JOURNAL OF
EMERGING INVESTIGATORS

Article

Predicting and explaining illicit financial flows in developing countries: A machine learning approach

Akshith Reddy Putta[1], Jeremy Bigness[2]
[1] Coppell High School, Coppell, Texas
[2] Agilent Technologies, Inc., Santa Clara, California

SUMMARY

Cross-border corruption and the illicit movement of financial assets, referred to as illicit financial flows (IFFs), have a strongly deleterious effect on the economies of developing nations. Over the past 20 years, there has been a concerted international effort to mitigate cross-border corruption; however, the most important economic and political factors leading to IFFs are unclear. In this work, we used multiple machine learning (ML) approaches—including linear regression, logistic regression, random forests, and neural networks—to predict the levels of corruption using various economic and political measures from the years 2009 to 2018. Furthermore, to make clear the relative importance of these factors, we used several ML model interpretation tools. We hypothesized that the artificial neural network (ANN) machine learning model can most effectively predict and explain IFFs in developing countries using economic and political indicators. Out of the various regression ML models, the ANN had the most success in predicting the IFFs, with a Pearson correlation coefficient of 0.97. The most important features, as quantified using Shapley values from the ANN and the feature importances of the random forest models were: aid percent of gross national income, population, human development indicator income, and government efficiency. Taken together, these models and their interpretation provide a method for predicting the IFFs as well as the features that drive them, enabling policymakers to focus on these factors to decrease corruption.

INTRODUCTION

Illicit financial flows (IFFs) are defined as any transfer of money that has been earned, moved, or stored illegally, posing a major challenge to developing countries (1). These funds often pass through the control of corrupt officials or corporations, diverting resources away from benefiting the country's citizens and ending up in tax havens. Tax havens are jurisdictions with low tax rates for non-residents, who tend to store their money there (2). IFFs are transported from these developing countries to tax havens to be stored. These tax havens are usually in small and affluent countries (2). It has been estimated that financial assets equal to 10% of global gross domestic product (GDP) are held by individuals in tax havens (3). Since hiding financial assets through IFFs is practiced mostly by wealthy and powerful members of society in developing countries, this practice may lead to widening financial inequalities that arise from such corruption, in addition to other social issues (4). Additionally, knowledge of others evading taxes may lead to less tax compliance among the general populace (5).

At the national level, analyzing and discovering indicators of corruption can enable leaders to address the problem of illegal international money flows through policy. As an example, if certain factors are linked to a high amount of IFFs, policymakers can address this issue accordingly. Estimating levels of IFFs of different regions could allow policymakers to identify regions with high levels of corruption, and it may also allow businesses to estimate the difficulties they will face when operating in a region.

Due to this substantial issue, large supranational organizations such as the United Nations (UN) and the Organization for Economic Cooperation and Development (OECD) have made an effort to address the issue of IFFs (6). Forty-four countries have signed the OECD Anti-Bribery Convention, which came into effect in February 1999 (6). The convention targets bribery of foreign public officials to foster equal grounds for international bribery prevention efforts. Similar anti-IFF efforts have been undertaken by the UN, with the adoption of the United Nations Convention Against Corruption (UNCAC), which came into effect in December 2005 (6).

Prior works have also used machine learning (ML) to predict corruption, both at the level of individual countries and globally (7–9). For example, López-Iturriaga *et al.* focused on predicting corruption across Spanish provinces using features such as regional government expenditure, debt levels, and local political variables (9). Similarly, Lima *et al.* examined corruption in Brazil by analyzing municipal audit results and political alignment factors (7). On the global scale, ML models are applied using cross-country datasets, though their focus remained largely on national corruption indicators and overlooked cross-border flows of financial assets, a crucial component of illicit financial flows (IFFs) (8). These local studies primarily target within-country corruption patterns, relying on area-specific features such as governmental reports and data collected by local governments, which limit the generalizability of their findings. Even the existing global corruption studies tend to overlook the cross-border dimension of IFFs, focusing instead on national-level corruption metrics without fully addressing how illicit assets move internationally. Currently, there is insufficient research examining how broad political and economic metrics, such as those used in this study, influence IFFs across developing countries.

