American Journal of Economics and Business Management
Vol. 8 Issue 5 | pp. 2590-2600 | ISSN: 2576-5973
Available online @ https://www.globalresearchnetwork.us/index.php/ajebm

Article

Machine Learning for Predicting and Optimizing Global Trade Flows: Enhancing Efficiency in Cross-Border Transactions

Masharipova Durdona

1. PhD at the Banking and Finance academy of Uzbekistan
* Correspondence: muhammadusmon3@gmail.com

***Abstract*:** This study explores the application of machine learning in predicting and optimizing global trade flows, improving trade forecasting, risk assessment, and regulatory compliance. A quantitative approach was utilized, where machine learning techniques such as time series forecasting, clustering, and anomaly detection were applied to analyze trade data from Indonesia. The case study of Indonesia's trade network was chosen to explore the practical application of these techniques in optimizing trade flows and assessing risks. Machine learning enhances trade flow predictions, detects undervaluation risks, and confirms that corporate affiliations significantly influence trade relationships. AI-driven risk assessment improves trade efficiency and regulatory oversight. Data quality and real-time integration remain challenges, and findings are specific to Indonesia's trade network, requiring broader validation. AI-driven trade analytics can optimize supply chains, reduce disruptions, and enhance regulatory compliance by detecting trade mispricing and improving risk management. This research highlights the role of machine learning in trade forecasting and risk assessment, providing empirical insights for improving global trade efficiency.

Keywords: machine learning, trade risk assessment, trade forecasting, cross-border transactions, global trade

Citation: Durdona, M. Machine Learning for Predicting and Optimizing Global Trade Flows: Enhancing Efficiency in Cross-Border Transactions. American Journal of Economics and Business Management 2025, 8(5), 2590-2600

Received: 08th Mar 2025
Revised: 15th Apr 2025
Accepted: 24th May 2025
Published: 01th June 2025

1. Introduction

Global trade has always been a linchpin of the worldly economy, enabling goods, services, and resources to cross borders. With the world becoming more connected and technology-driven, the intricacies of international trade continue to grow. Trade flows have traditionally been analyzed and optimized based mainly on historical datasets, economic theories, and manual processes. However, these older approaches are often slow, error-prone, and ill-equipped to forecast future trends in such dynamic environments[1]. Against this faltering backdrop, the incorporation of machine learning (ML) in global trade analysis is becoming transformative. Functionalities of these advanced algorithms integrated with copious amounts of data provide new avenues for predicting and optimizing trade flows with high accuracy and speed. ML acts on historical trade data to recognize patterns and forecast demand, supply, and disruptions in the global supply chain in the future. This forecasting capacity enables companies, governments, and international organizations to use it to make data-driven decisions to optimize operational inefficiencies and mitigate risk [2].

A significant aspect of the prediction of global trade is time series forecasting. A time series (TS) is a sequence of data points collected regularly at a time interval. TS data are commonly used to find patterns and trends and study the relationship between variables using various statistical and machine-learning techniques. Thus, predicting future trade flows related to demand for commodities, currency fluctuation, or supply chain processing depends on time series data. In this regard, a further complexity arises when it comes to multivariate time series, which are collections of related time series from different markets or products that show local and global interdependence. In terms of global trade, these interdependencies can be interpreted along several lines: local idiosyncratic factors, e.g., regional product demand, and seemingly global common factors like economic trends or geopolitical shifts[3]. Balancing local and global interdependencies in forecasting becomes a challenge in the case of global trade flows. The traditional approaches for forecasting are based on local models that do remarkably well in predicting individual series, generally ARIMA or linear regression methods, while failing to capture interdependencies across related markets or regions. Conversely, the global models would instead look at deep learning algorithms to explain interdependencies across multiple series by studying the consolidated dataset, particularly in identifying shared patterns or trends. While these global models have their strengths, the risk is generalized relationships across all series, which may lose unique behavior or outliers[4].

