PREDICTING DIGITAL FRAUD RISK USING SUPPORT VECTOR MACHINE CLASSIFIER: A CASE STUDY OF UNIVERSITAS TERBUKA STUDENTS

Fonda Leviany[1], Kurnia Sari Kasmiarno[2], Ika Nur Laily Fitriana[3]

[1] Data Science Study Program, Faculty of Science and Technology, Universitas Terbuka, Indonesia
[2] Islamic Economics Study Program, Faculty of Economics and Business, Universitas Terbuka, Indonesia
[3] Statistics Study Program, Faculty of Science and Technology, Universitas Terbuka, Indonesia
Corresponding author: fonda.leviany@ecampus.ut.ac.id

Abstract

As a higher education institution that operates through open and distance learning, Universitas Terbuka depends significantly on digital infrastructure to support both its academic activities and student financial transactions. Students, as active users of digital technology are also vulnerable to various forms of digital fraud, such as phishing, identity theft, and scams related to social media or financial applications. Early detection of the risk of digital fraud is essential, allowing institutions to implement preventive strategies and deliver more targeted education on digital security. This study aims to develop a machine learning predictive model to identify the potential risk of digital fraud among students at Universitas Terbuka. A supervised machine learning approach using multiclass Support Vector Machine (SVM) was applied to a dataset collected from UT students who were actively studying and had reported income or employment status. Relevant features were selected and pre-processed to train the SVM model for multi-class classification of digital fraud risk levels. The model demonstrated a good classification performance, indicating that machine learning can effectively support early detection systems for digital fraud in open and distance learning environments. This study contributes to the emerging literature on fraud detection in the education sector by applying a Support Vector Machine model tailored to the context of a large-scale distance learning university.

Keywords: Digital Fraud, Predicting, Students, Support Vector Machine, Universitas Terbuka.

Introduction

The internet has become an inseparable part of modern life, especially since the COVID-19 pandemic accelerated digitalization across various sectors. As of January 2021, there were 4.66 billion active internet users worldwide, accounting for approximately 59.5% of the global population, with 92.6% accessing the internet via mobile phones (Zhang, et al., 2025). Alongside the rise in digital activity, cybersecurity risks have also increased. One of the most common threats is phishing—fraudulent attempts where perpetrators impersonate trusted entities to obtain sensitive information from victims (Jain & Gupta, 2017). Sectors such as online payment systems, e-commerce, and social media are particularly vulnerable to these attacks. Many users fall victim due to a lack of understanding of URL structures, impulsive behavior, and the increasingly convincing appearance of phishing websites that closely resemble legitimate ones (Zhang, et al., 2025; Volkamer, Renaud, Reinheimer, & Kunz, 2017). This trend is further supported by reports from the Anti-Phishing Working Group (APWG), which recorded a surge in phishing attacks from around 200,000 cases in April 2021 to nearly 385,000 cases in less than a year (Zhang, et al., 2025).

Students, as part of the productive age group and active users of digital technology, are increasingly vulnerable to various forms of digital fraud, including phishing, misuse of personal data, and scams through social media and financial apps. This issue is becoming more relevant considering platforms like Instagram, TikTok, and X—widely used by younger generations—have become prime targets for digital criminals. PhishLab reported a 103% increase in phishing attacks on Instagram in 2021 (Qahri-saremi, 2023). Meanwhile, APWG (2023) stated that 22.3% of all phishing attacks in the second quarter of 2023 targeted social media platforms. Demographically, more than 61.1% of Instagram users are aged 18–34, and TikTok reports that 68.3% of its users fall within the same age group (Mouncey & Ciobotaru, 2025). Therefore, students—who generally fall within this age range—require improved digital literacy and early detection capabilities against digital crimes. Low awareness and limited ability to recognize phishing patterns on social media can make students easy targets for cybercriminals.

However, efforts to mitigate digital fraud risks remain limited and largely reactive. Early detection of digital fraud risks is crucial for educational institutions to take preventive measures and provide more targeted digital security education. Most research on digital crime has focused on evaluating user behavior after falling victim to phishing, as seen in studies by Bera and Dan (2025) and Molinaro and Matthew (2018). These studies are post-event and do not offer predictive approaches that could be used preventively. On the other hand, some research has begun to explore early detection of digital fraud risks, although most still rely on conventional statistical methods. For example, Butavicius et al. (2022) used ANOVA and linear regression techniques to analyze factors influencing phishing detection ability. Meanwhile, Lee et al. (2025) developed a phishing detection system based on ensemble learning with high accuracy. However, their research focused on technical analysis of general phishing emails, without considering individual user characteristics or demographic factors that may influence digital fraud risk. Thus, the approach does not address personal dimensions or the context of open and distance higher education, which features diverse demographics and unique digital vulnerabilities.