There are several adjacent studies relevant to unlawful economic activities more broadly (10–13). These studies include which applying machine learning to illicit trade flows, money laundering detection through financial analysis, and

https://doi.org/10.59720/24-030

bribe payments in India (10–13). Although these studies are not directly focused on cross-border IFFs in the sense of illicit money transfers between nations, they provide important theoretical background suggesting that unlawful economic behaviors may indeed be influenced by political and economic metrics. Recognizing this connection further supports the importance of our research focus on IFFs, while also highlighting the gap: few, if any, studies have systematically explored how these broad national-level metrics shape the dynamics of cross-border illicit financial flows specifically. By situating our work within both the direct IFF literature and these adjacent domains, we aim to provide a stronger theoretical foundation for understanding how political and economic factors can affect financial crime patterns at the international level.

For the dataset features, we obtained political indicators defined by the Worldwide Governance Indicators (WGI), such as voice and accountability, political stability, absence of violence/terrorism, government effectiveness, and others (14). Global Financial Integrity (GFI) is a think tank that keeps track of IFFs and various other financial crimes, and this is where we acquired the data about the levels of IFFs, measured in millions of United States dollars (USD) (1, 15, 16). Other social and economic factors in this dataset were defined by a previous study (**Table A1**) (17). We chose these datasets because they offer broad metrics that could be easily calculated. For the purpose of standardization, we limited the analysis to the years 2009 to 2018. These years were selected because this is the range in which data from all three datasets overlap. For all three datasets, the values are collected every year for all countries.

To identify factors leading to high IFFs, we used multiple types of ML models—including linear and logistic regression, random forests, and neural networks—to predict IFFs given an initial set of features. We used increasingly more complex models as we proceed with our analysis, starting with simpler linear and logistic regression, then random forest models, and finally neural networks. ML is a branch of artificial intelligence that can be used to analyze datasets and uncover patterns that would be difficult to detect using traditional methods.

Linear and logistic regression models are both valued for their simplicity and interpretability, making them useful tools for analyzing economic and political factors related to illicit financial flows (IFFs). Linear regression provides a clear baseline by modeling linear relationships between input features and a continuous target variable, with coefficients offering direct insights into the strength and direction of associations. It has been successfully applied in cross-country corruption studies, highlighting its utility in related domains (8). Logistic regression, commonly used for binary classification, estimates the likelihood of outcomes such as high or low IFF levels based on economic and political indicators (8). It assumes linearity in the log odds and has been effectively employed to model behaviors like tax compliance, shedding light on patterns of financial misconduct (5). However, while both models offer ease of interpretation, their reliance on linear assumptions limits their ability to capture complex, nonlinear interactions within multifaceted datasets.

Random forest and logistic regression models offer complementary strengths for analyzing corruption and illicit financial flows (IFFs). Random forest regressors are well-suited for capturing nonlinear relationships and complex interactions by aggregating multiple decision trees, offering robustness against overfitting and providing feature importance measures such as mean decrease in impurity (MDI) (8). Though they can be computationally intensive and less interpretable than linear models, they have been effectively applied to corruption prediction in diverse contexts, including Brazil and global analyses, demonstrating their ability to model complex socioeconomic dynamics (7, 8). In contrast, logistic regression is valued for its simplicity and interpretability in binary classification tasks, making it useful for identifying factors associated with high or low IFF levels. It assumes linearity in the log odds, which enables straightforward estimation based on political and economic indicators (8), and has been used in prior research to predict tax compliance and related behaviors (5). However, its linear assumptions can limit its effectiveness in modeling the intricate relationships found in large, multifaceted datasets. Artificial neural networks (ANNs), both regressors and classifiers, have been used to detect corruption by capturing complex, nonlinear relationships between regional economic and political features and outcomes such as illicit financial flows (IFFs). For instance, ANN models have been applied in Spain to predict provincial-level corruption using such regional indicators (9). These models excel at learning intricate patterns from data and offer strong predictive performance. However, they often require substantial training data and are commonly viewed as “black-box” models due to their limited interpretability (21). To address this limitation, explainability techniques like SHapley Additive exPlanations (SHAP) have proven effective in offering detailed insights into the contribution of individual features to model predictions (17, 21). As demonstrated across these studies, machine learning models offer significant advantages for identifying illicit financial flows (IFFs), particularly through their ability to capture complex relationships that traditional methods struggle to address. Despite these strengths, prior research has primarily focused on national-level corruption or financial fraud detection, often relying on region-specific datasets with limited applicability to cross-border transactions. Additionally, studies seldom incorporate broad economic and political indicators to analyze systemic IFF patterns. By addressing these gaps, we aimed to provide a more comprehensive approach to understanding the drivers of IFFs on a global scale.

