"Machine Learning and Big Data Analytics for Cybersecurity Threat Detection"

Ahmed. N.,(2021).

September 2024

1 Introduction

2 Summary of the Paper

2.1 Motivation/Purpose/Aims/Hypothesis

The paper is driven by the overwhelming volume and sophistication of cyber threats that challenge established cybersecurity techniques. The objective is to examine how cybersecurity threat detection capabilities can be improved by machine learning and big data analytics. By using massive amounts of data and complex algorithms to detect and react to threats more precisely, it is hypothesized that incorporating these cutting-edge techniques can increase the efficacy of threat detection systems.

3 Contribution

The paper offers a thorough review of machine learning techniques and big data analytics methods specifically tailored for cybersecurity.

- **Machine Learning Techniques**: Considered are several machine learning algorithms, including supervised learning (such as regression and classification), unsupervised learning (such as anomaly detection and clustering), and reinforcement learning.
- **Big Data Analytics:** Large-scale dataset handling and analysis methods, such as distributed computing, real-time analytics, and data mining, are reviewed.
- **Practical Insights**: By showing how these methods are applied in actual situations, case studies offer practical insights into the efficiency of these methods as well as their implementation issues.

- **Framework Proposal:** To improve the efficacy of current cybersecurity infrastructures, it suggests a conceptual framework for incorporating ML and big data analytics.

4 Methodology

- **Review of the Literature**: Information about ML and big data analytics techniques applied in cybersecurity is gathered by a thorough analysis of previously published research papers, publications, and reports.
- **Classification of Techniques:** The study discusses the theoretical foundations, advantages, disadvantages, and appropriateness of several machine learning and analytics techniques for a range of cyber risks.
- **Case Studies:** To demonstrate how the studied methodologies are applied, real-world case studies are examined and analyzed. These case studies offer proof of real-world implementation achievements and difficulties.
- **Comparative Analysis:** A number of criteria, including detection accuracy, false positive rates, computing efficiency, and scalability, are used to compare the performance of various methodologies. This comparison analysis facilitates comprehension of the relative benefits and drawbacks of each technique.

5 Conclusion

The study concludes that big data analytics and machine learning significantly enhance Cybersecurity threat detection by improving speed and accuracy through vast data and complex algorithms. However, it also highlights ongoing challenges, such as data protection, computational demands, and the need for constant adaptation. The study proposes a structured framework for integrating these technologies into existing cybersecurity systems.

6 Critiques or Limitations

6.1 1st Critique/Limitation

Data privacy and moral concerns: It's possible that the study may not discuss the privacy issues and ethical ramifications of gathering and analyzing vast amounts of data. In order to comply with laws like the CCPA and GDPR, cybersecurity applications that handle sensitive and personal data must properly address privacy concerns.

6.2 2nd Critique/Limitation

Scalability and Real-World Challenges: The study may not fully address how its proposed methodologies scale in diverse and dynamic real-world scenarios, where varying network architectures and inconsistent data quality could impact their effectiveness.

6.3 3rd Critique/Limitation

Focus on Established Techniques: The study may not go in-depth into new or cutting-edge technologies like quantum computing or sophisticated federated learning approaches, which may present new cybersecurity opportunities. Instead, it may concentrate mostly on established ML and big data techniques.

7 Synthesis

7.1 1st Potential/Idea of a New/Follow-Up/Extension paper

Efficiency and Optimization: Improving the computational effectiveness of machine learning models for extensive cybersecurity applications may be the subject of a future study. This could entail investigating techniques for hardware acceleration, model pruning, or distributed learning to improve real-time threat detection capabilities while controlling resource limitations.

7.2 2nd Potential/Idea of a New/Follow-Up/Extension Paper

Privacy-Preserving Methods: The creation and assessment of cybersecurity privacy-preserving machine learning methods may be a further topic for future study. This can require investigating methods such as homomorphic encryption, differential privacy, or secure multi-party computation to guarantee the security of sensitive data while permitting efficient threat detection.