Furthermore, a study by Birthriya et al. (2025) attempted to fill this gap by combining classical statistical approaches such as logistic regression with machine learning algorithms like random forest to predict the likelihood of individuals falling victim to phishing based on psychographic and demographic data. Although this approach is more personalized than previous studies, it has not been directed toward building a detection system specifically for students in online higher education contexts, where students heavily rely on digital technology for learning, communication, and financial transactions. This condition makes students highly vulnerable to various forms of digital fraud. Therefore, a significant gap remains in the literature—namely, the absence of studies that specifically develop early detection models for digital fraud using machine learning, particularly with the Support Vector Machine (SVM) algorithm, tailored to the context of open and distance higher education institutions. To the best of the researcher's knowledge, this is the first study to apply the SVM approach to detect potential digital fraud risks in large-scale online learning environments such as those operated by Universitas Terbuka.

This study proposes a machine learning method using the Support Vector Machine (SVM) Classifier, which is known for its effectiveness in data classification and detecting non-linear data patterns (Rizki & Darip, 2025). SVM also demonstrates better performance in data classification compared to other methods (Maulana, Fahmi, Imran, & NutrianaHidayati, 2024). The SVM algorithm was used by Eldo et al. to detect fraud in online transactions, where the data consisted of legitimate and fraudulent transactions. The results showed that SVM could detect fraudulent transactions with a high accuracy rate of up to 95% (Eldo, Ayuliana, Suryadi, Chrisnawati, & Judijanto, 2024). SVM also outperformed the Neural Network (NN) method in predicting stock prices saham (Gunawan, Putri, & Kurniawan, 2024), surpassed the Decision Tree method in sentiment analysis of text data and breast cancer classification (Rokhman, Berlilana, & Arsi, 2021; Imaduddin, Hermansyah, & Salsabilla, 2021)and proved more reliable and effective than logistic regression in predicting financial data (Qin, 2021).

In this study, several variables are used to build a predictive model for digital fraud risk among Universitas Terbuka students, who are part of an open and distance learning higher education institution. These variables include financial literacy scores and digital financial literacy scores, which reflect students' understanding of financial management and digital security, as well as demographic variables such as education, gender, age, occupation, income, marital status, and work experience. Data were collected from active students with employment or income status and then used in a machine learning approach based on Support Vector Machine (SVM) to classify the level of digital fraud risk. By considering these factors, the model aims to enable early risk detection so that the university can develop preventive strategies and provide more targeted digital security education.

This study aims to develop a predictive model capable of accurately and adaptively identifying potential digital fraud among students. The results of this research are expected to serve as a foundation for developing policies to enhance digital security literacy within open higher education environments, particularly among students.

Methods

This study utilized data collected from 450 Universitas Terbuka (UT) students through a structured questionnaire designed to measure indicators relevant to predicting digital fraud risk. The indicators include financial literacy score, digital financial literacy score, highest education level, gender, age, work profile, income, marital status, and work experience. The variable used in this study can be seen in Table 1. The questionnaire was validated and deemed reliable, supported by significant Pearson correlation values and a Cronbach's Alpha of 0.665. According to Sugiyono, a Cronbach's Alpha value above 0.6 indicates acceptable reliability.

Table 1 Operational Variables

Symbol	Name of Variable	Type of Data
Y	Digital Fraud Risk	Categoric with 3 levels: • Low • Medium • High
X_1	Financial Literacy Score	Integer
X_2	Digital Financial Literacy Score	Integer
X_3	Highest Education Level	Categoric with 4 levels: • Secondary • Diplomas • Bachelors • Masters

Symbol	Name of Variable	Type of Data
X_4	Gender	Categoric with 2 levels: • Male • Female
X_5	Age	Integer
X_6	Work Profile	Categoric with 9 levels: • Employee • Self-employed • Teacher • Civil servant • Freelancer • Admin • Laborer • Internship • Others
X_7	Income	Integer
X_8	Marital Status	Categoric with 3 levels: • Unmarried • Divorced • Married
X_9	Professional Experience	Integer

Prior to analysis, the data underwent preprocessing steps such as handling missing values, transforming variables, and standardizing the dataset. Descriptive statistics were then applied to explore the data distribution and characteristics. The predictive modeling was conducted using the Support Vector Machine (SVM) classifier to categorize students based on their risk level of becoming victims of digital fraud. SVM is a classification technique that identifies the optimal hyperplane to separate data classes, where the best hyperplane is defined by the largest margin—the shortest distance between the data points and the hyperplane. The data points closest to the hyperplane are known as support vectors (James, Witten, Hastie, & Tibshirani, 2013; Amelia, Soleh, & Rahardiantoro, 2022).

