

Artificial Intelligence Model for Detecting Tax Evasion Involving Complex Network Schemes

Nuryani^{1*}, Achmad Benny Mutiara², I Made Wiryana³, Detty Purnamasari⁴, Souza Nurafrianto
Windiartono Putra⁵

^{1, 2, 3, 4}Dept. of Information Technology, Universitas Gunadarma, Indonesia

¹Dept. of Information and Communication Technology, Directorate General of Taxation, Indonesia

⁵Dept. of Information Technology, Universität des Saarlandes, Germany

¹yeyen07@gmail.com, ¹nuryani70@kemenkeu.go.id, ²amutiara@staff.gunadarma.ac.id, ³mwiryana@staff.gunadarma.ac.id,

⁴detty@staff.gunadarma.ac.id, ⁵s.windiartono@uni-saarland.de

*Corresponding Author

Article Info

Article history:

Submission June 19, 2024

Revised July 29, 2024

Accepted September 05, 2024

Published September 09, 2024

Keywords:

Artificial Intelligence

Big Data

Machine Learning

Social

Network Analysis



ABSTRACT

Tax evasion through complex network schemes poses a significant challenge to tax authorities, leading to substantial revenue losses. **This paper aims to** develop and evaluate an artificial intelligence model designed to detect tax evasion within complex corporate networks, providing a comprehensive overview and prediction of tax avoidance behaviors. **Employing a systematic** literature review and document analysis of applicable tax regulations, the study utilizes Social Network Analysis (SNA) as a primary technique for mapping and analyzing taxpayer networks. The process involves matching taxable identities, constructing taxpayer graphs, extracting features, and developing a machine learning model. **The proposed** architectures and processes demonstrate the potential for tax authorities to enhance their capabilities in detecting tax evasion involving complex networks, with the machine learning model effectively identifying features related to both individual and network characteristics of taxpayers. **The findings** suggest that the integration of artificial intelligence and big data analytics can significantly improve the detection of tax evasion in complex corporate structures, offering valuable tools for tax authorities to better enforce tax compliance.

This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.



DOI: <https://doi.org/10.34306/att.v6i3.436>

This is an open-access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>)

©Authors retain all copyrights

1. INTRODUCTION

The development of information technology is not only used by the private sector but also by government agencies to enhance operations and public services. The Directorate General of Taxes (DGT) has integrated technology into its tax system for services, supervision, law enforcement, and data collection. One ongoing challenge is tax avoidance and evasion by non-compliant taxpayers, which results in significant revenue losses for the state.

Tax evasion through related party transaction schemes is particularly damaging. According to the OECD, developing countries lost 4.4% of their total tax revenue to such schemes. In 2019, Indonesia relied on taxes for 83% of its revenue, making evasion through complex networks a critical issue for the Ministry of Finance and the DGT [1]. These practices not only reduce revenue but also undermine tax fairness.

Tax avoidance and evasion strategies are becoming more sophisticated. One of the main areas of concern for tax authorities is the use of corporate networks to facilitate these schemes. Tax evasion through related party transactions often appears as legitimate business dealings between affiliated companies, making detection difficult [2].

These schemes are complex, diverse, and involve hidden interactions aimed at evading tax obligations. This growing complexity requires tax authorities to continuously evolve their strategies to combat evasion effectively. The rapid advancement of information technology has become essential not only for the private sector but also for government agencies to enhance public services and operational efficiency [3]. For tax authorities like the Directorate General of Taxes (DGT), integrating information technology into tax systems is crucial for improving service delivery, supervision, law enforcement, and tax data collection [4]. However, challenges such as tax avoidance and evasion continue to pose significant threats to state revenue, with tax evasion through complex related-party transaction schemes being particularly problematic [5]. The OECD estimated that developing countries lost 4.4% of total tax revenues due to such schemes in 2016, making it a critical issue for countries like Indonesia, where tax revenue is vital, accounting for 83% of state revenue in 2019 [6].

Tax evasion strategies are evolving, becoming increasingly sophisticated, often involving corporate networks that are difficult to detect with traditional audit techniques. The use of big data analytics and machine learning offers new opportunities to address these challenges, enabling the detection of tax evasion within large and complex data sets [7]. This research proposes an innovative approach that leverages internal and external data to map taxpayer networks and identify potential tax evasion activities. This research contributes to several key Sustainable Development Goals (SDGs), particularly SDG 16: Peace, Justice, and Strong Institutions, by enhancing the transparency and accountability of tax systems through the detection of tax evasion. By leveraging advanced technologies like big data analytics and machine learning, the study also supports SDG 9: Industry, Innovation, and Infrastructure, fostering innovation in public sector operations. Furthermore, by ensuring that all entities contribute their fair share to public revenues, this research promotes SDG 10: Reduced Inequalities and SDG 8: Decent Work and Economic Growth, helping to create a fairer and more inclusive economic environment. Lastly, the research aligns with SDG 17: Partnerships for the Goals, as it highlights the importance of international cooperation in combating tax evasion, particularly in cases involving multinational corporations [8], [9], [10].

In this paper, we propose an approach to detect tax avoidance and evasion within networks of affiliated taxpayers that can disrupt national revenue targets. The approach involves mapping taxpayer networks using internal and external data in the form of multi-graphs, reflecting their financial positions and relationships. The use of big data analytics is essential due to the large volume of data, which requires parallel processing [11]. Tax evasion detection covers various data, including taxpayer financial characteristics and relationships like capital ownership and family ties. Data from the Directorate General of Taxes (DGT) includes 49 million taxpayers, 400 million annual tax invoices, and 295 million periodic tax returns. DGT also incorporates external data such as import and export notifications [12].

1.1. Background of the Problem

Tax evasion encompasses a broad spectrum of criminal activities within the realm of taxation. According to the penal provisions in the General Tax Provisions Law (UU KUP), tax offenses cover all pillars of compliance, including registration, reporting, payment, and accuracy of reporting [13]. The limited resources available to address tax offenses necessitate prioritization and focus in handling such cases. Currently, the Directorate General of Taxes (DGT) employs 562 Civil Servant Investigators (PPNS) [14]. The insufficient number of PPNS requires the DGT to be strategic in selecting and prioritizing tax offenses to be addressed. The handling of large data volumes and the available technology at DGT involves using the Compliance Risk Management (CRM) approach for law enforcement [4]. The CRM for law enforcement generates a risk map of tax offenses from taxpayers across Indonesia to prioritize taxpayers for enforcement actions based on their risk levels, thereby enhancing resource efficiency at DGT. The current law enforcement CRM encompasses only three risk events: TBTS (False Invoices), Unpaid Withholding Tax, and Restitution Attempts [15]. However, the Affiliated-Transaction-Based Tax Evasion (ATTE) cases are not yet included in this CRM. The three risk events in the law enforcement CRM are:

- Risk Event TBTS (False Invoices)

The development of law enforcement CRM includes specific variables that describe this event, used to capture TBTS tax offenses, such as taxpayer profiles, transactions based on tax invoices, reporting behavior of tax returns (SPT), and business activities related to exports and imports [16].

