustness. Standard approaches involve synonym replacement, paraphrasing, random character insertion or deletion, and adding benign tokens. Noise injection (e.g. random misspellings or token substitution) mimics the natural variations in SMS and teaches the model to not overfit to very specific forms of messages. Both methods intend to expose the classifier to a more extensive spectrum of input variations, thus reducing the susceptibility with adversarial perturbations while preserving human-readable message semantics.

7.3. ROBUST FEATURE ENGINEERING AND PREPROCESSING

Robust feature engineering focuses on designing representations that are less sensitive to adversarial perturbations. Techniques include:

- **Character- and subword-level embeddings** to handle spelling variations and homoglyphs.
- **Contextual embeddings** from transformer-based models to capture semantic meaning beyond exact word matches.
- **Preprocessing pipelines** that normalize text, remove irrelevant noise, or correct common misspellings before classification.

By improving input representations, these methods reduce the model's vulnerability to minor manipulations in SMS text.

7.4. MODEL REGULARIZATION AND ENSEMBLE METHODS

Regularization approaches including dropout, weight decay and adversarial regularization are effective to avoid overfitting with single patterns for better generalizability and resistance against adversarial inputs. In order to undermine the effect of attacks on a single model, we can combine predictions from different classifiers that may adopt various architectures or feature representations, a strategy referred to as an ensemble method. Methods such as majority voting, stacking or boosting increase robustness because misclassification of one model can be compensated by the others. They are particularly successful in countering model-specific attacks, and are widely used in spam detection systems for high-stakes tasks. When properly aggregated, these defense mechanisms together can greatly improve the resiliency of SMS spam classifiers, but at a cost to computational cost and model complexity.

8. CHALLENGES AND LIMITATIONS

8.1. SHORT LENGTH AND INFORMAL NATURE OF SMS TEXTS

Natural language in SMS is short by nature (with an average length typically not exceeding 160 characters for a single message) and includes informal content, abbreviations, acronyms, slang terms and punctuations with interchangeable use of personal and impersonal pronouns. This compactness makes the model highly sensitive to small adversarial perturbations: a single word or character can turn a sentence into an attack. More generally, the informal and noisy language used in SMS text complicates feature extraction and semantic understanding, which mitigates the performance of existing traditional and deep learning approaches for robustness in adversarial settings.

8.2. TRADE-OFF BETWEEN ROBUSTNESS AND MODEL PERFORMANCE

Robustness against adversarial examples is usually achieved at the cost of sacrificing performance in standard scenario. Methods such as adversarial training, data augmentation and ensemble modeling can enhance robustness but they also bring additional challenges to training complexity, clean-accuracy loss or inference slowdown. The trade-off between high classification accuracy, low false negatives and robustness to attacks is still an area of focus in SMS spam filtering.

8.3. COMPUTATIONAL AND DEPLOYMENT CONSTRAINTS IN REAL-TIME SYSTEMS

SMS spam filtering mechanisms are often employed in real-time scenarios, like mobile networks or messaging services where low response time and power consumption are crucial. Methods for enhancing robustness, particularly those based on deep learning models or ensemble workings or requiring extensive adversarial training, can be computationally expensive and challenging to scale. Memory, computation power, and energy of mobile or edge devices are also limited, thus even more the real-time operation problem being hard without quality decay. These little idiosyncrasies emphasize the difficulties on building a strict SMS spam classifiers with high-rate of accuracy and resilience, reiterating the demand for effective lightweight robust ways to prevent spam.

9. FUTURE RESEARCH DIRECTIONS

9.1. DEVELOPMENT OF LIGHTWEIGHT YET ROBUST MODELS

Further studies might concentrate on the development of robust and yet computationally efficient SMS spam classifiers. Lightweight architectures, like mobile transformer or slim neural network, can achieve powerful adversarial robustness and leverageful for deployment on mobile and edge devices. Other model compression techniques, such as sparsity, quantization, and knowledge distillation could be applied to reduce the size and latency of the models without a considerable loss of robustness and thus improve real-time spam detection at scale.