The stages of this research are carried out as follows:

1. Data collection, from Universitas Terbuka students.
2. Data preprocessing, including checking for missing values, transforming character data type into categorical format, and standardizing the data.
3. Exploratory data analysis, using descriptive statistics to understand the data distribution and characteristics.
4. Splitting the dataset, into training and testing sets with various ratios (90:10, 80:20, 70:30, 60:40, and 50:50).
5. Modeling using Support Vector Machine (SVM) with a one-against-one approach for both linear and non-linear models, followed by performance evaluation.
6. Selecting the best-performing SVM model based on evaluation metrics.
7. Predicting digital fraud potential using the selected model.
8. Providing recommendations and offering suggestions for targeted digital education initiatives.

Results and Discussions

The dataset underwent a preprocessing procedure, which confirmed the absence of missing values. Additionally, six predictor variables initially formatted as character types were systematically converted into categorical variables and subsequently encoded into numerical form. Following this transformation, data standardization was performed to ensure uniformity across all variables for subsequent analysis.

Table 2 Original Data

i	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
1	Medium	6	5	Secondary	Female	24	Admin	1200000	Unmarried	13
2	High	5	4	Secondary	Female	22	Admin	1500000	Unmarried	24
3	Low	4	5	Secondary	Female	19	Admin	3000000	Unmarried	3
	…	…	…	…	…	…	…	…	…	…
450	High	6	5	Bachelors	Male	28	Self-employed	8000000	Unmarried	60

Table 3 Standardized Data

i	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
1	Medium	0.012	1.132	Secondary	Female	-0.055	Admin	-0.766	Unmarried	-0.531
2	High	-1.080	-0.006	Secondary	Female	-0.468	Admin	-0.679	Unmarried	-0.277
3	Low	-2.172	1.132	Secondary	Female	-1.087	Admin	-0.243	Unmarried	-0.762
	…	…	…	…	…	…	…	…	…	…
450	High	0.012	1.132	Bachelors	Male	0.770	Self-employed	1.211	Unmarried	0.555

The pie chart presents the percentage distribution of dependent variables associated with digital fraud risk among students. It categorizes individuals into three risk levels: low, middle, and high. The largest portion of the chart, shaded in dark orange, represents the low-risk group, indicating that the majority of respondents fall into this category. A smaller segment, colored in medium orange, corresponds to the middle-risk group, while the smallest portion, shown in light orange, represents the high-risk group. This pie chart suggests that although most students are considered to have a low potential for becoming victims of digital fraud, a notable proportion still falls into moderate and high-risk categories, highlighting the need for targeted digital security education.

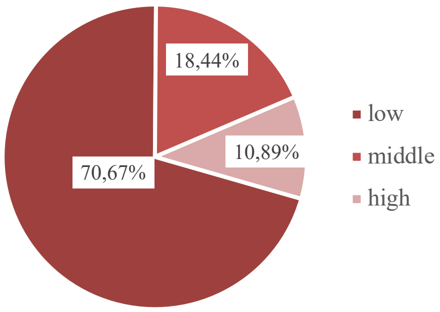

Figure 1 The Distribution of Digital Fraud Risk Among Students

To evaluate the performance of the Support Vector Machine (SVM) model, the dataset consisting of 450 observations was divided into training and testing sets using various proportions. Five different splitting ratios were applied: 90:10, 80:20, 70:30, 60:40, and 50:50. Under the 90:10 split, 405 records were used for training and 45 for testing. The 80:20 split allocated 359 records for training and 91 for testing. The 70:30 configuration used 315 training records and 135 testing records, while the 60:40 split involved 270 training and 180 testing records. Lastly, the 50:50 split evenly distributed the data, with 225 records each for training and testing. Among these configurations, the selected data partition was the one that yielded the best performance on the testing data, ensuring optimal model generalization and predictive accuracy.