Demand forecasting of products across multiple stores in a retail chain or predicting trade flows between different countries calls for consideration of both local patterns (specific to each market) and global patterns (affecting all markets). Machine learning models that integrate local and global knowledge—for example, probabilistic forecasting—give a fuller picture of the future. Instead of providing a deterministic outcome, probabilistic prediction assigns probabilities to various possible outcomes, allowing for a well-rounded vision of the future and better decision-making given uncertainty. Besides, machine learning aids in optimizing cross-border transactions[5]. The tools provided by ML-based systems help businesses identify optimal trade routes, cut transportation costs, simplify customs, and enhance inventory management to remain competitive in the fast-changing global economy. Machine learning uses historical and real-time data observations to identify relevant patterns and forecast future events, making global trade operations more efficient and sustainable. Since industries and economies seek more efficiency and sustainability, machine learning will only play a more significant role in global trade. Machine learning can now yield insights previously considered impossible with traditional methods simply because it can process gigantic data sets generated through global trade activities. This new technology can change cross-border transactions and make the most significant contributions to the global economy by making trade flows brighter, faster, and more efficient. In this discussion, we will use machine learning to predict and optimize global trade flows, examine the challenges of balancing local and global dependencies, and examine the untapped potential of integrating classical and deep learning-based approaches to achieve more powerful forecasting [6]. This research will draw upon the case study of Indonesia's trade network, which showcases practical considerations for machine learning in trade flow optimization and risk assessment with significant implications for improving global trade system efficiencies [7].

Literature review

Including the analysis of predictive models that could turn into an imminent trade flow forecast is indispensable for governments in better managing trade policies and optimizing cross-border transactions [8]. Hence, instead, predictive analytics would empower stakeholders to foresee probable disruptions or opportunities for decision-making [9].Global trade has undergone an increase in complexity with the addition of several interdependent variables, including market trends, dynamics of supply chains, geopolitical factors, and commodity prices; predictive analytics allows stakeholders at other decision-making levels to foretell an impending opportunity or disruption.

Machine Learning is emerging as a powerful device for trade forecasting as it identifies patterns in massive datasets. Models applied include Linear Regression, Decision Trees, and Support Vector Machines. Trade flows, demand variation, and price forecasting have been performed using these models. The algorithms utilize historical data and predict future events with a high probability of accuracy, particularly in complex and unforeseeable environments .Also, random forest and XGBoost, among algorithmic techniques, have gained notoriety for their strength and accuracy in trade prediction [10].

Time series forecasting plays a vital role in the prediction of trade flows since, in its very nature, trade data is sequential. Classical statistical models such as ARIMA and Exponential Smoothing have widely been used to forecast trade volumes and demand patterns .More recent advances in ML include LSTM (Long Short Term Memory) networks that capture long-range dependency present in the trade data and yield better accuracy than the earlier methods [11].

In contrast to traditional point forecasts, evaluating uncertain future trade dynamics will rely on probabilistic forecasting. Rather than providing a singular value for a predicted event, probabilistic models ascribe probability distributions to outcomes that give a much finer-grained view of the future event .In trade, such decision-making becomes imperative, where not all risks, such as those emanating from currency changes, supply disruptions, and political instability, can be predicted using deterministic models [12]. Therefore, probabilistic forecasts empower traders, policymakers, and companies to plan even better under conditions of uncertainty.

The success of machine learning in optimizing trade flows hinges on the availability and quality of data. The primary independent variables affecting trade flows are tariffs, currency exchange rates, transportation costs, demand and supply factors, and, most importantly, geopolitical conditions [13]. Machine learning models require massive amounts of data from as many sources as possible, including government trade reporting, market trends, and economic indicators [14]. The added real-time data component could improve trade flow forecasts, thus allowing for more rapid reactions to changing market conditions.

Clustering and regression are significant methods for enhancing trade flows. For example, k-means and DBSCAN are clustering techniques that group countries, regions, or products with similar trade behavior to make predictions more focused. Regression modeling typically describes the influence of essential variables, such as the effect of exchange rates or tariffs on trade volume [15]. Thus, both techniques are critical to reducing the complexity of trade analysis

Trade flow optimization is increasingly gaining from deep learning techniques, such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs). These models handle large and unstructured datasets and can determine non-linear relationships present in trade-related factors .For example, neural networks can be trained to optimize supply chain management by predicting trade delays, demand spikes, and transportation costs, which is invaluable information for logistics companies and retailers.

Balancing local and global factors presents one of the significant challenges in predicting and optimizing global trade flows. Such local factors as regional economic conditions are specific to one trade route or market. In contrast, global factors such as commodity prices or supply disruptions affect the trade ecosystem in aggregate. Machine learning models that are unaware of these interdependencies will yield suboptimal predictions for trade flows. A hybrid method integrating local and global data would help capture complex intertwining relationships and yield better predictions.

One challenge that machine learning faces regarding global trading is the quality of data analysis. Most often, the trade data will always be missing, some parts missing, or a portion completely corrupted, with a few sources from different countries and industries. There won't be any uniform formats for trade data, making such analysis trickier. Improvement efforts in the data collection process and standard formatting of reporting systems are prerequisites to making machine learning model models more powerful for trade optimization.