We addressed this gap by clarifying the relationship between broad political and economic metrics and the levels of IFFs, offering insights into which factors facilitate cross-border illegal transactions. We used a combination of interpretable models, such as regression and random forests, as well as neural networks supplemented with Shapley values, to determine the most important factors driving each model’s predictions. We hypothesized that the ANN ML model could most effectively predict and explain IFFs in developing countries using these features. This is because ANNs, when paired with SHAP, offer high explainability, can learn from high-dimensional data, and model complex relationships, even in noisy datasets like the one used in this study. Our hypothesis was supported by the results: the ANN regressor and classifier outperformed the other models. This superior performance arises from the ANN’s layered structure, which allows it to capture intricate patterns, automatically extract relevant features, and reduce the need for manual feature

https://doi.org/10.59720/24-030

engineering — an important advantage given the risk of overfitting and error in traditional algorithms.

RESULTS

To evaluate the predictive performance of various machine learning models in identifying key economic and political factors influencing IFFs, we trained and tested multiple models, including linear regression, random forests, and ANNs. Model performance was assessed using the Pearson correlation coefficient for regression tasks, given its ability to capture both the direction and strength of relationships between predicted and actual values. Additionally, we employed SHAP to interpret model outputs and determine the most influential features contributing to predictions. By applying these models to a dataset comprising economic and demographic indicators, we aimed to uncover the primary drivers of IFFs and assess which algorithm best captures their complex, non-linear relationships.

The most important features, according to the SHAP study above, for the performance of the models are aid percent of gross national income (GNI), population, and gross domestic product per capita (GDPPC) (**Figure 1**). Note that this was done with the same dataset that has been used for training the ML models.

We first examined the overall distribution of IFF values, and the key factors most strongly correlated with them to contextualize the subsequent modeling results. Some interesting trends in the IFF distribution include a large concentration of values around 350 million USD, with apparent outliers reaching approximately 6,500 million USD (**Figure 2**). Among the most influential factors in the models are aid as a percentage of GNI and population, both of which show strong correlations with IFF levels (**Figure 1**). Aid as a percentage of GNI represents the proportion of a country's gross national income derived from financial aid, and while most countries fall within the 0–5% range, certain outliers—such as one around 38%—receive exceptionally high levels of aid. Population, on the other hand, has the highest Pearson correlation coefficient among all features and is the most important factor in the ANN model (**Figure 3B**).

The Pearson correlation coefficient was used to gauge the accuracy of the regression models because it provides a more comprehensive and interpretable assessment of regression model performance, considering both the direction and strength of the relationship between predicted and actual values, when compared to metrics like the accuracy and mean squared error (7). Linear regression had a Pearson correlation coefficient of 0.850. The random forest regression model performed better, with a Pearson correlation coefficient of 0.856. The ANN regressor model achieved the highest Pearson correlation coefficient of 0.970 (**Figure 4**).

The F-1 score metric was chosen to gauge the classification models because it has a balance between precision and recall, giving a complete picture of a classification model's performance (7). The logistic regression performed at the same level as the linear regression, with a weighted average F-1 score of 0.23. Therefore, it can be concluded that logistic regression does perform as well when trying to predict values binarized by the mean, due to outliers greatly skewing the model. Random forest had good accuracy on the dataset, with the first and third quartiles having the best individual F-1 scores of 0.94. Quartiles 2 and 4 were predicted less accurately, with F-1 scores of 0.90 and 0.91, respectively. The last classification model was the ANN classifier, which had a weighted average F-1 score of 0.94 (**Figure 5**).

Model feature importances were acquired from SHAP for three separate models: the random forest classifier, random forest regressor, and ANN (**Figure 1, 6**). For the random forest models, only the features that contributed the most to the MDI (population, aid percent of GNI, and GDP per capita), which is sometimes called the Gini value, are displayed (**Figure 6**). These three important features in the random forest Regressor and Classifier were also important in the ANN (**Figure 1**). Therefore, it is logical to conclude that the features that led to an increase in the rate of IFFs in the ANN model would do the same in the random forest regressor. The most important features were selected according to their magnitudes of MDI contribution along all three models, though the ANN was given the highest precedence. Also, these features were common across all three models as being the most important.