- Risk Event Unpaid Withholding Tax

This event also has specific variables used as predictors for unpaid withholding tax offenses, such as periodic tax return reporting behavior, taxpayer payment behavior, and taxpayer business transactions [17].

- Risk Event Restitution Attempts

There are variables used to predict restitution attempt offenses, such as taxpayer profiles, export-import business transactions, and taxpayer payment behavior. The law enforcement CRM has three risk events (TBTS, Unpaid Withholding Tax, and Restitution Attempts) with equal weights, and the mapping is done for each risk event before combining them into a unified risk map [18]. The law enforcement efforts conducted by the DGT are currently limited to identifying tax evasion based on historical data that has already been entered into the DGT system. These rules only consider individual taxpayer characteristics, whereas most current tax avoidance involves taxpayer networks. Thus, characteristics involving taxpayer networks should also be considered as variables to predict taxpayer compliance risk [19].

1.2. Previous Research

Research on tax evasion has long garnered attention. The utilization of big data analytics to support the handling of tax evasion cases involving taxpayer networks represented in graph form has also attracted considerable interest [20]. Recent studies on employing big data analytics for detecting complex tax evasion schemes have been reviewed and compared to gather information on tax evasion schemes, analytic techniques, and technologies used in detecting these schemes [21]. A comparison of the background issues, tax evasion schemes, and analytical techniques used in previous research is presented in Table 1.

Table 1. Comparison of Backgrounds of Tax Evasion Schemes and Analysis Techniques

No	Title	Author	Background	Tax Avoidance/Tax Evasion Scheme	Method	Data Used	Solution/Tools	Evaluation of Model Performance
1	Big data analytics for Nabbing Fraudulent Transactions in Taxation System	Mehta et al	The tax evasion strategy executed by a group of individuals involves circular trading to reduce indirect tax payments (VAT and GST). Manual inspections are inadequate for detecting fictitious transactions within circular trading due to large data sizes, the complexity of transaction sequences conducted	Businesses collect VAT from sales transactions without issuing tax invoices, then issue fictitious tax invoices to other parties (with values lower than the actual transactions without invoices). These tax invoices are used by other parties as tax credits, thereby reducing the amount of tax payable by those parties. Shell company networks are	Graph/ Social Network Analysis, Map Reduce, Algorithm for identifying fictitious transactions using Benford Analysis.	Formation of a graph from sales transactions, outlier detection.	R Programming with Hadoop	-

No	Title	Author	Background	Tax Avoidance/Tax Evasion Scheme	Method	Data Used	Solution/Tools	Evaluation of Model Performance
			by shell company networks, and transactions lacking identification credentials in this evasion scheme.	created to obscure this evasion scheme with a series of fictitious transactions, making it difficult for tax auditors to detect.				
2	Identifying suspicious groups of affiliated transaction based tax evasion in big data	Ruan et al	Tax evasion through affiliate transactions is a new strategy employed by tax evaders. This tax evasion strategy results in significant losses to state revenue. The tax evasion scheme appears as legitimate transactions within the group of companies. These transactions are diverse, complex, and involve hidden interactions to avoid tax obligations. Current tax inspection techniques are time-consuming with lengthy processes, thus ineffective in detecting tax evasion through affiliate transaction schemes.	1. Sales from entities with low tax rates to affiliates with high tax rates at prices higher than market value. 2. Sales from entities with high tax rates to affiliates with low tax rates at prices lower than market value.	Graph/ Social Network Analysis, Matching suspicious network patterns based on recognized tax evasion patterns, Classification using decision tree classifier algorithms to classify groups engaged in tax evasion and those that are not based on differences in tax rates, corporate network patterns, and abnormal tax burdens compared to industry norms.	Formation of nodes and edge list, formation of a graph.	Spark	10 fold inner cross validation loop
3	Predicting tax avoidance by means of social network analytics	Lismont et al	Tax avoidance not only yields the benefit of reduced tax payments but also carries risks such as penalties or fines imposed by tax authorities.	Graph/Social Network Analysis, Classification using the random forest algorithm.	Formation of node and edge lists, formation of the graph, feature selection, and extraction.	-	Comparison of accuracy rates and ROC curves	

No	Title	Author	Background	Tax Avoidance/Tax Evasion Scheme	Method	Data Used	Solution/Tools	Evaluation of Model Performance
4	TaxVis: a Visual System for Detecting Tax Evasion Group.	Yu et al	Despite being aware of these risks, taxpayers continue to engage in tax avoidance practices. Factors that drive taxpayers to pursue tax avoidance include considerations within social network analysis. Tax revenue losses in China due to tax evasion reach 22%. The current trend in tax evasion involves scenarios that encompass networks of multiple companies. The use of pattern matching techniques has limitations because the process of identifying suspicious network patterns requires the involvement of tax inspectors. The rapid evolution of tax evasion patterns makes it increasingly difficult for tax inspectors to keep up with them. Therefore, due to these limitations, detecting tax evasion through pattern matching requires a new, more effective approach.	Tax avoidance through taxpayer networks/groups.	Social network analysis, network embedding using the node2vec method, classification comparison among LightGBM, random forest, SVM, & AdaBoost. Identifying using rules regarding common anomalies such as transaction volume anomalies, mismatched product categories traded within the taxpayer's business scope, transactions between parent and subsidiary companies.	Feature extraction, Formation of the graph-network of company group transactions.	TaxVis Platform (an application developed by researchers, as an output of the study).	Kurva ROC

Based on the comparison in Table 1, it is evident that the four pieces of literature share similarities in highlighting the increasing complexity of tax evasion schemes. Currently, tax evasion schemes involve networks of taxpayers and transactions among affiliated parties, making it difficult to identify tax avoidance and evasion activities [22]. The literature also reveals that tax authorities face resource limitations and technical constraints in detecting tax evasion schemes involving complex corporate networks.

The tax evasion schemes in the literature exhibit differences/variations. There are at least two variations of tax evasion schemes identified in these four pieces of literature, as follows:

- Tax evasion in indirect taxes such as Value Added Tax (VAT) and Goods and Services Tax (GST) involves the issuance and use of invoices not based on actual transactions (fictitious). This scheme includes the use of dummy companies to complicate detection processes by tax authorities [23].
- Tax Evasion in Direct Taxes Tax evasion in direct taxes, such as income tax, involves profit shifting from taxpayers with higher tax rates to affiliated taxpayers with lower tax rates. These affiliated taxpayers may be in the same country or across borders. Different tax rates in the same country can arise from progressive tax rates, compensable losses, or tax facilities. Furthermore, differences in tax rates can also arise due to variances between countries. Multinational corporations operate business units across several different countries, and transfer pricing is a method commonly utilized by large companies to measure the performance of each business line [24].

Most of the previous research on tax evasion detection has considered aspects of taxpayer networks. It was felt that the proposals in previous research were still not comprehensive because they only involved some parties, even though the tax evasion scheme involved many quite complicated relationships, so this research aims to define the architecture and process to produce a comprehensive machine learning model design for detecting tax evasion [25].