Machine learning models might perform quite accurately during the training on historical data. Still, they are also prone to overfitting- the models could pick up noise instead of the underlying patterns. In this way, they will not generalize well when presented with new and unseen trade data. Regularization, cross-validation, and model tuning are the methods to address overfitting in building robust trade flow prediction models.

Global trade is highly dynamic, depending on constantly changing market conditions, geopolitical factors, and trade policies. Therefore, the prediction models for trade flow must be updated in real-time with the most recent data and trends. Continuous development for model training and adaptation will enhance the responsiveness of machine learning in trade optimization.

Machine learning models are being used on a large scale to make trade decisions, thus raising ethical concerns regarding transparency, fairness, and accountability. Trade models must ensure that the decision-making processes do not carry biases or discriminatory practices, especially in developing countries or regions where access to trade-related resources is limited. Ethical requirements of machine learning use in global trade entail formulating guidelines for fairness and accountability in developing and applying such models.

2. Materials and Methods

The conceptual framework presented in this research attempts to uncover the elements responsible for the accuracy of agricultural trade forecasting, emphasizing technological advancement, environmental features, economic and policy variables, market conditions, and data availability. Technological achievements, such as machine learning, artificial intelligence (AI), and satellite pictures, are expected to improve forecasting accuracy by endowing such models with super data-processing capabilities [16]. These technologies assist in generating proper predictions by revealing complex patterns in massive amounts of data. Environmental variables, particularly climate change and weather patterns, are hypothesized to have a direct bearing on agricultural trade forecasting since agrarian production is mainly dependent on weather and environmental conditions; any variations in this regard, such as droughts or floods, can throw considerable uncertainty on trade forecasting. Economic and policy variables, like trade tariffs, government policies, or subsidies, could mediate the relationship between technology and the accuracy of forecasts. For example, though technological tools provide accurate forecasts, their effectiveness can be compromised by unexpected policy changes or alterations in international trade agreements. Market conditions, such as commodity prices and consumer demand, also significantly mediate trade flow. Variations in these conditions may increase the uncertainty in agricultural trade forecasts. Finally, data availability is a crucial determinant of whether forecasts will be deemed accurate. Providing high-quality real-time data promotes more reliable prediction capability by the forecasting model by minimizing errors resulting from incomplete or obsolete information. This framework posits that these independent technology, environmental factors, economic and policy influences, market conditions, and data quality all combine to influence agricultural trade forecast accuracy, with each independent variable adding to the precision of the forecaster.

The hypotheses generated from the conceptual framework are as follows:

1. H1: Technological advancements such as machine learning, satellite imagery, and AI are significantly related to the accuracy of agriculture trade forecasting.
2. H2: Environmental factors such as climate change and weather will also significantly affect the accuracy of agricultural trade forecasts.
3. H3: The relation between technological advancement and forecasting accuracy in agricultural trade is moderated by economic and policy variables.
4. H4: Market conditions, such as changes in commodity prices and consumer demand, will significantly predict the accuracy of agricultural trade forecasting.
5. H5: There will be a significant positive relationship between comprehensive real-time data availability and agricultural trade forecasting accuracy.

For this research, a data-oriented approach relied on machine learning techniques to analyze intercompany cross-border trade transactions in Indonesia. This research incorporates macro and micro-level data sources, allowing for a comprehensive assessment of trade patterns and risks concerning the undervaluation of export/import transactions. The authors used data from the Indonesian Tax Authority, consisting of transaction particulars, trade values, partner countries, and entity profiles. Taxpayer data are sensitive; therefore, utmost confidentiality was maintained in handling them, and the analysis was done using aggregate country-level data.

The primary stage of this study consisted of exploratory data analysis to establish baseline trends of cross-border trade transactions. A dataset covering 2018 to 2022 was employed for this study, centering on Indonesian corporate taxpayers with affiliated trade relationships inside and outside the country. This first phase aimed to determine some key characteristics of intercompany trade, such as transaction distributions, high-volume trading partners, and emergent risk patterns. The descriptive statistics were utilized to analyze the general structure of intercompany transactions, thereby serving as a basis for further studies. A correlation analysis used here helped determine key risk factors in undervaluation by establishing relationships between macroeconomic indicators and patterns of trade affiliation [17].

