SUPPLY CHAIN CHARACTERISTICS AS PREDICTORS OF CYBER RISK: A MACHINE-LEARNING ASSESSMENT

Beempaka Navyasri, M. Tejaswi
M.Tech Student, Assistant Professor
Department of Computer Science and Engineering
Brilliant Institute of Engineering & Technology, Abdullapur (V), Abdullapurmet (M), Rangareddy (D), Hyderabad - 501 513
Tejaswireddyteju@gamil.com

Abstract This project provides the first large-scale data-driven analysis to evaluate the predictive power of digital supply chain attributes for assessing risk of cyber attack data breaches. Motivated by rapid increase in the complexity of digital supply chains and related third-party enabled cyberattacks, the project provides the first quantitative empirical evidence that digital supply chain attributes are significant predictors of enterprise cyber risk. The analysis leverages externally observable cybersecurity ratings that aim to capture the quality of the enterprise internal cybersecurity management, but augments these with original supply chain features that are inspired by observed third-party cyberattack scenarios, as well as concepts from network science research. The main quantitative result of the project is to show that these supply chain network features add significant detection power to predicting enterprise cyber risk, relative to merely using enterprise-only attributes. In particular, compared to a base model that relies only on internal enterprise features, the supply chain network features improve the out-of-sample AUC by 2.3%. Given that a cyber data breach on a specific enterprise is a low probability high impact risk event, these improvements in the prediction power have significant value. Additionally, the model highlights several cybersecurity risk drivers related to third-party cyberattack and breach mechanisms and provides important insights as to what key metrics should be monitored and what interventions might be effective to mitigate these risks.

INTRODUCTION

The rapid growth of digital supply chains has increased enterprise exposure to cyber threats. Many cyberattacks originate through third-party vendors. Traditional cyber risk models focus only on internal enterprise security. This project presents a large-scale data-driven analysis of digital supply chain risk. Externally observable cybersecurity ratings are used to assess internal security posture. These are augmented with original supply chain network features. Network science concepts capture interdependencies among organizations. The model predicts the risk of cyber data breaches. Supply chain attributes significantly enhance predictive performance. Compared to enterprise-only models, AUC improves by 2.3 percent. This improvement is valuable for rare but severe cyber events. The approach provides empirical evidence of third-party risk importance. Feature analysis identifies key cyber risk drivers. The system highlights metrics for continuous monitoring. The methodology is systematic and reproducible. Out-of-sample testing ensures robustness. The framework improves cyber risk detection. It supports proactive mitigation strategies. The model aids enterprise decision-making. It helps prioritize cybersecurity investments. The approach is scalable. It integrates data science with cybersecurity. The project contributes to cyber analytics literature.

It bridges network science and risk management. The framework enhances predictive accuracy. The study demonstrates real-world relevance. The project supports cyber resilience initiatives. Overall, it presents an effective enterprise cyber risk prediction model.

Objectives of the study

The primary objective of this project is to evaluate the predictive power of digital supply chain attributes in assessing enterprise cyber risk. The project aims to develop a data-driven model to predict the likelihood of cyber data breaches. It focuses on understanding third-party enabled cyberattacks risks. The system integrates enterprise cybersecurity ratings with supply chain network features. Network science concepts are applied to model complex digital supply chains. The objective includes improving prediction accuracy for low-probability high-impact cyber events. Enhancing out-of-sample AUC is a key goal. The project aims to identify critical cyber risk drivers. It supports proactive risk monitoring and mitigation. The framework assists organizations in managing third-party risks. The model provides quantitative insights into cyber vulnerabilities. Automation of cyber risk assessment is emphasized. The approach improves decision-making for cybersecurity investments. The project contributes to cyber risk analytics research. Overall, it aims to strengthen enterprise cybersecurity resilience.