1.3. Tax Evasion Schemes

Advancements in information and communication technology provide opportunities for organizations to enhance competitiveness and add value to their strategies. These improvements influence how organizations interact with their business environment [26]. The Directorate General of Taxes (DGT) uses these advancements to not only offer services but also to oversee, audit, and enforce law, particularly against non-compliant taxpayers suspected of tax evasion. Document studies have identified four key tax evasion schemes that are the focus of DGT's examination and enforcement activities [27].

- Issuance and Use of False Tax Invoices (TBTS)

This tax evasion scheme involves issuing tax invoices without actual transactions. This research focuses on TBTS invoices in circular trading scenarios, where a taxable entrepreneur (PKP) delivers goods or services to end consumers, collects VAT, but does not issue tax invoices [28]. The invoices are useless to end consumers as they do not file VAT returns. The PKP Seller then creates tax invoices to sell to other PKPs, forming a series of fictitious transactions in a circular network, where no VAT is paid, complicating detection by tax auditors.

- Unpaid Withholding Tax

The criminal offense of unpaid withholding tax is described in Article 39, paragraph 1, letter i of the UU KUP, which states: *"Anyone who deliberately fails to remit taxes that have been withheld or collected, thereby causing losses to state revenue, is punishable by imprisonment for at least six months and at most six years and a fine of at least two times the amount of unpaid tax and at most four times the amount of unpaid tax."*

- Restitution Attempts

The criminal offense of restitution attempts is based on Article 39, paragraph 3 of the UU KUP, which states: *"Anyone who attempts to commit the criminal acts of misusing or unlawfully using a Tax Identification Number (NPWP) or Taxable Entrepreneur Registration (PKP) as referred to in paragraph (1) letter b, or submitting a tax return and/or information that is not true or incomplete, as referred to in compensation or tax credits, is punishable by imprisonment for at least six months and at most two*

years and a fine of at least two times the amount of restitution requested and/or compensation or credits claimed and at most four times the amount of restitution requested and/or compensation or credits claimed.”

- **Affiliated-Transaction-Based Tax Evasion (ATTE)**

According to Article 33, paragraph (1) of Government Regulation No. 55 of 2022, affiliated transactions involve relationships due to ownership, control, or family ties. In tax evasion, transfer pricing refers to setting unreasonable prices within affiliated companies to reduce corporate income tax [29]. The goal is to shift profits to other affiliated companies to offset losses or gain tax incentives, minimizing the group’s overall tax. Transfer pricing can occur domestically or internationally, both aimed at reducing the group’s tax liability.

The present study seeks to find out the moderating effect of a taxpayer’s network on his/her tax avoidance by establishing the correlation between the centrality of the network and tax avoidance measures [30]. More precisely, the aim of the study focuses on establishing whether taxpayers’ level of connectivity or degree centrality within the ownership structure affects their ETR, often used as a measure of tax evasion.

Based on the information on the corporate owners’ identities, the research builds an ownership network where nodes are taxpayers, and links between them are ownership connections. Finally, degree centrality which determines the number of direct contacts a taxpayer holds in this network for every taxpayer is computed [31].

Therefore, we propose that a taxpayer’s network, based on their degree centrality in an ownership network, influences their ETR. Through degree centrality, the taxpayers whose graph G has more ownership connections are considered to have more opportunities and resources in deploying the strategies in avoiding taxes [32]. This may lead to a lower ETR in comparison with the other taxpayers with a lower degree centrality, though the latter may have fewer connections and fewer opportunities to practice evasion.

2. RESEARCH METHOD

This research is conducted in a qualitative way, because qualitative methods are good for studying a phenomenon in more depth. It involves a literature review and information gathering through qualitative data collection techniques [33]. The model’s performance was evaluated using several metrics, including accuracy, precision, recall, and F1-score, with confidence intervals calculated to assess the reliability of these metrics. It involves a literature review and information gathering through qualitative data collection techniques. Information is collected through document studies by reviewing the relevant tax laws to understand the definition of affiliated relationships [34]. The use of method in each research stages are as below:

- Problem Formulation (Qualitative): Document Study
- Theoretical Framework Development (Qualitative): Systematic Literature Review, Literature Study, Interview
- Design of Machine Learning Model (Qualitative): Systematic Literature Review, Interview

2.1. Hypothesis

2.1.1. Null Hypothesis (H_0)

One cannot distinguish any differences in effective tax rate between taxpayers with low and high levels of degree centrality in the ownership network.

$$\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k \quad (1)$$

Where:

- μ_i represents the mean Effective Tax Rate (ETR) for the i -th class of taxpayers, categorized based on their degree centrality in the ownership network.
- k is the total number of taxpayer classes.

2.1.2. Alternative Hypothesis (H_1)

For the taxes paid to the ownership network, it was revealed that taxpayers with higher degree centrality have a significantly lower ETR than those with lower degree centrality. This hypothesis will be tested using Analysis of Variance (ANOVA) to establish if variations of ETR by the classes of taxpayers based on their degree centrality are statistically significant [35].

$$\exists i, j \quad \text{such that} \quad \mu_i \neq \mu_j \quad (2)$$

Where:

- At least one pair (i, j) exists such that the mean ETR for class i is significantly different from the mean ETR for class j .

2.2. Document Study

Document study was carried out by examining the Income Tax Law, to find out the definition of an Affiliated-Transaction-Based Tax Evasion. The study uses Analysis of Variance (ANOVA) to analyze the differences in the effective tax rates between distinct classes of taxpayers according to their degree centrality [36]. Therefore the hypothesis of the study is: The centrality of taxpayers in the network of taxpayers, which means, how ‘connected’ they are to others in the network, affects their ETR and the corresponding tax avoidance behaviors in a way that those with higher centrality behave differently from those of lower centrality [37].

2.3. Interview

Interview is used in several research steps, including problems definition, understanding the nature of tax evasion. Interview was conducted with involved peoples with deferent roles [38]. It aims to obtain information regarding the tax evasion schemes currently facing by DGT, as well as the infrastructure that supports current data analysis. The target respondents of the interviews are shown as the Table 2.

Table 2. Hypothesis Test Results

No	Respondent	Total	Information collected
1	Tax Investigator	1 person	Scheme for issuing and using tax invoices that are not based on actual transactions
2	Data Analyst	1 person	Data sources, data quality, available tools
3	Data Warehouse Developer	1 person	Current data warehouse and Hadoop architecture

The table 2 summarizes interviews conducted with key individuals to gather insights on tax evasion schemes and DGT’s data infrastructure. A Tax Investigator provided details on schemes involving the issuance of tax invoices not based on actual transactions. A Data Analyst contributed information on data sources, quality, and available analysis tools. Finally, a Data Warehouse Developer explained the current data warehouse setup and Hadoop architecture, helping evaluate the infrastructure’s capacity for detecting complex tax evasion schemes. These interviews provided essential perspectives from investigative, analytical, and technical angles.