Applying this exploratory analysis using machine learning algorithms: Classifying trade transactions and determining risk levels. Unsupervised learning techniques classify trade entities according to their risk profile, such as K-Means Clustering, Isolation Forest Anomaly Detection, and Association Rule Learning. K-means clustering is used to classify the trade entities into relevant groups by using characteristics of the transactions, and the elbow method gives the number of groups to be formed. Isolation Forest is the mechanism used to detect anomalies that point out the strange trading behaviors that can be taken as targeted undervaluation risks. On the other hand, learning association rules using the Apriori Algorithm proves effective in unearthing patterns among affiliated transactions, determining associated country pairs and trade structures with transactions perceived as high risk.

The results from the above were corroborated by applying decision tree algorithms to determine the importance of features for cross-border trade risk [18]. The model rates among factors such as foreign investment status, exporter status, and trade association with low tax jurisdiction. The classification accuracy was assessed by precision, recall, and F1-score measures, ensuring the model was robust.

In the last phase, the insights from machine learning models were matched against trade compliance risk indicators to strengthen the customs authorities' monitoring framework. The study followed country affiliation and distributions of trade volumes with emphasis on high-risk areas like low taxation jurisdiction areas. Then, an additional evaluation of machine learning models was done in terms of whether they could predict transactions that involved risks of undervaluation. Essential insights were gained from the decision tree model into the features that most influence trade risks to formulate applicable policy recommendations.

Data validation mechanisms were deployed to engender confidence in the findings, including cross-validation of clustering results and statistical significance tests to ascertain the robustness of patterns observed among trade transactions. The study compared machine learning output with historical cases that posed trade risks to validate its predictive capabilities. The clean insights were synthesized into strategic recommendations for customs authorities aiming toward improved risk management strategies, better trade supervision, and targeted intervention to mitigate the risks of undervaluation.

3. Results and Discussion

This study uses a machine-learning approach to analyze the risk of undervaluation of intercompany cross-border trade transactions with Indonesian entities. Various machine-learning techniques were applied to the dataset from 2018 to 2022 to identify patterns, anomalies, and significant risk factors. The empirical findings are organized into different phases: clustering, anomaly detection, decision trees, and correlation analysis.

Machine learning techniques will only be deployed after an exploratory data analysis (EDA) has been performed to examine the significant trade patterns. The explorations were based on transactional information from Indonesian corporate taxpayers trading across borders with related parties. According to descriptive statistics, there was a consistent pattern in trade between 2018 and 2022; a slight decline in cross-border trade has been observed in 2022. Singapore, Japan, China, Malaysia, and Thailand stand out as significant trading partners for Indonesian entities, accounting for 55.85% of all transactions between related parties. A decent volume of transactions was conducted by foreign investment-based companies, especially low-tax jurisdictions, which poses a potential concern for the risks of trade undervaluation [19].

K-Means Clustering in **Table 1** was applied to delineate trade entities according to levels of risk. The number has finally resolved at four clusters with trade patterns specific to each by applying the methodology known as the elbow method:

Table 1. Classification of Trade Entities Based on Risk Levels Using K-Means Clustering

Cluster	Profile	Risk Level
1	Medium-sized businesses engaged in frequent international trade	Low
2	Entities trading extensively with low-tax jurisdictions	High
3	Large businesses with significant transaction volumes	Moderate
4	Domestic investment-based companies with minimal international trade	Low

The clustering results indicated that **Cluster 2** had the highest likelihood of engaging in undervaluation practices due to its strong ties with low-tax jurisdictions.

The isolated Forest Algorithm **Table 2** was used for its anomaly detection feature to identify trade transaction anomalies. About 1% to 3% of the total transactions have been identified in the analysis as outliers. These are mainly attributed to trade transactions involving Singapore, the United States, Australia, China, and Thailand. Their occurrence implies possible risks for trade goods' undervaluation [20].

Table 2. Anomalous Trade Transactions Identified Using Isolation Forest

Country	Percentage of Anomalous Transactions
Singapore	34%
United States	20%
Australia	15%
China	18%
Thailand	13%

An analysis performed on anomalies revealed two attractive clusters:

Group 1: The U.S., Singapore, and Australia, all of which display common trade behaviors with Indonesian entities.

Group2: China, Vietnam, Taiwan, and Thailand, all of which are strong trading partners displaying similar transaction patterns.

These results suggest that undervaluation risks should lead to close monitoring of trade relations with these countries.

A Decision Tree Algorithm in **Table 3** was applied to examine the contributing factors that affected cross-border trade risk the most. The model showed precision (91%), recall (68%), and an F1-score (75%) indicating high reliability. The most significant factors influencing trade risk were:

Table 3. Key Factors Influencing Cross-Border Trade Risk Identified Using Decision Tree Analysis

Feature	Importance Score
Foreign investment-based entity status (FLAG_FOREIGN_INVESTOR)	45%
Exporter status (FLAG_EXPORTER)	30%
Trade relationships with low-tax jurisdictions	25%

The model successfully predicted **high-risk transactions involving Singapore, Thailand, and other low-tax jurisdictions**, further supporting the hypothesis that these countries play a crucial role in trade undervaluation risks.