Project scope

This project focuses on predicting enterprise cyber risk using digital supply chain data. It analyzes cybersecurity risks arising from third-party relationships. The scope includes collecting externally observable cybersecurity ratings. Digital supply chain attributes are extracted and modeled. Network-based features inspired by real-world cyberattack scenarios are used. The project applies data-driven and statistical learning techniques. Enterprise-only attributes are used as a baseline. Supply chain network features are integrated for comparison. Model performance is evaluated using AUC metrics. Out-of-sample validation ensures generalization. The framework focuses on breach prediction. Real-time monitoring is outside the scope. The system operates in an offline analytical environment. The model addresses low-probability high-impact events. Scalability to large enterprise networks is considered. Visualization is not the primary focus. The scope includes feature importance analysis. Risk drivers are identified and interpreted. The project supports cybersecurity risk management strategies. Ethical data usage is assumed. The system is adaptable across industries. It supports third-party risk assessment. The project contributes empirical evidence to cybersecurity research. The scope includes documentation and result analysis. Deployment aspects are not covered. The model focuses on supervised learning. Regulatory compliance analysis is excluded. The framework aids strategic planning. The project improves predictive risk modeling. The study enhances cyber resilience understanding.

EXISTING SYSTEM

Existing systems rely primarily on internal enterprise cybersecurity metrics. Third-party risks are assessed qualitatively. Network relationships are not modeled. Predictive accuracy is limited. Risk detection is reactive rather than proactive.

Disadvantages

1. Ignores digital supply chain dependencies.
2. Limited prediction accuracy.

3. Poor detection of third-party risks.

4. Reactive cybersecurity posture.

PROPOSED SYSTEM

The proposed system integrates enterprise cybersecurity ratings with digital supply chain network features. Network science metrics are applied. A predictive model estimates breach risk. Out-of-sample validation ensures robustness. The system improves cyber risk detection.

Advantages

1. Captures third-party cyber risk.

2. Improved predictive accuracy.

3. Early risk identification.

4. Supports proactive mitigation.

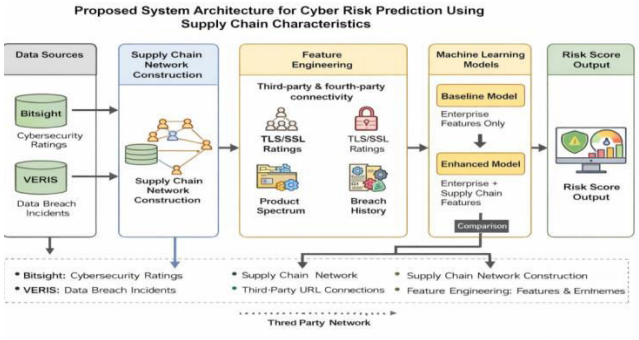
HARDWARE REQUIREMENTS

Processor	:	i3 or higher
Speed	:	2.9 GHz
RAM	:	4 GB (min)
Hard Disk	:	160 GB

SOFTWARE REQUIREMENTS

- **Operating system** : Windows 7 Ultimate
- **Coding Language** : Python
- **Back-End** : Django-ORM
- **Designing** : Html, css, javascript
- **Data Base** : MySQL (WAMP Server)

SYSTEM ARCHITECTURE



System Architecture

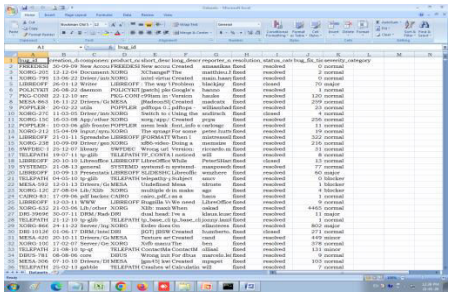
ALGORITHM USED

The system follows data collection, feature extraction, network modeling, predictive model training, and performance evaluation stages.

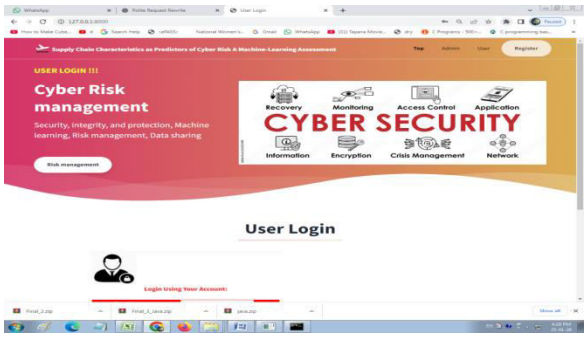
ALGORITHM TECHNIQUES

The project employs data-driven statistical learning and network science techniques. Supply chain attributes are modeled as network features. Supervised learning predicts cyber risk. AUC-based evaluation measures effectiveness. Feature importance analysis identifies risk drivers.