2.4. Literature Study

Articles were selected to gain insights into research objectives, methods, and other aspects discussed in articles on tax evasion, especially those related to computer sciences. Ten articles published between 2018 and 2023 were selected [39]. These articles were compared in terms of background, tax avoidance/evasion schemes, methods used, and data utilized. Continuously, a technical design is made and finally a conclusion is drawn [40].

2.5. Observation

Observations were carried out by accessing the DJP Integrated Data Warehouse, to find out available tax data, size and growth of tax data.

As preparation for the ANOVA test that will examine the relationship between the independent variable (taxpayer networks) and the dependent variable (tax avoidance behavior), some data manipulation is required. This includes normalizing the outdegree (degree centrality) and the Effective Tax Rate (ETR) before classifying the taxpayers [41]. The steps are as follows:

- **Data Collection:** Fiscal year 2022 tax return data, including information on commercial income, tax due, and ownership relations between taxpayers, is used. The dataset contains 503,287 records and 7 columns. Sensitive taxpayer identification information is excluded [14].
- **Outlier Detection and Removal:** The ETR values range from -1.27 (indicating tax refunds or negative tax rates) to a maximum of 1555.58%, which indicates the presence of extreme outliers. Outliers are detected and removed to improve the analysis.
- **Standardization:** It is important to ensure that all variables are on the same scale. Normalization improves the efficiency and readability of the ANOVA, increasing the chances of rejecting a false null hypothesis.
- **Outdegree (Degree Centrality) Standardization:**
 1. Calculate the mean (μ) and standard deviation (σ) of the outdegree values.
 2. Standardize the outdegree for each taxpayer using the formula:

$$\text{Standardized Outdegree} = \frac{\text{Outdegree} - \mu_{\text{Outdegree}}}{\sigma_{\text{Outdegree}}} \quad (3)$$

This ensures the outdegree values have a mean of 0 and a standard deviation of 1, making them comparable across taxpayers.

- **ETR (Effective Tax Rate) Standardization:**
 1. Calculate the mean (μ) and standard deviation (σ) of the ETR values.
 2. Standardize the ETR for each taxpayer using the formula:

$$\text{Standardized ETR} = \frac{\text{ETR} - \mu_{\text{ETR}}}{\sigma_{\text{ETR}}} \quad (4)$$

Standardizing the ETR allows for comparison across taxpayers on a uniform scale, simplifying the detection of significant differences during the ANOVA analysis.

2.5.1. Binning Process

Binning is a technique that classifies continuous variables into respective bins or classes. In this case, the standardized outdegree and ETR values are used to partition the taxpayer population into different groups for performing ANOVA.

2.5.2. Analysis of Variance (ANOVA)

ANOVA is a statistical technique used to compare the means of three or more groups to determine if there are significant differences between them. In this study, ANOVA will be applied to answer the research question, comparing the central tendencies of the ETR across taxpayer classes categorized by degree centrality. The steps are as follows:

To perform the ANOVA test, the first step is to formulate the hypothesis. The null hypothesis (H_0) states that all taxpayer classes have the same mean Effective Tax Rate (ETR), implying no effect of degree centrality on tax avoidance behavior. In contrast, the alternative hypothesis (H_1) suggests that at least one taxpayer class has a different mean ETR, indicating that degree centrality influences tax avoidance behavior [42]. Next, the mean ETR is computed for each taxpayer class based on their degree centrality, followed by calculating the overall mean ETR across all taxpayer classes. The sum of squares, specifically the between-group sum of squares (SSB), is then calculated to measure the variation in ETR between the taxpayer classes, which involves summing the squared differences between each class mean and the overall mean, weighted by the number of observations in each class [43]. Afterward, the F-statistic is determined by dividing the mean square between groups (MSB) by the mean square within groups (MSW). The calculated F-statistic is compared to a critical value from the F-distribution table based on the chosen significance level (e.g., $\alpha = 0.05$) and the degrees of freedom for the numerator ($k - 1$) and the denominator ($N - k$). If the F-statistic exceeds the critical value, the null hypothesis is rejected, indicating a statistically significant difference in mean ETRs between at least two taxpayer classes. If the null hypothesis is rejected, it suggests that degree centrality has a significant effect on tax avoidance behavior, as reflected in the differences in ETR among taxpayer classes.

3. FINDINGS

A literature study was conducted to explore current methods for detecting tax evasion, focusing on affiliated transactions, identity matching algorithms, and transfer pricing. Affiliated transactions, as defined by the Income Tax Act, involve relationships between parties that can impact transaction values. Identity matching algorithms, such as Levenshtein and Jaro-Winkler, measure name similarity based on required changes, while the Jaccard coefficient calculates dataset similarity by comparing intersections and unions. Transfer pricing, a form of tax evasion, shifts profits to lower-tax jurisdictions, with Social Network Analysis (SNA) commonly used to analyze taxpayer networks. Models are evaluated using techniques like 10-fold cross-validation and ROC curve comparison. The challenge of missing identity identifiers in corporate networks but do not address identity resolution, a key factor in graph analysis, where unique node identity is essential to avoid obscuring results [44]. This study will focus on identity matching and resolution to build accurate taxpayer network graphs, with a discussion on the utilization of big data analytics at the Directorate General of Taxes (DGT) to detect tax evasion in complex networks. The dataset after outlier removal process comprises 418,484 taxpayers, with key variables being outdegree and Effective Tax Rate (ETR). The outdegree shows that most taxpayers have a single ownership connection, with an average of 1.05 and a standard deviation of 0.76. The range extends from 1 to 166, but the majority are concentrated at 1, indicating minimal network complexity [45]. The ETR has an average of 3.28%, with a standard deviation of 0.0523, reflecting low variability. Most taxpayers have an ETR of 0%, with 75% paying 11% or less, and a maximum rate of 16%. Overall, the data suggests that most taxpayers have simple ownership structures and very low tax rates, with a few exceptions. The descriptive statistics of the dataset are as below:

Table 3. Descriptive Statistics of Taxpayer Dataset After Outlier Removal

	Outdegree	ETR
Count	418484.000000	418484.000000
Mean	1.050336	0.032790
Std	0.757554	0.052343
Min	1.000000	-1.270000
25%	1.000000	0.000000
50%	1.000000	0.000000
75%	1.000000	0.110000
Max	166.000000	0.160000

After standardization, both the outdegree and Effective Tax Rate (ETR) have means close to zero and standard deviations of one, making them directly comparable. The outdegree now has a mean of nearly zero, with most taxpayers clustered around this value. The standardized ETR also has a mean close to zero, but it shows more variability, with values ranging from 24.89 to 2.43.