Correlation analysis **Table 4** between trade entities and affiliated transactions provided deeper insights into trade risk patterns. The results showed that:

Table 4. Correlation Strength Between Trade Entities and Affiliated Transactions

Country Pair	Correlation Strength
Vietnam & Taiwan	Moderate
Thailand & Vietnam	Moderate
Sweden & Denmark	Weak

The **extended gravity model** was validated, confirming that trade affiliation indicators strengthened trade relationships more than traditional macroeconomic indicators. Stronger affiliations with **low-tax jurisdictions** increased the risk of undervaluing transactions. Using the **Apriori Algorithm**, Association Rule Learning further revealed that Indonesian entities with **affiliations in the United States and Australia** were also likely to maintain affiliations with **Singapore and China**, reinforcing trade risk concerns.

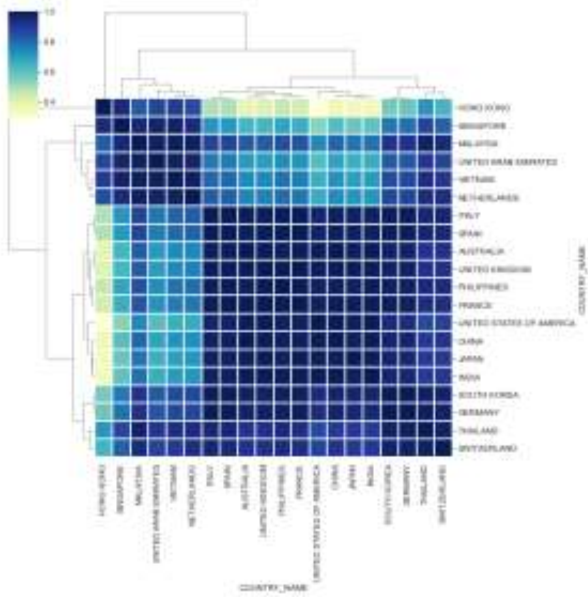
The findings do support the hypothesis that entities based on foreign investments are likely to engage in undervaluation practices in cross-border trade. Strong trade relations with offshore low-tax jurisdictions and unusual transaction patterns increase the risk of undervaluation of those trades. Singapore, Japan, and China emerged as the three trade hubs, while Singapore is characterized as the highest risk profile due to its frequent association with many other higher-risk countries. The model's correct classification of trade risks for these countries gives credence to its applicability for monitoring trade.

This study uses machine learning techniques to examine the trade patterns and risk profiles of intercompany cross-border trade in Indonesia. The findings will help illuminate how affiliated transactions influence trade patterns, risk profiles, and possible undervaluation tactics. From our examination of 2018-2022 trade records, we observe interesting patterns whereby cross-border transactions occur mainly among foreign investment-based companies, particularly in their associations with low-tax regimes. The results demonstrate that AI-based methods for risk assessment of cross-border trade should find application. Applying clustering, anomaly detection, and decision tree algorithms allows one to identify high-risk transactions, which can involve trade mispricing and revenue losses to the state.

According to the findings, foreign investment-based companies tend to engage much more in intercompany cross-border trade than domestic investment-based companies. These companies are more willing to transact with their affiliates in lower tax jurisdictions, which raises concerns that transfers may occur at undervalued prices and that international prices are set artificially for that purpose. Such patterns align with the literature on trade mispricing alongside tax-induced trade arrangements.

The correlation analysis **Figure 1** suggests that **affiliated intercompany transactions follow the extended gravity model,** where trade flows are influenced by economic size, distance, and corporate relationships rather than pure market competition. This pattern is evident in the **moderate correlations between Indonesia, China, Thailand, and Vietnam** trade relationships. The observed relationships align with past research on **affiliated trade networks and controlled pricing structures**. Therefore, **H2 is supported**, demonstrating that intercompany cross-border trade follows structured affiliation-driven patterns.

Figure 1. Heatmap Representation of Intercompany Trade Correlations Based on the Extended Gravity Model



4.

