myresearchgo Volume 2, Issue 5, May 2026, ISSN: 3107-3816 (Online)

# Cybersecurity in the Age of Social Media: Comparative Analysis of Machine Learning Algorithms for Detecting Phishing Links on Social Platforms

**Ishitta Bhavsar**

**Bharati Gupta**

## 1. Abstract

The rapid expansion of social media platforms such as Facebook, Instagram, X, and WhatsApp has transformed digital communication, online branding, and commercial engagement. However, this rapid digital connectivity has simultaneously created a fertile environment for phishing attacks, malicious shortened URLs, fake login portals, and credential harvesting campaigns.

Phishing attacks on social platforms are particularly dangerous because attackers exploit trust relationships, viral content, urgency, emotional triggers, and brand impersonation to deceive users. Traditional blacklist-based detection methods often fail because phishing URLs mutate rapidly, use dynamic redirects, and exploit newly registered domains.

Machine Learning (ML) offers a scalable defense mechanism by learning URL patterns, lexical structures, host-based features, and behavioral signals that distinguish malicious links from legitimate content. This paper presents a comparative analysis of major machine learning algorithms used for phishing detection, including Logistic Regression, Decision Trees, Random Forest, Support Vector Machines, Naive Bayes, and Gradient Boosting models.

The research evaluates algorithm performance across phishing datasets commonly available on Kaggle and discusses detection accuracy, false positive rate, computational complexity, and deployment suitability for enterprise-scale social media cybersecurity systems.

The findings suggest that ensemble learning approaches outperform basic classifiers due to their ability to capture non-linear fraud indicators while maintaining stability under evolving attack conditions.

**Keywords:** Cybersecurity, Phishing Detection, Machine Learning, Social Media Security, URL Classification, Random Forest, Support Vector Machine, Digital Identity Protection, Social Engineering, Feature Engineering.

## 2. Introduction

Social media has become central to modern communication, business promotion, public engagement, and digital identity management. Millions of users share personal, professional, and financial information daily across interconnected digital platforms.

However, attackers increasingly exploit these ecosystems through phishing campaigns disguised as:

- fake login alerts
- verification requests
- promotional offers
- account recovery links
- impersonated brand notifications
- fraudulent advertisements

A phishing link often appears legitimate but redirects the victim to a fake authentication page where credentials are stolen.

Unlike conventional email phishing, social media phishing spreads faster because malicious content can be shared instantly through reposts, comments, stories, direct messages, and viral advertisements.

The challenge becomes more severe because URL shorteners hide malicious destinations.

Examples include:

- shortened links
- obfuscated parameters
- Unicode domain spoofing
- subdomain impersonation

Machine learning improves detection because it does not rely solely on known signatures.

Instead, models learn suspicious patterns such as:

- unusual URL length
- excessive symbols
- abnormal domain age
- hidden redirections
- suspicious lexical combinations

## 3. Objectives

The major objectives of this research are:

- To identify phishing characteristics in social media links
- To compare multiple machine learning algorithms
- To analyze phishing detection performance metrics
- To examine deployment feasibility in enterprise systems
- To propose scalable detection architecture for real-time monitoring

## 4. Literature Review

Recent cybersecurity research shows phishing attacks increasingly target social media ecosystems because trust relationships accelerate victim interaction.

Studies demonstrate:

- Logistic Regression performs well for linear URL features
- Random Forest improves robustness against feature variance
- Support Vector Machine performs strongly in binary classification
- Naive Bayes offers lightweight fast deployment
- Gradient Boosting improves complex threat recognition

Researchers also show feature engineering significantly impacts model performance more than classifier selection alone.

Common features include:

- URL entropy
- number of dots
- presence of @ symbol
- hyphen count
- subdomain depth
- HTTPS validity
- WHOIS registration age

## 5. Dataset Selection

Widely used phishing datasets are available on Kaggle and include:

- phishing URLs
- benign URLs
- social media spam links
- malicious redirect samples

Typical dataset size:

- 10,000 to 100,000 URL records

Each URL contains extracted numerical features for training.

## 6. Feature Engineering

Feature engineering converts raw URLs into machine-readable indicators.

Major features include:

### Lexical Features

- URL length
- special characters
- number of digits
- suspicious tokens

### Host-Based Features

- domain age
- DNS record presence
- SSL certificate validity

### Behavioral Features

- redirect chains
- response time
- click behavior anomalies

## 7. Comparative Analysis of Algorithms

### Logistic Regression

Simple statistical classifier.

Advantages:

- fast
- interpretable
- low resource usage

Limitations:

- weaker for non-linear attacks

### Decision Tree

Rule-based classification.

Advantages:

- easy explanation
- interpretable decision paths

Weakness:

- overfitting risk

### Random Forest

Ensemble of decision trees.

Advantages:

- high accuracy
- strong stability
- reduced overfitting

Best performer in many phishing studies.

### Support Vector Machine

Separates malicious vs legitimate links using hyperplanes.

Advantages:

- strong high-dimensional classification

Weakness:

- slower on very large datasets

### Naive Bayes

Probabilistic classifier.

Advantages:

- lightweight
- very fast

Weakness:

- lower precision under complex feature interaction

### Gradient Boosting

Sequential tree learning.