9.2. ADAPTIVE DEFENSES AGAINST EVOLVING ADVERSARIAL STRATEGIES

Threat actors change tactics over evasion systems – defenses must evolve as well. As a direction for future research we also believe that it may be promising to investigate dynamic adversarial training, continual learning and online model updating (i.e., retraining periodically the classifiers on newly observed attack patterns). Embedding adaptability in models will enable the models to be robust against new attacks without requiring extensive user intervention.

9.3. MULTILINGUAL AND CODE-MIXED ADVERSARIAL ROBUSTNESS

Given the worldwide popularity of SMS usage and the ubiquity of multilingual and code-mixed messages, future work should generalize these adversarial robustness initiatives to languages other than English. Models need to understand linguistic variations, transliterations and mixing of languages, while remaining robust to adversarial manipulations. Cross-lingual embeddings, multilingual transformer models and language-neutral preprocessing methods could also contribute significantly to obtaining robustness on various SMS datasets.

9.4. INTEGRATION OF EXPLAINABILITY FOR DETECTING ADVERSARIAL BEHAVIOR

Explainable AI (XAI) methods offer a way to interpret models' predictions and assist in the detection of suspicious/adversarially modified messages. In the future, one could add plug-in attention visualization (or feature attribution or saliency mapping) to identify anomalies due to adversaries' perturbation. Interpretability enhances trust and transparency, and can serve as a tool for proactive defences by elucidating weaknesses in SMS spam classification models. By exploring these directions, future studies can design more robust, adaptive and universally applicable SMS spam detection systems against increasingly advanced adversarial attacks.

10. CONCLUSION

In this work, we have explored the influence of adversarial attacks on SMS spam classification models and approached their robustness improvement. Hateful content has been furtively under the spotlight because adversaries can regrettably misuse adversarial examples (AEs) to modify data for a number of existing systems, including image and audio recognition-related tools. Robustness measurement in metrics such as attack success rate, robust accuracy, and false negative rate indicates that even slight alterations to the text can compromise the robustness. Defense mechanisms including adversarial training, data augmentation, robust feature extraction, model regularization and ensemble methods provide promising prospects for enhancing the resilience. Nevertheless, there are remained challenges such as short and informal of SMS text design, the trade-off between robustness and accuracy of model and the limitation of computation in online service. The findings from this study have implications for the need to design robust SMS spam detection systems which are effective both on clean data and robust against adversarial tampering. Building on these with light, adaptive, multilingual and explainable solutions, we believe that future systems can successfully realize secure and reliable spam filtering in practical mobile communication environments that are immune against arising threats.

REFERENCES

[1] B. Narra, D. V. K. R. Buddula, H. Patchipulusu, N. Vattikonda, A. Gupta, and A. R. Polu, "The Integration of Artificial Intelligence in Software Development: Trends, Tools, and Future Prospects," SSRN Electronic Journal, 2025, doi: https://doi.org/10.2139/ssrn.5596472.

[2] A. K. Gupta et al., "Leveraging deep learning models for intrusion detection systems for secure networks," Journal of Computer Science and Technology Studies, vol. 6, no. 2, pp. 199-208, 2024, doi: https://doi.org/10.32996/jcsts.2024.6.2.22

[3] R. P. Achuthananda et al., "Evaluating machine learning approaches for personalized movie recommendations: A comprehensive analysis," Journal of Contemporary Education Theory & Artificial Intelligence, pp. 1-8, 2024.