Table 4 Data Split

Split Ratio	Training Data	Testing Data
90:10	405	45
80:20	359	91
70:30	315	135
60:40	270	180
50:50	225	225

Table 5 Accuracy Performance of SVM Models Across Different Data Splits and Kernel Type

Split	Linear		Polynomial		RBF		Sigmoid	
Ratio	Train	Test	Train	Test	Train	Test	Train	Test
90:10	0.7111	0.6889	0.7111	0.7111	0.7062	0.7111	0.684	0.6889
80:20	0.7103	0.7033	0.7131	0.7033	0.7075	0.7033	0.688	0.6813
70:30	0.7111	0.7037	0.7111	0.7037	0.7079	0.7037	0.6889	0.6815
60:40	0.7074	0.7056	0.7074	0.7056	0.7074	0.7056	0.6852	0.6889
50:50	0.7067	0.7067	0.7067	0.7067	0.7067	0.7067	0.6889	0.6978
Average	0.7093	0.7016	0.7099	0.7061	0.7071	0.7061	0.6870	0.6877

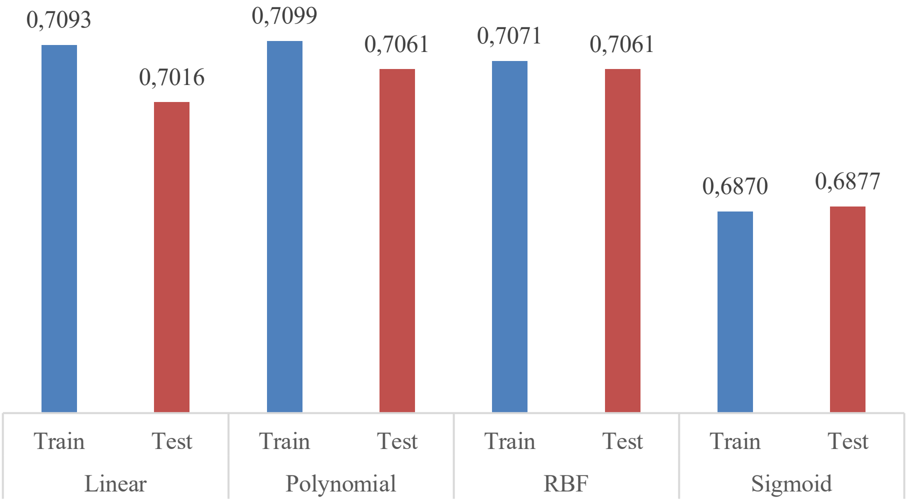

Figure 2 SVM Multiclass Accuracy Comparison Based on Data Split

The performance evaluation of the multiclass SVM model across different data split ratios and kernel types reveals that the polynomial and RBF kernels consistently achieved the highest testing accuracy, particularly under the 90:10 split ratio, where both reached an accuracy of 0.7111. This supports the notion that a larger training dataset contributes to better model generalization and predictive performance, as suggested by Nugroho (2022). Among these, the polynomial kernel slightly outperformed RBF in terms of average testing accuracy across all splits (0.7061 vs. 0.7061, with polynomial showing more consistent results). This finding aligns with previous studies indicating that the polynomial kernel often yields superior classification performance compared to RBF in certain contexts (Muflikhah, Haryanto, Soebroto, & Santoso, 2018).

Table 6 Confusion Matrix for Training Data

Actual	Prediction		
	Low	Medium	High
Low	286	0	0
Medium	74	1	0
High	43	0	1

Table 7 Confusion Matrix for Testing Data

Actual	Prediction		
	Low	Medium	High
Low	32	0	0
Medium	8	0	0
High	5	0	0

The results of the Support Vector Machine (SVM) classifier in predicting digital fraud risk levels among Universitas Terbuka students show a strong tendency toward classifying observations into the low-risk category. As seen in the confusion matrix for the training data, the model correctly classified 286 students as low risk, while misclassifying 74 medium-risk and 43 high-risk individuals into the same category. Only one observation each was correctly classified as medium and high risk, indicating a significant imbalance in classification performance across categories. A similar pattern is observed in the testing data, where 32 students were correctly predicted as low risk, but 8 medium-risk and 5 high-risk individuals were also misclassified as low risk. No observations were classified into the medium or high categories during testing.

These results suggest that while the model performs well in identifying low-risk individuals, it struggles to distinguish between medium and high-risk categories, likely due to class imbalance in the dataset. This limitation is common in multiclass classification problems, especially when one class dominates the data distribution. Despite this, the overall testing accuracy achieved by the polynomial and RBF kernels under the 90:10 split ratio remains the highest, indicating that these kernels are more effective in generalizing the model to unseen data.