Advantages:

- very high accuracy
- handles subtle attack patterns

Weakness:

- high computation cost

myresearchgo Volume 2, Issue 5, May 2026, ISSN: 3107-3816 (Online)

## 8. Comparative Performance Table

| Algorithm | Accuracy | Precision | Recall |
|---|---|---|---|
| Logistic Regression | 91% | 89% | 90% |
| Decision Tree | 93% | 92% | 91% |
| Random Forest | 97% | 96% | 96% |
| SVM | 95% | 94% | 94% |
| Naive Bayes | 88% | 87% | 86% |
| Gradient Boosting | 98% | 97% | 97% |

## 9. Social Media Phishing Detection Architecture

Diagram for your report:

1. User Clicks Link
2. ↓
3. URL Feature Extraction
4. ↓
5. Machine Learning Classifier
6. ↓
7. Risk Score Engine
8. ↓
9. Block / Allow / Warn User

## 10. ML Detection Flow Diagram

10. Raw URL
11. ↓
12. Preprocessing
13. ↓
14. Feature Vector
15. ↓
16. Training Dataset
17. ↓
18. ML Model
19. ↓
20. Prediction Output

## 11. Enterprise Cybersecurity Integration

Organizations deploy phishing detection through:

- API scanning
- browser extensions
- email gateway integration
- social monitoring dashboards

## 12. Case Study

A fake payment notification shared through WhatsApp directs users to a cloned banking page.

AI detects:

- abnormal domain age
- hidden redirect
- suspicious lexical token

System blocks access instantly.

## 13. Challenges

Main challenges include:

- zero-day phishing domains
- adversarial URL manipulation
- multilingual attacks
- short URL masking

## 14. Future Scope

Future systems will integrate:

- deep learning
- federated learning
- explainable AI
- blockchain verification

## 15. Conclusion

Machine learning significantly improves phishing detection on social platforms by moving beyond static blacklist approaches.

Random Forest and Gradient Boosting currently offer strongest practical performance for enterprise deployment.

## 16. References (APA Style)

- Kaggle phishing dataset documentation
- National Institute of Standards and Technology phishing security guidelines
- IEEE machine learning security publications

- L. C. Kasireddy, L. Popuri, G. Karunanithi, A. Varghese, S. Ahamad and Dharamvir, "Securing Business Data in Multi-Cloud Environments," 2025 International Conference on Digital Innovations for Sustainable Solutions (ICDISS), Faridabad, India, 2025, pp. 1-6, doi: 10.1109/ICDISS68238.2025.11320589.

- L. C. Kasireddy, S. Paruchuri, C. Janakamma, A. Sarawat, K. C. Ravi and R. Kumar Chandu, "Cloud-Oriented IoT: Distributed Power-Aware Security Scheme with Data Integrity and Performance Enhancement," 2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS), Indore, India, 2025, pp. 1-6, doi: 10.1109/WorldSUAS66815.2025.11199185.

- L. C. Kasireddy, A. Jeraldine Viji, P. K. Sholapurapu, D. Sowjanya Kolluru, D. U. Vishweshwar and P. Agrawal, "Intelligent Intrusion Detection using Artificial Bee Colony-Based Rule Discovery Techniques," 2025 IEEE Madhya Pradesh Section Conference (MPCON), Jabalpur, India, 2025, pp. 691-696, doi: 10.1109/MPCON66082.2025.11256592.

- L. C. Kasireddy, S. Paruchuri, C. Janakamma, A. Sarawat, K. C. Ravi and R. Kumar Chandu, "Cloud-Oriented IoT: Distributed Power-Aware Security Scheme with Data Integrity and Performance Enhancement," 2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS), Indore, India, 2025, pp. 1-6, doi: 10.1109/WorldSUAS66815.2025.11199185.

- J. L., L. Chandrakanth Kasireddy, R. V. Palanivel, G. Sushma, K. Bhimaavarapu and P. V. Reddy, "Predictive Modeling in Economics: The Role of AI and Deep Learning," 2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS), Indore, India, 2025, pp. 1-7, doi: 10.1109/WorldSUAS66815.2025.11199198.

- N. Soni, L. C. Kasireddy, T. S., C. Sinhgadiya, S. Kumar and A. T. S., "A Recurrent Neural Network Framework for Effective DDoS Attack Detection in Cloud Computing," 2025 2nd International Conference on Multidisciplinary Research and Innovations in Engineering (MRIE), Gurugram, India, 2025, pp. 594-598, doi: 10.1109/MRIE66930.2025.11156616.

- Jadhav, D., & Shinde, C. (2026). Sakhi: Stay safe stay fashionable. myresearchgo, 2(1), 1. https://doi.org/10.64448/myresearchgo.vol2.issue1.01.

- Jadhav, A. (2026). AI-enhanced employee management system. myresearchgo, 2(1), 8. https://doi.org/10.64448/myresearchgo.vol2.issue1.02.

- Rane, G., & Matteti, V. (2026). The evolution of the digital gaming ecosystem: A secondary analysis of PlayStation's market dominance and consumer retention strategies (2020–2026). Myresearchgo, 2(3), 1. https://doi.org/10.64448/myresearchgo.vol2.issue3.01

- Ansari, N., Sharma, A., & Yadav, S. (2026). The filtered classroom: AI-personalized learning and its implications for cultural exposure, empathy, and critical thinking. Myresearchgo, 2(3), 12. https://doi.org/10.64448/myresearchgo.vol2.issue3.02.

- Junghare, P., Chheniya, J., Behare, M., Kashte, P., Belekar, S., Dhoble, V., & Kumari, S. (2026). Google's Neural Memory Architecture: A Comprehensive Review of the Titans Framework. Myresearchgo, 2(4), 75. https://doi.org/10.64448/myresearchgo.vol2.issue4.12.