[4] A. R. Polu et al., "Analyzing The Role of Analytics in Insurance Risk Management: A Systematic Review of Process Improvement and Business Agility," IRJEMS International Research Journal of Economics and Management Studies, vol. 2, no. 3, pp. 325-332, 2025, Doi: https://doi.org/10.56472/25835238/IRJEMS-V2I1P142

[5] V. Tamilmani et al., "A Review of Cyber Threat Detection in Software-Defined and Virtualized Networking Infrastructures," International Journal of Technology, Management and Humanities, vol. 10, no. 4, pp. 136-146, 2024, doi: https://doi.org/10.21590/ijtmh.10.04.15

[6] V. Bitkuri et al., "A Survey on Blockchain-Enabled ERP Systems for Secure Supply Chain Processes and Cloud Integration," International Journal of Technology Management and Humanities, vol. 10, no. 02, pp. 52–65, Jun. 2024, doi: https://doi.org/10.21590/ijtmh.2024100209.

[7] Jaya Vardhani Mamidala et al., "Machine Learning Approaches to Salary Prediction in Human Resource Payroll Systems," Journal of Computer Science and Technology Studies, vol. 7, no. 10, pp. 528–536, Oct. 2025, doi: https://doi.org/10.32996/jcsts.2025.7.10.52.

[8] P. Waditwar, "Reimagining procurement payments: From transactional bottlenecks to strategic value creation," World Journal of Advanced Research and Reviews, vol. 28, no. 1, pp. 588–598, Oct. 2025, doi: https://doi.org/10.30574/wjarr.2025.28.1.3480.

[9] A. Attipalli et al., "Privacy Preservation in the Cloud: A Comprehensive Review of Encryption and Anonymization Methods," International Journal of Multidisciplinary on Science and Management IJMSM, vol. 1, no. 1, pp. 35-44, 2024, doi: https://doi.org/10.71141/30485037/V1I1P106

[10] S. J. Enokkaren et al., "Artificial Intelligence (AI)-Based Advance Models for Proactive Payroll Fraud Detection and Prevention, International Journal of Machine Learning, AI & Data Science Evolution, vol. 1, no. 1, pp. 1-11, 2024, doi: https://doi.org/10.63665/ijmlaidse-y1f1a001

[11] M. S. V. Tyagadurgam, "AI-Powered Cybersecurity Risk Scoring for Financial Institutions Using Machine Learning Techniques," Journal of Artificial Intelligence & Cloud Computing, pp. 1–9, Dec. 2024, doi: https://doi.org/10.47363/jaicc/2024(3)452.

[12] P. Waditwar, "The Intersection of Strategic Sourcing and Artificial Intelligence: A Paradigm Shift for Modern Organizations," Open Journal of Business and Management, vol. 12, no. 06, pp. 4073–4085, 2024, doi: https://doi.org/10.4236/ojbm.2024.126204.

[13] D. Rajendran, Venkata Deepak Namburi, Vetrivelan Tamilmani, A. Arjun, V. Maniar, and Rami Reddy Kothamaram, "Middleware Architectures for Hybrid and Multi-cloud Environments: A Survey of Scalability and Security Approaches," Asian Journal of Research in Computer Science, vol. 19, no. 1, pp. 106–120, Jan. 2026, doi: https://doi.org/10.9734/ajrcos/2026/v19i1808.

[14] P. Waditwar, "De-Risking Returns: How AI Can Reinvent Big Tech's China-Tied Reverse Supply Chains," Open Journal of Business and Management, vol. 14, pp. 104-124, 2026, doi: 10.4236/ojbm.2026.141007

[15] V. Maniar, R. R. Kothamaram, D. Rajendran, V. D. Namburi, V. Tamilmani, and A. A. S. Singh, "A Comprehensive Survey on Digital Transformation and Technology Adoption Across Small and Medium Enterprises," European Journal of Applied Science, Engineering and Technology, vol. 3, no. 6, pp. 238–250, Dec. 2025, doi: https://doi.org/10.59324/ejaset.2025.3(6).18.