The 25th and 50th percentiles of the ETR are at 0.626, indicating that a significant portion of taxpayers have below-average ETRs, while the 75th percentile is at 1.48, showing that the top 25% have notably higher tax rates. Overall, the standardization process has effectively centered both variables, but the ETR still exhibits considerable variation across taxpayers [46]. The detail descriptive statistics of the standardize data are as below:

Table 4. Descriptive Statistics of Standardized Outdegree and ETR

	Outdegree	ETR
Count	4.184840e+05	4.184840e+05
Mean	1.216289e-16	1.967226e-16
Std	1.000001e+00	1.000001e+00
Min	-6.644607e-02	-2.488954e+01
25%	-6.644607e-02	-6.264459e-01
50%	-6.644607e-02	-6.264459e-01
75%	-6.644607e-02	1.475082e+00
Max	2.177399e+02	2.430322e+00

Taxpayers are categorized into three classes based on their standardized outdegree values. The classi-

fication is as follows:

- Class 3: Taxpayers with a standardized outdegree greater than or equal to 4 are placed in this class. These taxpayers have the highest degree centrality within the ownership network, indicating a more complex network structure.
- Class 2: This class includes taxpayers whose standardized outdegree falls between 0 and 4. These individuals have moderate network centrality, suggesting a more typical or average level of ownership connections.
- Class 1: Taxpayers with a standardized outdegree less than 0 are categorized into this class. They have the lowest network centrality, indicating simpler ownership structures with fewer connections.

This categorization helps to analyze and compare the behavior of taxpayers across different levels of network complexity.

Table 5. Classification of Taxpayers by Standardized Outdegree

Outdegree Class	Number of Taxpayer
1	412437
2	3729
3	2318

Based on the ANOVA analysis using the standardized data, the results indicate a significant difference in Effective Tax Rate (ETR) across the three taxpayer classes categorized by their standardized outdegree. The F-statistic obtained from the ANOVA is 75.06, which measures the ratio of the variance between the classes to the variance within the classes. A higher F-statistic suggests a greater difference between the groups compared to within-group variation [47]. The p-value associated with this F-statistic is extremely small, at approximately 6.34×10^{-33} . This p-value is far below the conventional significance level (e.g., 0.05), indicating that the differences in ETR across the different taxpayer classes are statistically significant. Therefore, we can confidently reject the null hypothesis, which would have suggested that there is no difference in ETR between the classes based on their standardized outdegree. In summary, the ANOVA results strongly suggest that the level of network centrality, as measured by standardized outdegree, has a significant impact on a taxpayer's Effective Tax Rate.

3.1. Tax Evasion Schemes Faced by DGT

Based on interviews and document studies conducted, there are two tax evasion schemes that are the focus of DGT's examination and law enforcement activities:

- Issuance and Use of Tax Invoices Not Based on Actual Transactions (TBTS) In this scheme, the taxable entrepreneur (PKP) sells goods or services to the end consumer and collects VAT but does not issue a tax invoice. The tax invoice is not valuable to the end consumer as they do not report the VAT Return (SPT) and do not credit the input tax listed in the tax invoice. Subsequently, the PKP issues a tax invoice for the sale to sell to another PKP. A series of fake transactions are created to complicate detection by tax examiners, forming a circular network, resulting in no VAT being paid from these false transactions. Figure 1 provides a simple illustration of circular trading; however, in reality, the false transaction network used to conceal the sale of TBTS invoices is more complex.

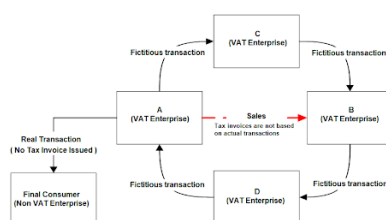


Figure 1. Illustration of Circular Trading

- Related Party Transactions or Transfer Pricing

Transfer pricing in the context of tax evasion refers to the policy of setting unreasonable prices within an affiliated group of companies to reduce corporate income tax payments. The focus of transfer pricing schemes is on shifting taxable profits to affiliated companies, thereby reducing overall tax payments due to differences in tax rates, loss compensation at the receiving company, or tax facilities resulting in lower taxed transferred profits. Transfer pricing within a group can be divided into domestic transfer pricing and international transfer pricing. These schemes are used to shift profits to affiliated companies, reducing the total tax paid by the group.

The definition of related parties is stipulated in the Income Tax Law. According to Article 18 paragraph (4) of the Income Tax Law, a special relationship is considered to exist if:

1. A taxpayer directly or indirectly owns at least 25% of another taxpayer.

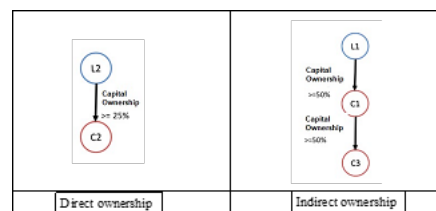


Figure 2. Ownership of 25%

Figure 2 depicts a simple network where one entity (L2) owns 25% or more of another entity (C2). This ownership structure indicates that L2 exerts significant control over C2, making it a part of an affiliated network for tax purposes.

2. The relationship between taxpayers involves at least 25% ownership in two or more other taxpayers

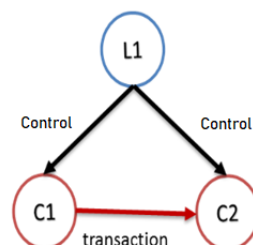


Figure 3. Ownership of 25% in more than one taxpayer

Figure 3 represents a more intricate network where one entity, L1, holds 25% or more ownership in two separate entities, C1 and C2. This scenario is more complex because L1 not only controls C1 but also exerts significant influence over C2. This dual ownership allows L1 to coordinate transactions between C1 and C2 to further its objectives, including tax avoidance or evasion.

In the context of the study, this figure highlights how the control exerted by a single parent entity over multiple subsidiaries increases the complexity of the network. Tax authorities need advanced tools like artificial intelligence and social network analysis to detect these sophisticated evasion strategies, as traditional audit techniques may fail to uncover hidden relationships and manipulated transactions [48]. This figure also underscores the importance of analyzing ownership percentages in tax audits, as control over multiple entities can significantly impact tax compliance.

3. Several taxpayers are under the same control.

Controlling several taxpayers This figure represents a more complex network where L1 controls both C1 and C2, with ownership percentages greater than or equal to 50% between them. Additionally, transactions occur between C1 and C2, indicating possible transfer pricing or tax evasion practices within the network, as profits may be shifted between these entities to lower tax liabilities.

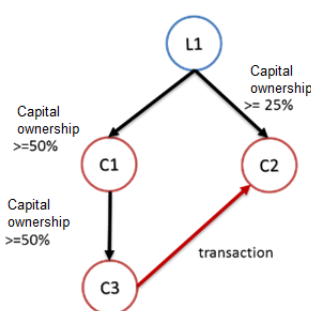


Figure 4. Controlling several taxpayers

4. Family Relationship.

Family relationships by blood or marriage in a straight line or one degree sideways. Special relationships can also occur from combinations of these relationships. Companies engaging in tax evasion schemes will attempt to conceal special relationships within a complex network. For example, special relationships involving combined connections are as follows:

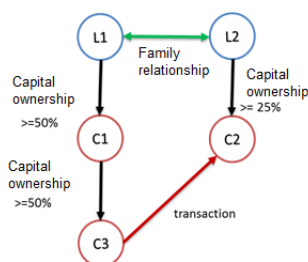


Figure 5. Family relationship

C2 and C3 have a special relationship because both companies are owned by individuals who have familial ties and hold more than 25% of the shares in each company. The fairness of the pricing of transactions between C2 and C3 may be influenced by this special relationship.