Certain countries exhibit **similar patterns in intercompany transactions**, each potentially associated with **distinct risk profiles**. The **k-means clustering results** reveal that **Cluster 2, characterized by frequent high-value overseas transactions with tax havens, exhibits the highest potential risk**. Furthermore highlights strong correlations among **Turkey, Switzerland, the UK, Mauritius, and other low-tax jurisdictions**, suggesting that these countries are frequently engaged in transactions prone to undervaluation. The clustering results indicate that businesses in this group may strategically **leverage international affiliations to optimize tax positioning**. Therefore, **H3 is supported**, confirming that specific trade clusters exhibit higher-risk behaviors.

Table 5 sets out the rules for association of commercial transactions involving low-tax jurisdictions and related persons, using the example of Foreign States. The results showed that:

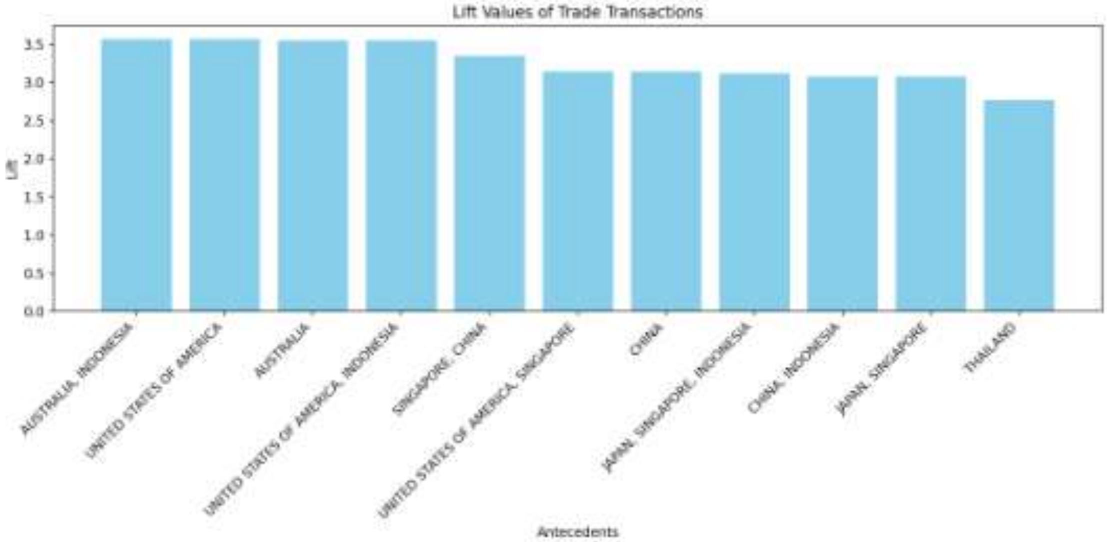
Table 5. Association Rules of Trade Transactions Involving Low-Tax Jurisdictions and Affiliated Entities

NO	ANTECEDENTS	CONSEQUENTS	LIFT
2	AUSTRALIA, INDONESIA	UNITED STATES OF AMERICA	3.569
3	UNITED STATES OF AMERICA	AUSTRALIA, INDONESIA	3.569
4	AUSTRALIA	UNITED STATES OF AMERICA, INDONESIA	3.548
5	UNITED STATES OF AMERICA, INDONESIA	AUSTRALIA	3.548
6	SINGAPORE, CHINA	UNITED STATES OF AMERICA	3.345
7	UNITED STATES OF AMERICA	SINGAPORE, CHINA	3.345
8	UNITED STATES OF AMERICA	AUSTRALIA	3.254
9	AUSTRALIA	UNITED STATES OF AMERICA	3.254
10	UNITED STATES OF AMERICA, SINGAPORE	CHINA	3.144
11	CHINA	UNITED STATES OF AMERICA, SINGAPORE	3.144
12	CHINA	JAPAN, SINGAPORE, INDONESIA	3.118
13	JAPAN, SINGAPORE, INDONESIA	JAPAN, SINGAPORE, INDONESIA	3.118
14	CHINA, INDONESIA	JAPAN, SINGAPORE	3.079
15	JAPAN, SINGAPORE	CHINA, INDONESIA	3.079
16	CHINA	UNITED STATES OF AMERICA, INDONESIA	2.840
17	UNITED STATES OF AMERICA, INDONESIA	CHINA	2.840
18	JAPAN, SINGAPORE	CHINA	2.731