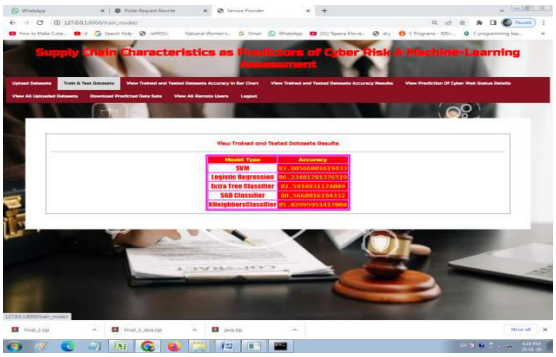
RESULTS AND ANALYSIS



Data Sets

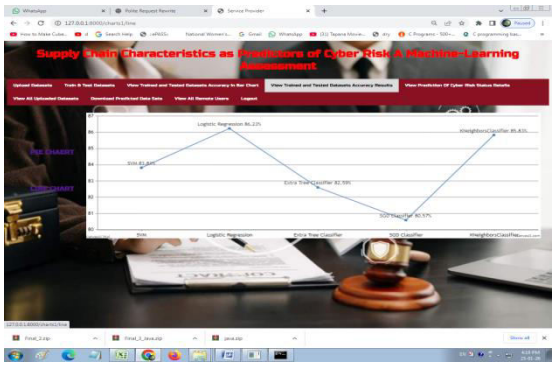


Home Page

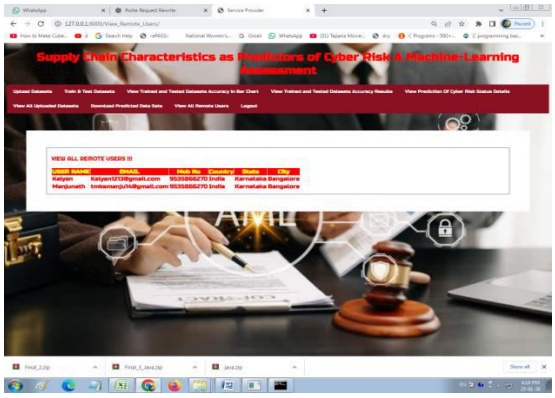


ML Accuracy

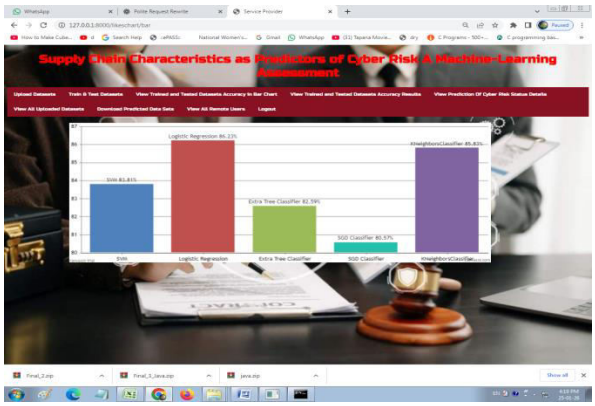
Login Page



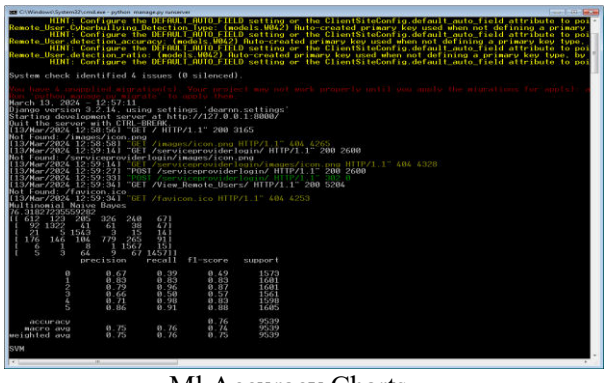
Graph Showing Ml Accuracy



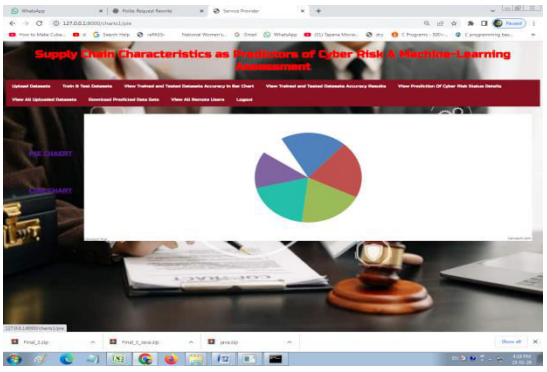
Admin Menu Operations



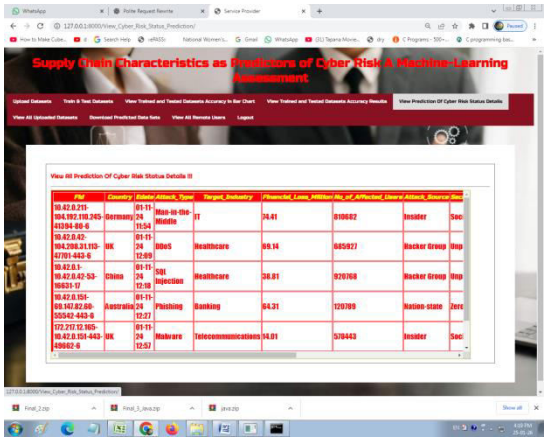
Line Graph Showing Ml Accuracy



Ml Accuracy Charts



Pie Graph Showing Ml Accuracy



List of Fares with Status

Prediction of Claims

Prediction Results

CONCLUSION

This project demonstrates the importance of digital supply chain attributes in predicting enterprise cyber risk. By integrating network-based supply chain features with cybersecurity ratings, predictive accuracy is significantly improved. The 2.3 percent AUC gain is critical for rare cyber breach events. The approach provides quantitative evidence of third-party risk impact. Network science concepts effectively capture digital dependencies. The model identifies key cyber risk drivers. It supports proactive monitoring and mitigation. The framework outperforms enterprise-only models. The methodology is scalable and robust. It contributes to cybersecurity analytics research. The project enhances cyber risk visibility. It aids strategic decision-making. Automation improves efficiency. The framework supports resilience planning. The approach is adaptable across sectors. The findings have practical relevance. The model improves early detection capability. It strengthens enterprise cyber defense strategies. Overall, the project offers a powerful cyber risk prediction solution.

REFERENCES