3.2. Proposed Analytical Process for Detecting Tax Evasion

The steps required in analyzing tax evasion begin with constructing a graph depicting various relationships between taxpayers. Nodes in this graph represent taxpayers and other related parties. Edges represent numerous relationships, such as ownership, debt relationships, shared directors and commissioners, family relationships between individuals, transactions of goods or services between taxpayers, and import/export transactions by taxpayers. The steps include:

3.2.1. Graph Construction

Graph construction involves identity matching, resolution, and record linkage as part of Master Data Management activities. This is necessary due to data quality issues forming the relationships between taxpayers.

- Identity Matching and Resolution

Transactions without identity identifiers pose a challenge in graph construction. Transaction data involving relationships between taxpayers and other parties, both in annual tax return and external data, often lack validation of identity identifiers like NPWP and NIK for Indonesian taxpayers. Additionally, transactions involving foreign taxpayers typically only list names, addresses, and countries, which are prone to errors and inconsistencies. This negatively impacts the graph's accuracy.

The proposed solution is to implement identity matching and resolution processes as part of Master Data Management. Identity matching cannot be accomplished solely with SQL; Taxpayer data that cannot

be recognized based on the main identifying attribute, will be matched based on the similarity of values from several other identifying attributes. The comparison involves the weighting of each attribute used as well as several algorithms that are often used in identity matching. Taxpayer identity matching compares the similarity of 7 attributes, and for these attributes it is necessary to give a weight that determines how much similarity in these attributes determines the data match with the taxpayer reference [49]. The attributes used in identity matching include the location and characteristics of the person or organization. The possibility of finding a number of people with similar characteristics such as name or place of birth in a wider location (e.g. city) is greater than in a narrower location (e.g. neighborhood). Therefore, the weight for similarity of the broader location (city) of the 2 entities being compared is given a smaller value than the narrower location. The resulting graph, built from nodes and edges processed through identity matching and resolution, shows high accuracy.

- Transformation into Graph Data Format

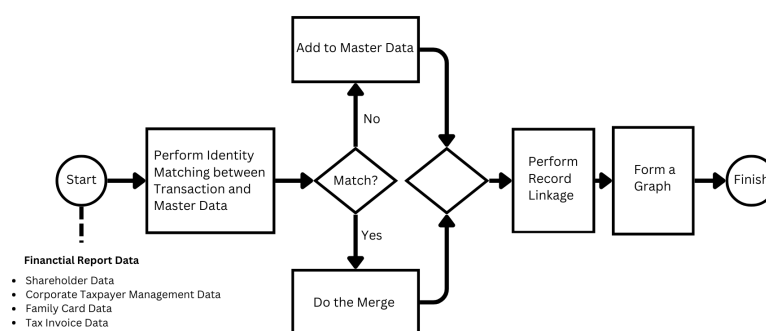


Figure 6. Taxpayer Graph Construction

Figure 6 illustrates the process of constructing a taxpayer network graph through identity matching and data merging. It starts by collecting financial data, including shareholder, taxpayer management, family card, and tax invoice data. The system performs identity matching between transaction and master data. If a match is found, the records are merged; if not, unmatched entities are added to the master data [50]. Afterward, record linkage is performed to connect data across various sources, forming a graph where nodes represent taxpayers and edges represent their relationships, such as ownership or transactions. This graph enables tax authorities to analyze complex networks for detecting tax evasion schemes.

3.2.2. Analysis for Detecting Tax Evasion

Tax evasion detection analysis is performed using the constructed graph data. Based on previous research, tax evasion detection can be approached in two ways:

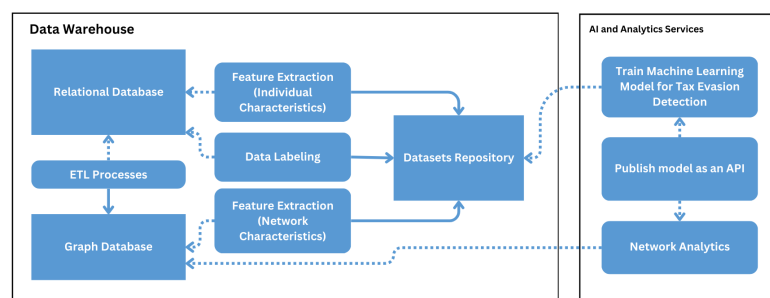


Figure 7. Process Design and Architecture

- Network Pattern Matching

Network pattern matching is based on predefined tax evasion network patterns. If the graph data is stored in a graph database like Neo4J, searching for taxpayer networks based on a pattern can be done by writing queries called Cypher.

- Classification

Detection of tax evasion using this approach is accomplished by training a predictive model built with features that encompass both individual taxpayer characteristics, network metrics, and several indicators indicating conformity with characteristics of taxpayers involved in tax evasion.

4. MANAGERIAL IMPLICATIONS

This study provides significant managerial insights for tax authorities, particularly the Directorate General of Taxes (DGT), and other tax enforcement stakeholders. It underscores the importance of integrating AI and big data analytics into tax systems to detect tax evasion within complex corporate networks. Tax administrators should invest in AI-driven tools to map taxpayer networks, identify suspicious activities, and automate fraud detection. Leveraging Social Network Analysis (SNA) and machine learning can improve decision-making and compliance, reducing revenue loss.

The study also stresses the need for collaboration between DGT and financial institutions through data-sharing agreements for more effective network analysis. Enhancing information exchange and aligning with Sustainable Development Goals (SDGs) can promote transparency, accountability, and fairness in tax collection, contributing to economic equity and development.

5. CONCLUSION

Tax evasion schemes are becoming increasingly complex, involving corporate networks where perpetrators create intricate webs of companies and transactions to conceal their activities. The large volume of transactions and issues related to data quality have rendered current tax inspection techniques ineffective in detecting such evasion. This research addresses these challenges by proposing the use of big data analytics and artificial intelligence to build capabilities for detecting tax evasion in complex corporate networks.

The proposed approach involves the formation of graphs and the use of Social Network Analysis (SNA) for tax evasion detection, including schemes like issuing invoices without actual transactions and transfer pricing through special relationships. Graph formation incorporates identity matching and resolution, supported by the selection of identity matching algorithms. Our experiments demonstrate that combining the Levenshtein and Jaccard algorithms yields the best performance in terms of accuracy, precision, and recall. This method allows for more accurate mapping of taxpayer networks and better detection of evasion activities.

The findings suggest that network centrality plays a crucial role in tax evasion. Specifically, taxpayers with higher degree centrality, such as those in Taxpayer Class 3 (mean ETR of 1.5%), are more likely to engage in tax avoidance compared to Class 1 (mean ETR of 2.5%) and Class 2 (mean ETR of 4.5%). The ANOVA analysis, with an F-statistic of 75.06 and a p-value of 6.34×10^{-33} , confirms the statistical significance of these differences, highlighting the influence of network complexity on tax avoidance behavior.