[16] V. Tamilmani et al., “Automated Cloud Migration Pipelines: Trends, Tools, and Best Practices–A Survey,” Journal of Computer Science and Technology Studies, vol. 7, no. 11, pp. 121-134, 2025, doi: https://doi.org/10.32996/jcsts.2025.7.11.14

[17] A. Attipalli, R. Kendyala, J. Kurma, J. V. Mamidala, V. Bitkuri, and S. J. Enokkaren, “Survey on Evolution of Java Web Technologies and Best Practices: from Servlets to Microservices,” Asian Journal of Research in Computer Science, vol. 18, no. 11, pp. 172–187, Nov. 2025, doi: https://doi.org/10.9734/ajrcos/2025/v18i11786.

[18] Jaya Vardhani Mamidala et al., “Explainable Machine Learning Models for Malware Identification in Modern Computing Systems,” European Journal of Applied Science Engineering and Technology, vol. 3, no. 5, pp. 153–170, Oct. 2025, doi: https://doi.org/10.59324/ejaset.2025.3(5).13.

[19] Prajkta Waditwar, “AI-Driven Smart Negotiation Assistant for Procurement—An Intelligent Chatbot for Contract Negotiation Based on Market Data and AI Algorithms,” Journal of Data Analysis and Information Processing, vol. 13, no. 02, pp. 140–155, Jan. 2025, doi: https://doi.org/10.4236/jdaip.2025.132009.

[20] Raghuvaran Kendyala, Jagan Kurma, Jaya Vardhani Mamidala, Sunil Jacob Enokkaren, Avinash Attipalli, and Varun Bitkuri, “Framework based on Machine Learning for Lung Cancer Prognosis with Big Data-Driven,” European Journal of Technology, vol. 9, no. 1, pp. 68–85, Oct. 2025, doi: https://doi.org/10.47672/ejt.2787.

[21] A. B. Kakani, S. K. K. Nandiraju, S. K. Chundru, S. R. Vangala, R. M. Polam, and B. Kamarthapu, “Big Data and Predictive Analytics for Customer Retention: Exploring the Role of Machine Learning in E-Commerce,” SSRN Electronic Journal, 2025, doi: https://doi.org/10.2139/ssrn.5515281.

[22] P. Kulkarni, T. Siddharth, S. Pillai, P. Pathak, V. N. Gangineni, and V. Yadav, “Cybersecurity Threats and Vulnerabilities - A Growing Challenge in Connected Vehicles,” Lecture Notes in Networks and Systems, pp. 466–476, Nov. 2025, doi: https://doi.org/10.1007/978-3-032-03558-5_39.

[23] N. R. Vanaparthi, “Intelligent Finance: How AI is Reshaping the Future of Financial Services,” International Journal of Computer Engineering and Technology, vol. 16, no. 1, pp. 126–137, Jan. 2025, doi: https://doi.org/10.34218/ijcet_16_01_012.

[24] M. Sai, Venkataswamy Naidu Gangineni, Sriram Pabbineedi, Ajay Babu Kakani, K. Kireeti, and Sandeep Kumar Chundru, “Preventing Phishing Attacks Using Advanced Deep Learning Techniques for Cyber Threat Mitigation,” Journal of Data Analysis and Information Processing, vol. 13, no. 03, pp. 314–330, Jan. 2025, doi: https://doi.org/10.4236/jdaip.2025.133020.

[25] M. Penmetsa, J. R. Bhumireddy, R. Chalasani, S. R. Vangala, R. M. Polam, and B. Kamarthapu, “Adversarial Machine Learning in Cybersecurity: A Review on Defending Against AI-Driven Attacks,” European Journal of Applied Science, Engineering and Technology, vol. 3, no. 4, pp. 4–14, Jun. 2025, doi: https://doi.org/10.59324/ejaset.2025.3(4).01

[26] Ram Mohan Polam, Bhavana Kamarthapu, Mitra Penmetsa, Jayakeshav Reddy Bhumireddy, R. Chalasani, and Srikanth Reddy Vangala, “Advanced Machine Learning for Robust Botnet Attack Detection in Evolving Threat Landscapes,” Asian Journal of Research in Computer Science, vol. 18, no. 8, pp. 1–14, Aug. 2025, doi: https://doi.org/10.9734/ajrcos/2025/v18i8735.

