UNIVERSITI TEKNOLOGI MARA

CSP760

PREDICTING VULNERABILITY SUSCEPTIBILITY IN MALAYSIAN BANK USING SUPERVISED MACHINE LEARNING

STUDENT NOR ADANI BINTI KAMAL MOHAMAD NASIR (2024782087)

SUPERVISOR DR SITI ARPAH BINTI AHMAD

Date: 23 November 2025

UNIVERSITI TEKNOLOGI MARA

TABLE OF CONTENT

Paper	Dataset Source	Algorithm	Key Finding	Justification
Oliveira et al., 2023 – *Hybrid ML for Vulnerability Detection in Web Apps*	Web app vulnerability logs (CVE-related)	Random Forest, SVM, Naive Bayes	RF performs strongly for vulnerability detection; hybrid supervised models increase accuracy.	Supports RF as a strong classifier for structured vulnerability data.
Williams et al., 2020 – *Vulnerability Analysis & Prediction Framework*	Historical vulnerability metadata & severity scores	Artificial Neural Networks (ANN), Regression models	ANN delivers strong performance in predicting vulnerability severity.	Directly supports the use of Neural Network Regression for predictive scoring.
Asmar & Tuqan, 2024 – *ML for Cybersecurity in Digital Banks*	Digital banking cybersecurity datasets (financial institutions)	SVM, RNN, HMM, LOF	ML improves detection accuracy in banking security environments.	Financial institution context —ML is suitable for Malaysian banks.
Yosifova et al., 2021 – *Predicting Vulnerability Type in CVE Database*	**CVE Details dataset - 122,774 CVE records (expanded to 162,789 after exploding categories)**	Random Forest, Naive Bayes, Linear SVM	RF performs strongly on CVE classification.	Directly supports RF for CVE-based training, matching my dataset (NVD + Tenable + ExploitDB).
Sharadhi et al., 2022 – *ML Algorithms for Software Defect Prediction*	PROMISE defect dataset (NASA) - More than 40 datasets (sizes vary by module)	ANN, RF, SVM, Decision Trees	RF & ANN repeatedly outperform other ML algorithms.	Supports choosing RF + NN as my primary models due to strong precision in defect/cybersecurity prediction tasks.
Adani Kamal, 2025 – *Predicting Vulnerability Susceptibility in Malaysian Banks Using Supervised ML* (Your Study)	**Merged dataset: NVD + Tenable + ExploitDB**	Random Forest + Neural Network Regression	Combines global vulnerability metadata with exploit information + scanner data to build predictive model for banks.	Extends prior work by integrating multi-source real-world datasets to target Malaysian banking vulnerability susceptibility.

Paper / Framework	Asset Types Listed
FFIEC IT Handbook *(Financial Sector)*	Transaction systems, Settlement & clearing systems, Customer information databases, Identity systems, Network & security infrastructure
BNM RMiT (Bank Negara Malaysia) *(Regulatory Framework)*	Core Banking System (CBS), Payment systems (FPX, IBG, RENTAS, DuitNow), Internet/mobile banking, Fraud monitoring systems, ATM & card systems, Customer information systems (CIF), Third-party vendor systems, Cloud workloads used for banking
NIST SP 1800-25 (Financial Services)	Treasury systems, Fund transfer systems, Card processing systems, API gateways, IAM / MFA systems, network device such as (Firewalls, routers, switches)
SWIFT CSCF *(Banking Payment Network)*	SWIFT interface servers, HSM (Hardware Security Modules), Payment messaging servers, Secure communication nodes
ISO/IEC 27001 (General Framework)	Information assets, Servers & databases, Security appliances, Network devices, End-user devices

CSP760

THANK YOU

(ADANI,2025)

SUMMARY PROJECT

Problem	Research Questions	Hypothesis	Objectives	Scope	Significance
Malaysian banks rely heavily on **Vulnerability Assessment** (VA) tools such as Nessus, Qualys, and OpenVAS that only detect *existing* **vulnerabilities** but cannot *predict future threats*. This leaves financial institutions unprepared for advanced cyberattacks.	**RQ1:** What are the limitations of current **vulnerability assessment** tools in predicting future attack trends? **RQ2:** To what extent can machine learning models improve **vulnerability** prioritization and remediation by forecasting attack patterns? **RQ3:** How effective is the proposed ML model compared to traditional VA methods in predicting **vulnerability** susceptibility?	**H1:** Machine Learning algorithms can enhance the predictive capabilities of **vulnerability assessments**. **H2:** Random Forest (RF) and Neural Network Regression (NNR) are suitable models for forecasting **vulnerability susceptibility**.	1. To design and propose a machine learning model for predicting cyberattack susceptibility using **vulnerability Assessment** data. 2. To develop and implement a supervised ML-based predictive system using NVD + ExploitDB + Tenable datasets. 3. To evaluate the accuracy and performance of the model in forecasting potential vulnerabilities.	- **Data:** **Vulnerability** from NVD CVE vulnerability, exploitDB, Tenable datasets. - **Tools:** Orange Data Mining, Python (Jupyter Notebook). - Models: Random Forest (classification), Neural Network Regression (regression). - Evaluation: Accuracy, Precision, Recall, F1-score, MAE, RMSE.	The study enhances the cybersecurity posture of Malaysian banks by shifting from reactive to **predictive vulnerability** management. It supports AI-driven, enables faster remediation prioritization.