Given the model's tendency to favor the majority class, future research should consider techniques such as resampling (oversampling minority classes or undersampling the majority class, class weighting, or ensemble methods to improve classification performance across all risk levels. Additionally, incorporating parameter tuning and feature importance analysis may further enhance the model's ability to detect digital fraud risk more accurately and equitably across diverse student profiles.

Conclusion

The model demonstrated a good classification performance, indicating that machine learning can effectively support early detection systems for digital fraud in open and distance learning environments. This study contributes to the emerging literature on fraud detection in the education sector by applying a Support Vector Machine (SVM) classifier tailored to the context of a large-scale distance learning university, namely Universitas Terbuka.

Despite achieving high accuracy, particularly with the polynomial and RBF kernels under the 90:10 data split, the model showed a tendency to overpredict the low-risk category, as reflected in the confusion matrix. This highlights a limitation in distinguishing medium and high-risk individuals, likely due to class imbalance in the dataset. To address this, future research is recommended to apply resampling techniques like SMOTE to balance the class distribution. Implement hyperparameter tuning to optimize SVM performance, especially for the polynomial kernel is recommended. Also, expand the dataset and include additional behavioral or psychographic variables to improve model robustness. These recommendations aim to enhance the predictive capability and practical utility of machine learning models in safeguarding students from digital fraud in increasingly digital learning environments.

Acknowledgement

The authors would like to express their sincere gratitude to Universitas Terbuka for the support and opportunity to conduct this research. Special thanks are extended to the students who participated in the survey, whose responses provided valuable data for the development of the predictive model. We also acknowledge the contributions of colleagues and academic mentors who provided insightful feedback throughout the research process.

This study would not have been possible without the collaborative efforts of the research team. We are especially grateful for the technical and academic support that enabled the successful implementation of the Support Vector Machine classifier in analyzing digital fraud risk within the context of open and distance higher education.

References

Amelia, O. D., Soleh, A. M., & Rahardiantoro, S. (2022). Pemodelan Support Vector Machine Data Tidak Seimbang Keberhasilan Studi Mahasiswa Magister IPB. *Xplore, 2*(1), 33-40.

Bera, D., & Kim, D. J. (2025). The nexus of mindfulness, affect, and information processing in phishing identification: An empirical examination. *Information & Management, 62*, 1-27. doi:https://doi.org/10.1016/j.im.2025.104110

Birthriya, S. K., Ahlawat, P., & Jain, A. K. (n.d.). Intelligent phishing website detection: A CNN-SVM approach with nature-inspired hyperparameter tuning. *Cyber Security and Applications*. doi:https://doi.org/10.1016/j.csa.2025.100100

Butavicius, M., Taib, R., & Han, S. J. (2022). Why people keep falling for phishing scams: The effects of time pressure and deception cues on the detection of phishing emails. *Computers & Security*, 1-10. doi:https://doi.org/10.1016/j.cose.2022.102937

CRAN R. (2024). *Package 'e1071'.* Retrieved from CRAN R: https://cran.r-project.org/web/packages/e1071/e1071.pdf

Eldo, H., Ayuliana, Suryadi, D., Chrisnawati, G., & Judijanto, L. (2024). Penggunaan Algoritma Support Vector Machine (SVM) untuk Deteksi Penipuan pada Transaksi Online. *Jurnal Minfo Polgan*, 1627-1632.

Gunawan, Putri, A. R., & Kurniawan, R. D. (2024). Komparasi Penerapan Algoritma SVM dan NN dalam Memprediksi IHSG. *Jurnal BATIRSI*, 1-5.

Imaduddin, H., Hermansyah, B. A., & Salsabilla, F. A. (2021). COMPARISON OF SUPPORT VECTOR MACHINE AND DECISION TREE METHODS IN THE CLASSIFICATION OF BREAST CANCER. *Cyberspace: Jurnal Pendidikan Teknologi Informasi*, 22-30.

Jain, A. K., & Gupta, B. B. (2017). Phishing Detection: Analysis of Visual Similarity Based Approaches. *Hindawi: Security and Communication Networks*, 1-20. doi:https://doi.org/10.1155/2017/5421046

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learningwith Applications in R.* New York: Springer.

Karim, A. (2020). Perbandingan Prediksi Kemiskinan di Indonesia Menggunakan Support Vector Machine (SVM) dengan Regresi Linear. *Jurnal Sains Matematika dan Statistika, 6*(1), 107-112.