Moreover, the proposed integration of artificial intelligence into the Directorate General of Taxes (DGT)'s existing systems can greatly improve the effectiveness of tax enforcement, particularly in detecting sophisticated evasion strategies that are difficult to identify through traditional audit methods. This research also aligns with several Sustainable Development Goals (SDGs) by promoting transparency, accountability, and equity in tax systems. Future research should explore the scalability of this model across different jurisdictions and its application to a broader range of tax evasion schemes. Additionally, enhancing data quality and expanding the dataset to include more diverse taxpayer profiles could further improve the model's accuracy and reliability in detecting tax evasion.

6. DECLARATIONS

6.1. About Authors

Nuryani (NR)  <https://orcid.org/0009-0001-2079-5897>

Achmad Benny Mutiara (AB)  <https://orcid.org/0000-0003-0220-8289>

I Made Wiryana (IM)  -

Detty Purnamasari (DP)  <https://orcid.org/0000-0003-2806-720X>

Souza Nurafrianto Windiarsono Putra (SN)  -

6.2. Author Contributions

Conceptualization: NR, AB, IM, DP, and SN; Methodology: NR, AB, and DP; Software: IM; Validation: NR, AB, and SN; Formal Analysis: AB and DP; Investigation: NR and IM; Resources: AB; Data Curation: NR and DP; Writing Original Draft Preparation: NR, AB, and DP; Writing Review and Editing: NR, IM, and SN; Visualization: NR; All authors, NR, AB, IM, DP, and SN, have read and agreed to the published version of the manuscript.

6.3. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.4. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.5. Declaration of Conflicting Interest

The authors declare that they have no conflicts of interest, known competing financial interests, or personal relationships that could have influenced the work reported in this paper.

REFERENCES

- [1] J. Ruan, Z. Yan, B. Dong, Q. Zheng, and B. Qian, "Identifying suspicious groups of affiliated-transaction-based tax evasion in big data," *Information Sciences*, vol. 477, pp. 508–532, 2019.
- [2] R. A. Rahman, S. Masrom, and N. Omar, "Tax avoidance detection based on machine learning of malaysian government-linked companies," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 535–541, 2019.
- [3] H. Yu, H. He, Q. Zheng, and B. Dong, "Taxvis: a visual system for detecting tax evasion group," in *The World Wide Web Conference*, 2019, pp. 3610–3614.
- [4] P. Mehta, J. Mathews, S. Kumar, K. Suryamukhi, C. Sobhan Babu, and S. Kasi Visweswara Rao, "Big data analytics for nabbing fraudulent transactions in taxation system," in *Big Data–BigData 2019: 8th International Congress, Held as Part of the Services Conference Federation, SCF 2019, San Diego, CA, USA, June 25–30, 2019, Proceedings 8*. Springer, 2019, pp. 95–109.
- [5] B. Shi, B. Dong, Y. Xu, J. Wang, Y. Wang, and Q. Zheng, "An edge feature aware heterogeneous graph neural network model to support tax evasion detection," *Expert Systems with Applications*, vol. 213, p. 118903, 2023.
- [6] M.-J. Segovia-Vargas *et al.*, "Detection of shell companies in financial institutions using dynamic social network," *Expert Systems with Applications*, vol. 207, p. 117981, 2022.
- [7] J. Garcia-Bernardo, J. Witteman, and M. Vlaanderen, "Uncovering the size of the illegal corporate service provider industry in the netherlands: a network approach," *EPJ Data Science*, vol. 11, no. 1, p. 23, 2022.
- [8] Division for Sustainable Development Goals, Department of Economic and Social Affairs, United Nations, "Sustainable development goals un," 2024. [Online]. Available: <https://sdgs.un.org/goals>
- [9] The World Bank, "Sustainable development goals sdg," 2024. [Online]. Available: <https://www.worldbank.org/en/programs/sdgs>
- [10] —, "Our world in data: The united nations sustainable development goals (sdgs)," 2020. [Online]. Available: <https://ourworldindata.org/sdgs>

- [11] M. R. Anwar and L. D. Sakti, "Integrating artificial intelligence and environmental science for sustainable urban planning," *IAIC Transactions on Sustainable Digital Innovation (ITSIDI)*, vol. 5, no. 2, pp. 179–191, 2024.
- [12] C. Stamile, A. Marzullo, and E. Deusebio, *Graph Machine Learning: Take graph data to the next level by applying machine learning techniques and algorithms*. Packt Publishing Ltd, 2021.
- [13] E. Scifo, *Graph Data Science with Neo4j: Learn how to use Neo4j 5 with Graph Data Science library 2.0 and its Python driver for your project*. Packt Publishing Ltd, 2023.
- [14] N. Anwar, A. M. Widodo, B. A. Sekti, M. B. Ulum, M. Rahaman, and H. D. Ariessanti, "Comparative analysis of nij and nist methods for microsd investigations: A technopreneur approach," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 2, pp. 169–181, 2024.
- [15] G. S. Putra, I. I. Maulana, A. D. Chayo, M. I. Haekal, R. Syaharani *et al.*, "Pengukuran efektivitas platform e-learning dalam pembelajaran teknik informatika di era digital," *Jurnal MENTARI: Manajemen, Pendidikan dan Teknologi Informasi*, vol. 3, no. 1, pp. 19–29, 2024.
- [16] J. Lismont, E. Cardinaels, L. Bruynseels, S. De Groote, B. Baesens, W. Lemahieu, and J. Vanthienen, "Predicting tax avoidance by means of social network analytics," *Decision Support Systems*, vol. 108, pp. 13–24, 2018.
- [17] J. Jones, E. Harris, Y. Febriansah, A. Adiwijaya, and I. N. Hikam, "Ai for sustainable development: Applications in natural resource management, agriculture, and waste management," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 143–149, 2024.
- [18] M. Pereira, I. Guvlor *et al.*, "Implementation of artificial intelligence framework to enhance human resources competency in indonesia," *International Journal of Cyber and IT Service Management*, vol. 4, no. 1, pp. 64–70, 2024.
- [19] H. A. Winata and F. Simon, "Influence of profitability, audit quality, and corporate governance on earnings management," *APTISI Transactions on Management*, vol. 8, no. 2, pp. 93–104, 2024.
- [20] S. Lestari, S. Watini, and D. E. Rose, "Impact of self-efficacy and work discipline on employee performance in sociopreneur initiatives," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 2, pp. 270–284, 2024.
- [21] A. Ruangkanjanases, A. Khan, O. Sivarak, U. Rahardja, and S.-C. Chen, "Modeling the consumers' flow experience in e-commerce: The integration of ecm and tam with the antecedents of flow experience," *SAGE Open*, vol. 14, no. 2, p. 21582440241258595, 2024.
- [22] D. Nugroho and P. Angela, "The impact of social media analytics on sme strategic decision making," *IAIC Transactions on Sustainable Digital Innovation (ITSIDI)*, vol. 5, no. 2, pp. 169–178, 2024.
- [23] Y. S. Dewi, "Influence of type and dose of coagulants on vehicle wash wastewater," *ADI Journal on Recent Innovation*, vol. 6, no. 1, pp. 8–16, 2024.
- [24] U. Rahardja, N. Lutfiani *et al.*, "The strategy of improving project management using indicator measurement factor analysis (imf) method," in *Journal of Physics: Conference Series*, vol. 1477, no. 3. IOP Publishing, 2020, p. 032023.
- [25] L. W. Ming, J. Anderson, F. Hidayat, F. D. Yulian, and N. Septiani, "Ai as a driver of efficiency in waste management and resource recovery," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 128–134, 2024.
- [26] D. Bennet, S. A. Anjani, O. P. Daeli, D. Martono, and C. S. Bangun, "Predictive analysis of startup ecosystems: Integration of technology acceptance models with random forest techniques," *CORISINTA*, vol. 1, no. 1, pp. 70–79, 2024.
- [27] U. Rahardja, I. D. Hapsari, P. H. Putra, and A. N. Hidayanto, "Technological readiness and its impact on mobile payment usage: A case study of go-pay," *Cogent Engineering*, vol. 10, no. 1, p. 2171566, 2023.
- [28] D. S. Wuisan and T. Handra, "Maximizing online marketing strategy with digital advertising," *Startupreneur Business Digital (SABDA Journal)*, vol. 2, no. 1, pp. 22–30, 2023.
- [29] P. A. Sunarya, R. Refianti, A. B. Mutiara, and W. Octaviani, "Comparison of accuracy between convolutional neural networks and naïve bayes classifiers in sentiment analysis on twitter," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 5, 2019.
- [30] S. A. Hasan, W. N. Al-Zahra, A. S. Auralia, D. A. Maharani, R. Hidayatullah *et al.*, "Implementasi teknologi blockchain dalam pengamanan sistem keuangan pada perguruan tinggi," *Jurnal MENTARI: Manajemen, Pendidikan dan Teknologi Informasi*, vol. 3, no. 1, pp. 11–18, 2024.
- [31] E. Ligia, K. Iskandar, I. K. Surajaya, M. Bayasut, O. Jayanagara, and K. Mizuno, "Cultural clash: Investi-