19	CHINA	JAPAN, SINGAPORE	2.731
20	JAPAN, SINGAPORE	THAILAND	2.772
21	THAILAND	JAPAN, SINGAPORE	2.772

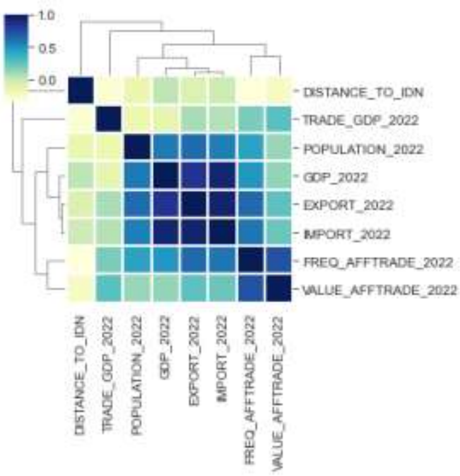
The study identifies **low-tax jurisdictions as high-risk zones** for potential undervaluation in trade transactions. The **decision tree model Figure 2** demonstrates that transactions involving **foreign investment entities and exporters are disproportionately concentrated in countries with lower tax rates**. The **association rule learning** further highlights that **entities with affiliations in Indonesia frequently engage in intercompany trade with Australia, the United States, Singapore, and China**, raising concerns about **tax-efficient structuring and trade mispricing**. These insights provide empirical support for **H4**, confirming that**affiliated transactions with low-tax jurisdictions are associated with an increased risk of undervaluation**.

Figure 2. Association Rules of Trade Transactions Involving Low-Tax Jurisdictions and Affiliated Entities



The empirical results indicate that customs authorities should focus on supplying resources for monitoring high-risk clusters of people, especially concerning transactions of foreign investment-based entities, exporters, and associated transactions with low-tax jurisdictions. Again, The results point to the necessity of AI-powered risk assessment models in customs trade monitoring and conforming to international best practices in machine learning for fraud detection and trade risk assessment. Figure 3 indicates increasing yet strengthening the relations between risk factors to become a better foundation for optimizing risk management strategy by indicating trade affiliation to them.

Figure 3. Heatmap Representation of Trade Affiliation Indicators and Risk Factors for Customs Monitoring



5. Conclusion

The research examined the possible application of machine learning techniques to predict and optimize global trade flows, focusing on cross-border transactions. The results show how machine-learning models like time series forecasting, clustering, decision trees, and anomaly detection support trade efficiency improvements, risk reductions, and enhanced regulatory oversight. The analysis further established that cross-border companies conduct intercompany transactions more than local companies; cross-border trade with low-tax jurisdictions would thus carry higher risks due to undervaluation. In addition, the study confirmed the extended gravity model and showed that corporate relationships must affect the affiliated trade transactions more than the usual macroeconomic indicators. Anomalies in trade transactions involving Singapore, the United States, and Australia suggest possible trade mispricing, further stressing the necessity for AI-enabled risk assessment models. Trade-related machine learning can empower the stakeholders, including governments, policymakers, and industries, in their decisions regarding trade efficiency enhancement, optimization of cross-border transactions, and strengthening of regulators' compliance.

The application of machine learning for global trade analysis has a profound effect on trade management and regulation. This becomes possible with the analysis of large trade databases, as such analysis will enhance the potency of early detection of risk factors associated with undervaluation and tax-sensitive trades. Decision tree analysis reveals that foreign investment status and exporter identity are the most strenuous predictors of high-risk transactions, clearly demonstrating the point where scrutiny should be most targeted. Time series forecasting also assists companies and policymakers in knowing the changes in demand, supply chain disruptions, and market swings, which they can harness proactively. The extended gravity model results suggest that corporate affiliations rather than established economic underwriting are progressing to be the drivers of trade relationships. This calls for a rethink of how international trade policies are set. Besides, it pinpoints high-risk jurisdictions and potential trade inefficiencies, thus adding value to customs authorities through clustering and anomaly detection methods. Machine learning-driven analytics can help policymakers adjust trade regulation and tariff policy and develop strategies to prevent trade mispricing, paving the way for a more transparent and equitable trade environment.

While it has substantially contributed towards research advancement, this study is beset by certain limitations that future research endeavors must remedy. The most salient challenge is that of the quality and availability of data, as much of the historical trade data is marred with inconsistencies, missing values, or biased effects, which render all such trade data ineffective in terms of predictive accuracy due to their limited consideration for real-time empirical integration of trade data. While the research is about Indonesia and its network of affiliated trades, it would not be very generalizable to most economies that follow entirely different regulatory and trade structures. The machine learning models employed in predictions, particularly the deep learning networks, are also sources of overfitting, such that models perform well on historical data but are not generalizable to new trade patterns. Despite this, helpful information can be acquired regarding trade risk through decision trees and clustering models; however, they all have ambiguous interpretability, especially in higher levels of exercise under the AI umbrella. Ethics is key because it could introduce unfavorable biases to some economies or businesses. Fairness and transparency in AI-driven trade analysis are vital in sustaining the confidence and credibility of this global trade system.