Lee, C., Kim, B., & Kim, H. (2025). The silence of the phishers: Early-stage voice phishing detection with runtime permission request. *Computers & Security, 152*, 1-15. doi:https://doi.org/10.1016/j.cose.2025.104364

Maulana, B. A., Fahmi, M. J., Imran, A. M., & NutrianaHidayati. (2024). Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive BayesdanSupport Vector Machine (SVM). *MALCOM: Indonesian Journal of Machine Learning and Computer Science, 4*(2), 375-384.

Molinaro, K. A., & Bolton, M. L. (2018). Evaluating the applicability of the double system lens model to the analysis of phishing email judgments. *Computers & Security, 77*, 128-137. doi:https://doi.org/10.1016/j.cose.2018.03.012

Mouncey, E., & Ciobotaru, S. (2025). Phishing scams on social media: An evaluation of cyber awareness education on impact and effectiveness. *Journal of Economic Criminology, 7*, 1-10. doi:https://doi.org/10.1016/j.jeconc.2025.100125

Nugroho, A. (2022). Analisa Splitting Criteria Pada Decision Tree dan Random Forest untuk Klasifikasi Evaluasi Kendaraan. *JSITIK: Jurnal Sistem Informasi dan Teknologi Informasi Komputer, 1*(1), 41-49. doi:https://doi.org/10.53624/jsitik.v1i1.154

OCC USA. (2025). *Online and Digital Scams*. Retrieved from OCC USA: https://www.occ.gov/topics/consumers-and-communities/consumer-protection/fraud-resources/online-and-digital-scams.html

Qin, R. (2021). Identification of Accounting Fraud Based on Support Vector Machine and Logistic Regression Model. *Complexity*, 1-11. doi:https://doi.org/10.1155/2021/5597060

Rizki, W., & Darip, M. (2025). PENGGUNAAN ALGORITMA SUPPORT VECTOR MACHINE UNTUK MENDETEKSI ANOMALI AKTIVITAS PENGGUNA PADA SISTEM INFORMASI KEUANGAN PT. DIGIDOKAT INDONESIA. *JATI: Jurnal Mahasiswa Teknik Informatika*, 4538-4546.

Rokhman, K. A., Berlilana, & Arsi, P. (2021). PERBANDINGAN METODE SUPPORT VECTOR MACHINE DAN DECISION TREE UNTUK ANALISIS SENTIMEN REVIEW KOMENTAR PADA APLIKASI TRANSPORTASI ONLINE. *JOISM : JURNAL OF INFORMATION SYSTEM MANAGEMENT*, 1-7.

Tantika, R. S., & Kudus, A. (2022). Penggunaan Metode Support Vector Machine Klasifikasi Multiclass pada Data Pasien Penyakit Tiroid. *Bandung Conference Series: Statistics*, *2 (2)*, pp. 159-166. Bandung. doi:https://doi.org/10.29313/bcss.v2i2.3590

Volkamer, M., Renaud, K., Reinheimer, B., & Kunz, A. (2017). User experiences of TORPEDO: TOoltip-poweRed Phishing Email DetectiOn. *Science Direct: Computers & Security*, 100-113.

Zhang, K., Wang, H., Chen, M., Chen, X., Liu, L., Geng, Q., & Zhou, Y. (2025). Leveraging machine learning to proactively identify phishing campaigns before they strike. *Journal of Big Data, 12*(124), 1-55.

Authors' Bibliography

Fonda Leviany, Kurnia Sari Kasmiarno, and Ika Nur Laily Fitriana are currently serving as lecturers at Universitas Terbuka, each specializing in different academic fields: Fonda in data science, Kurnia in sharia economics, and Ika in statistics.Fonda was born in Klungkung on September 28th, 1997; Kurnia was born in Surabaya on June 9th, 1994; and Ika was born in Jombang on June 15th, 1997.

Fonda and Ika both hold a Bachelor of Statistics (S.Stat.) and a Master of Statistics (M.Stat.) from Institut Teknologi Sepuluh Nopember, located in Surabaya, Indonesia. Kurnia earned her Bachelor's and Master's degrees in Sharia Economics from Universitas Airlangga, also in Surabaya.

Fonda's research interests include machine learning applications in education, particularly in the context of higher education and digital fraud detection. Kurnia's research centers on Islamic finance, digital economic behavior, and financial literacy in online learning environments, while Ika's research focuses on statistical modeling, data analysis, and educational data mining.