- gating how entrepreneurial characteristics and culture diffusion affect international interns' competency," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 2, pp. 182–198, 2024.
- [32] P. Sithole, E. Zirolla, and S. Lowel, "Artificial intelligence in literacy libraries a review of the literature," *International Journal of Cyber and IT Service Management*, vol. 4, no. 1, pp. 58–63, 2024.
- [33] V. Melinda, T. Williams, J. Anderson, J. G. Davies, and C. Davis, "Enhancing waste-to-energy conversion efficiency and sustainability through advanced artificial intelligence integration," *International Transactions on Education Technology (ITEE)*, vol. 2, no. 2, pp. 183–192, 2024.
- [34] D. Hernandez, L. Pasha, D. A. Yusuf, R. Nurfaizi, and D. Julianingsih, "The role of artificial intelligence in sustainable agriculture and waste management: Towards a green future," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 150–157, 2024.
- [35] C. Lukita, L. D. Bakti, U. Rusilowati, A. Sutarman, and U. Rahardja, "Predictive and analytics using data mining and machine learning for customer churn prediction," *Journal of Applied Data Sciences*, vol. 4, no. 4, pp. 454–465, 2023.
- [36] M. F. Nur and A. Siregar, "Exploring the use of cluster analysis in market segmentation for targeted advertising," *IAIC Transactions on Sustainable Digital Innovation (ITSIDI)*, vol. 5, no. 2, pp. 158–168, 2024.
- [37] S. Audiah, Y. P. A. Sanjaya, O. P. Daeli, M. Johnson *et al.*, "Transforming energy and resource management with ai: From theory to sustainable practice," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 158–163, 2024.
- [38] E. Sana, A. Fitriani, D. Soetarno, M. Yusuf *et al.*, "Analysis of user perceptions on interactive learning platforms based on artificial intelligence," *CORISINTA*, vol. 1, no. 1, pp. 26–32, 2024.
- [39] T. Hidayat, D. Manongga, Y. Nataliani, S. Wijono, S. Y. Prasetyo, E. Maria, U. Raharja, I. Sembiring *et al.*, "Performance prediction using cross validation (gridsearchcv) for stunting prevalence," in *2024 IEEE International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*. IEEE, 2024, pp. 1–6.
- [40] B. Any, S. Four, and C. Tariazela, "Technology integration in tourism management: Enhancing the visitor experience," *Startupreneur Business Digital (SABDA Journal)*, vol. 3, no. 1, pp. 81–88, 2024.
- [41] N. Lutfiani, N. P. L. Santoso, R. Ahsanitaqwm, U. Rahardja, and A. R. A. Zahra, "Ai-based strategies to improve resource efficiency in urban infrastructure," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 121–127, 2024.
- [42] J. L. Willson, A. Nuche, and R. Widayanti, "Ethical considerations in the development of ai-powered healthcare assistants," *International Transactions on Education Technology (ITEE)*, vol. 2, no. 2, pp. 109–119, 2024.
- [43] P. A. Sunarya, U. Rahardja, S. C. Chen, Y.-M. Lic, and M. Hardini, "Deciphering digital social dynamics: A comparative study of logistic regression and random forest in predicting e-commerce customer behavior," *Journal of Applied Data Sciences*, vol. 5, no. 1, pp. 100–113, 2024.
- [44] S. R. P. Junaedi, D. Edmond *et al.*, "Successful digital marketing techniques for business development," *Startupreneur Business Digital (SABDA Journal)*, vol. 3, no. 1, pp. 19–25, 2024.
- [45] M. Mohanty, S. R. Jena, and S. K. Misra, "Mathematical modelling of engineering problems," *Journal homepage: <http://iieta.org/journals/mmep>*, vol. 8, no. 3, pp. 409–417, 2021.
- [46] N. Lutfiani, A. Ivanov, N. P. L. Santoso, S. V. Sihotang, and S. Purnama, "E-commerce growth plan for msme's sustainable development enhancement," *CORISINTA*, vol. 1, no. 1, pp. 80–86, 2024.
- [47] K. Mirdad, O. P. M. Daeli, N. Septiani, A. Ekawati, and U. Rusilowati, "Optimizing student engagement and performance using ai-enabled educational tools," *CORISINTA*, vol. 1, no. 1, pp. 53–60, 2024.
- [48] C. davis Davis *et al.*, "Artificial intelligence in education: Enhancing learning experiences through personalized adaptation," *International Journal of Cyber and IT Service Management*, vol. 4, no. 1, pp. 26–32, 2024.
- [49] U. Rahardja, A. Sari, A. H. Alsalamy, S. Askar, A. H. R. Alawadi, and B. Abdullaeva, "Tribological properties assessment of metallic glasses through a genetic algorithm-optimized machine learning model," *Metals and Materials International*, vol. 30, no. 3, pp. 745–755, 2024.
- [50] S. Edilia and N. D. Larasati, "Innovative approaches in business development strategies through artificial intelligence technology," *IAIC Transactions on Sustainable Digital Innovation (ITSIDI)*, vol. 5, no. 1, pp. 84–90, 2023.