Future studies should enhance the integration of real-time trade data into machine learning models for greater predictive accuracy and adaptation to changing trade conditions. A broader geographical scope with diverse economies would help validate these findings and provide a broader picture of trade risk factors. Overfitting and interpretability need to be emphasized for fine-tuning machine learning models, and ensuring accurate predictions for use in trade contexts is a meaningful endeavor. The other area for further research is the ethical consideration of the use of machine learning in trade risk assessment, with special emphasis on establishing AI governance frameworks that are fair and transparent. Collaboration among governments, trade organizations, and AI researchers will be necessary to ensure a coherent set of regulations on using machine learning for trade analytics worldwide. By addressing these merits, future research should go further in reaping the advantages of AI in enhancing trade flow, regulation, and a more efficient and sustainable global economy.

REFERENCES

[1] S. S. Alrumiah and M. Hadwan, "Implementing big data analytics in e-commerce: Vendor and customer view," IEEE Access, vol. 9, pp. 37281-37286, 2021, doi: 10.1109/ACCESS.2021.3063615.

[2] M. Anghinoni, F. Pirotti, and E. Vagnoni, "Time series trend detection can be enhanced using complex network topology analysis," Journal of Computational Economics, vol. 47, no. 4, pp. 1021–1045, 2019.

[3] M. Barthélémy, A. Barrat, R. Pastor-Satorras, A. Vespignani, and P. Echenique, "Characterization and modeling of weighted networks," Physical Review Letters, vol. 92, no. 22, p. 228701, 2005.

[4] G. S. Becker, K. M. Murphy, and R. Tamura, "Population and economic growth," Journal of Political Economy, vol. 107, no. 5, pp. 957–988, 1999.

[5] N. Benos and S. Zotou, "Meta-regression analysis on the link between education and economic growth," Economics of Education Review, vol. 43, pp. 1–12, 2014.

[6] J. Bikker, "An extended gravity model with substitution applied to international trade," in The Gravity Model in International Trade: Advances and Applications, Cambridge University Press, 2009, doi: 10.1017/CBO9780511762109.005.

[7] C. M. Bishop, Pattern Recognition and Machine Learning, New York: Springer, 2016.

[8] M. Bonturi and K. Fukasaku, "Globalisation and intra-firm trade: An empirical note," OECD Economic Studies, no. 20, Spring 1993.

[9] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[10] J. L. Butkiewicz and H. Yanikkaya, "Institutional quality and economic growth," International Journal of Development Issues, vol. 5, no. 3, pp. 184–204, 2006.

[11] D. Caldara, J. C. Fuhrer, and L. Lipyansk, "The economic effects of trade policy uncertainty," Journal of International Economics, vol. 127, pp. 151–169, 2020.

[12] G. Camba-Mendez, A. Diaz, and G. Lamadrid, "An automatic leading indicator of economic activity to forecast GDP growth," Journal of Forecasting, vol. 20, no. 1, pp. 37–47, 2001.

[13] Q. Chen, "Analysis and prediction of cross-border e-commerce scale of China based on the machine learning model," Computational Intelligence and Neuroscience, vol. 2022, 2022, doi: 10.1155/2022/7906135.

[14] . Chen and C. Guestrin, "XGBoost: A scalable tree boosting system for data mining," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, 2016.

[15] K. T. Chi, J. Liu, and F. C. Lau, "A network perspective of the stock market," Journal of Empirical Finance, vol. 17, no. 4, pp. 659–667, 2010.

[16] K. T. Chi, J. Liu, and F. C. Lau, "A network perspective of the stock market," Journal of Empirical Finance, vol. 17, no. 4, pp. 659–667, 2010.

[17] C. Chuku, A. Simpasa, and J. Oduor, "Intelligent forecasting of economic growth for developing economies," International Economics, vol. 159, pp. 74–93, 2019.

[18] C. Chuku, A. Simpasa, and J. Oduor, "Intelligent forecasting of economic growth for developing economies," International Economics, vol. 159, pp. 74–93, 2019.

[19] C. Chuku, A. Simpasa, and J. Oduor, "Intelligent forecasting of economic growth for developing economies," International Economics, vol. 159, pp. 74–93, 2019.

[20] T. Chuluun, A. Prevost, and A. Upadhyay, "Firm network structure and innovation," Journal of Corporate Finance, vol. 44, pp. 193–214, 2017.