[1] M. A. Jamil, M. Arif, N. S. A. Abubakar, and A. Ahmad, ‘‘Software testing techniques: A literature review,’’ in Proc. 6th Int. Conf. Inf. Commun. Technol. Muslim World (ICT4M), Nov. 2016, pp. 177–182.

[2] W. Y. Ramay, Q. Umer, X. C. Yin, C. Zhu, and I. Illahi, ‘‘Deep neural network-based severity prediction of bug reports,’’ IEEE Access, vol. 7, pp. 46846–46857, 2019.

[3] W. Wen, ‘‘Using natural language processing and machine learning techniques to characterize configuration bug reports: A study,’’ M.S. thesis, College Eng., Univ. Kentucky, Lexington, KY, USA, 2017.

[4] J. Polpinij, ‘‘A method of non-bug report identification from bug report repository,’’ Artif. Life Robot., vol. 26, no. 3, pp. 318–328, Aug. 2021.

[5] S. Adhikarla, ‘‘Automated bug classification.: Bug report routing,’’ M.S. thesis, Fac. Arts Sci., Dept. Comput. Inf. Sci., Linköping Univ., Linköping, Sweden, 2020.

[6] K. C. Youm, J. Ahn, and E. Lee, ‘‘Improved bug localization based on code change histories and bug reports,’’ Inf. Softw. Technol., vol. 82, pp. 177–192, Feb. 2017.

[7] N. Safdari, H. Alrubaye, W. Aljedaani, B. B. Baez, A. DiStasi, and M. W. Mkaouer, ‘‘Learning to rank faulty source files for dependent bug reports,’’ in Proc. SPIE, vol. 10989, 2019, Art. no. 109890B.

[8] A. Kukkar, R. Mohana, A. Nayyar, J. Kim, B.-G. Kang, and N. Chilamkurti, ‘‘A novel deep-learning-based bug severity classification technique using convolutional neural networks and random forest with boosting,’’ Sensors, vol. 19, no. 13, p. 2964, Jul. 2019.

[9] A. Aggarwal. (May 2020). Types of Bugs in Software Testing: 3 Classifications With Examples. [Online]. Available: https://www.scnsoft.com/software-testing/types-of-bugs

[10] A. Kukkar and R. Mohana, ‘‘A supervised bug report classification with incorporate and textual field knowledge,’’ Proc. Comput. Sci., vol. 132, pp. 352–361, Jan. 2018.