[27] B. Kamarthapu, M. Penmetsa, J. R. Bhumireddy, R. Chalasani, S. R. Vangala, and R. M. Polam, “Data-Driven Detection of Network Threats using Advanced Machine Learning Techniques for Cybersecurity,” SSRN Electronic Journal, 2025, doi: https://doi.org/10.2139/ssrn.5515400.

[28] M. Penmetsa, J. R. Bhumireddy, R. Chalasani, S. R. Vangala, R. M. Polam, and B. Kamarthapu, “Effectiveness of Deep Learning Algorithms in Phishing Attack Detection for Cybersecurity Frameworks,” SSRN Electronic Journal, 2025, doi: https://doi.org/10.2139/ssrn.5515385.

[29] S. K. K. Nandiraju et al., “Towards Early Forecast of Diabetes Mellitus via Machine Learning Systems in Healthcare,” European Journal of Technology, vol. 9, no. 1, pp. 35-50, 2025.

[30] R. M. Polam, B. Kamarthapu, A. B. Kakani, S. K. K. Nandiraju, S. K. Chundru, and S. R. Vangala, “Predictive Modeling for Property Insurance Premium Estimation Using Machine Learning Algorithms,” SSRN Electronic Journal, 2025, doi: https://doi.org/10.2139/ssrn.5515382.

[31] P. Waditwar, “Agentic AI and sustainable procurement: Rethinking anti-corrosion strategies in oil and gas,” World Journal of Advanced Research and Reviews, vol. 27, no. 3, pp. 1591–1598, Sep. 2025, doi: https://doi.org/10.30574/wjarr.2025.27.3.3298.

[32] R. Vadisetty et al., “Cyber Warfare and AI Agents: Strengthening National Security Against Advanced Persistent Threats (APTs),” International Conference on Intelligence-Based Transformations of Technology and Business Trends, Cham: Springer Nature Switzerland, pp. 578-587, 2025.

[33] S. K. Chundru, M. S. V. Tyagadurgam, V. N. Gangineni, S. Pabbineedi, A. B. Kakani, and S. K. K. Nandiraju, “Analyzing and Predicting Anaemia with Advanced Machine Learning Techniques with Comparative Analysis,” International Journal of Applied Information Systems, vol. 13, no. 1, pp. 28–36, Aug. 2025, doi: https://doi.org/10.5120/ijais2025452027.

[34] Ram Mohan Polam, Bhavana Kamarthapu, Mitra Penmetsa, Jayakeshav Reddy Bhumireddy, R. Chalasani, and Srikanth Reddy Vangala, “Advanced Machine Learning for Robust Botnet Attack Detection in Evolving Threat Landscapes,” Asian Journal of Research in Computer Science, vol. 18, no. 8, pp. 1–14, Aug. 2025, doi: https://doi.org/10.9734/ajrcos/2025/v18i8735.

[35] B. Kamarthapu, M. Penmetsa, J. R. Bhumireddy, R. Chalasani, S. R. Vangala, and R. M. Polam, “Data-Driven Detection of Network Threats using Advanced Machine Learning Techniques for Cybersecurity,” SSRN Electronic Journal, 2025, doi: https://doi.org/10.2139/ssrn.5515400.

[36] Narasimha Rao Vanaparthi, “Why digital transformation in fintech requires mainframe modernization: A cost-benefit analysis,” International Journal of Science and Research Archive, vol. 14, no. 1, pp. 1052–1062, Jan. 2025, doi: https://doi.org/10.30574/ijsra.2025.14.1.0161.

[37] Ajay Babu Kakani, K. Kireeti, Sandeep Kumar Chundru, Srikanth Reddy Vangala, Ram Mohan Polam, and Bhavana Kamarthapu, “Leveraging NLP and Sentiment Analysis for ML-Based Fake News Detection with Big Data,” SSRN Electronic Journal, Jan. 2025, doi: https://doi.org/10.2139/ssrn.5515418.

[38] Prajkta Waditwar, “Quantum-Enhanced Travel Procurement: Hybrid Quantum–Classical Optimization for Enterprise Travel Management,” World Journal of Advanced Engineering Technology and Sciences, vol. 17, no. 3, pp. 375–386, Dec. 2025, doi: https://doi.org/10.30574/wjaets.2025.17.3.1572.

[39] Narasimha Rao Vanaparthi, “REGULATORY COMPLIANCE IN THE DIGITAL AGE: HOW MAINFRAME MODERNIZATION CAN SUPPORT FINANCIAL INSTITUTIONS,” vol. 8, no. 1, pp. 383–396, Jan. 2025, doi: https://doi.org/10.34218/IJRCAIT_08_01_033.

[40] P. Waditwar, “AI-Driven Procurement in Ayurveda and Ayurvedic Medicines & Treatments,” Open Journal of Business and Management, vol. 13, no. 03, pp. 1854–1879, 2025, doi: https://doi.org/10.4236/ojbm.2025.133096.

[41] N. Rao, “The Roadmap to Mainframe Modernization: Bridging Legacy Systems with the Cloud,” International Journal of Scientific Research in Computer Science Engineering and Information Technology, vol. 11, no. 1, pp. 125–133, Jan. 2025, doi: https://doi.org/10.32628/cseit25111214.

[42] D. Prabakar, N. Iskandarova, N. Iskandarova, D. Kalla, K. Kulimova, and D. Parmar, “Dynamic Resource Allocation in Cloud Computing Environments Using Hybrid Swarm Intelligence Algorithms,” 2025 International Conference on Networks and Cryptology (NETCRYPT), pp. 882–886, May 2025, doi: https://doi.org/10.1109/netcrypt65877.2025.11102314

[43] Subuddi Nagaraju, Prashant Johri, Prakash Putta, D. Kalla, Sultonmakhmud Polvanov, and N. V. Patel, “Smart Routing in Urban Wireless Ad Hoc Networks Using Graph Attention Network-Based Decision Models,” pp. 212–216, May 2025, doi: https://doi.org/10.1109/netcrypt65877.2025.11102255

[44] D. Kalla, A. S. Mohammed, V. N. Boddapati, N. Jiwani, and T. Kiruthiga, “Investigating the Impact of Heuristic Algorithms on Cyberthreat Detection,” 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), pp. 450–455, Nov. 2024, doi: https://doi.org/10.1109/icaiccit64383.2024.10912106.

[45] Rahul Vadisetty, Anand Polamarasetti, and D. Kalla, “Automated AI-Driven Phishing Detection and Countermeasures for Zero-Day Phishing Attacks,” Lecture notes in networks and systems, pp. 285–303, Jan. 2026, doi: https://doi.org/10.1007/978-981-96-8632-2_16.

[46] Preeti Nagrath, I. Saini, M. Zeeshan, Komal, Komal, and D. Kalla, “Predicting Mental Health Disorders with Variational Autoencoders,” Lecture notes in networks and systems, pp. 38–51, Oct. 2025, doi: https://doi.org/10.1007/978-3-032-03751-0_4.

[47] World Bank, Small and medium enterprises (SMEs) finance. World Bank Group, 2020. [Online]. Available; https://www.worldbank.org/en/topic/smefinance

[48] H. Zhai, M. Yang, and K. C. Chan, “Does digital transformation enhance a firm’s performance? Evidence from China,” Technology in Society, vol. 68, p. 101841, Feb. 2022, doi: https://doi.org/10.1016/j.techsoc.2021.101841.

[49] J. A. Restrepo-Morales, J. A. Ararat-Herrera, D. A. López-Cadavid, and A. Camacho-Vargas, “Breaking the digitalization barrier for SMEs: a fuzzy logic approach to overcoming challenges in business transformation,” Journal of Innovation and Entrepreneurship, vol. 13, no. 1, Nov. 2024, doi: https://doi.org/10.1186/s13731-024-00429-w.