The most important features that predict a higher rate of IFFs were population, aid percent of GNI, human development index (HDI) income, and government effectiveness (**Figure 1)**. There are other factors that play a smaller role, such as the HDI education ratings for each country, and the percent of population which has vulnerable employment (not having consistent wages, etc). The most important features that predicted a lower rate of IFFs were the voice and accountability, control of corruption, and political stability.

DISCUSSION

Our results indicate that the ANN outperforms both linear regression and random forest models in predicting IFFs, supporting our hypothesis that a deep-learning approach is best suited for capturing the complex relationships between economic and political indicators. The ANN regressor achieved a Pearson correlation coefficient of 0.970, significantly higher

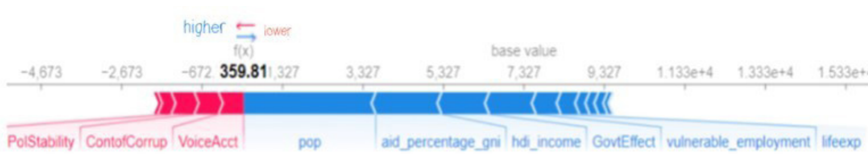

Figure 1: Plot of the Shapley values for the artificial neural network regressor. Values closer to 0 signify that the feature has less impact on the predicted outcome, while larger absolute values indicate stronger influence. Negative Shapley values suggest that the presence of the feature decreases the predicted rate of IFFs, whereas positive values indicate an increase. The bolded number, 359.81, represents the average predicted rate of IFFs across the dataset and is a reference point for understanding the impact of individual features. The light grey numbers accompanying the arrows denote increase or decrease of IFFs associated with each feature, highlighting their relative importance in contributing to the model's predictions. The light grey text is the actual IFF values by average. In the context of the red text "lower," the red features lead to lower levels of IFFs, and vice versa for the blue text "higher".

https://doi.org/10.59720/24-030

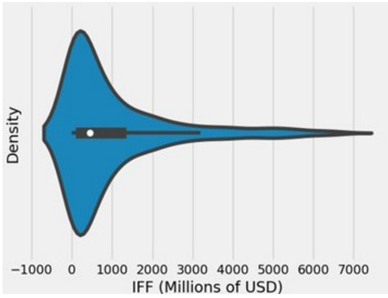

Figure 2: Violin plot of the illicit financial flows (IFFs). Distribution of the IFFs. The white dot is the median of all the values, and the black rectangle represents the interquartile range (IQR) of the data distribution. It spans from the first quartile (25th percentile) to the third quartile (75th percentile), with a line inside representing the median (50th percentile). A kernel density estimation (KDE) was used to estimate the probability density function of the data at different points along the x-axis. The graph was produced with the default values from the Seaborn library. This data was obtained from previously published data (11).

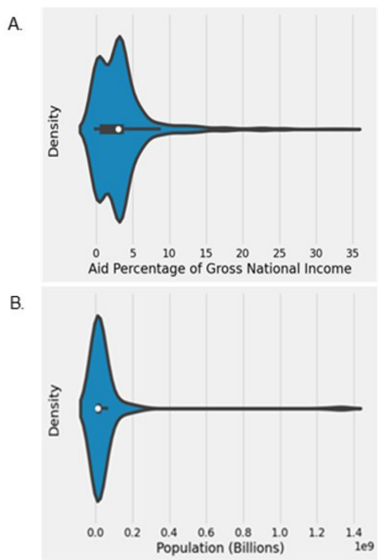

Figure 3: Violin plots of the aid percentage of gross national income (GNI) and population. Distribution of **A)** aid percent of GNI and **B)** population. The graphs were produced with the default values from the Seaborn library. These features were obtained from reference 12.

than that of the linear regression (0.850) and random forest regression (0.856). This demonstrates that while simpler models can identify general trends, they struggle to capture intricate nonlinear interactions present in the data. The ANN classifier also exhibited the highest weighted average F1-score (0.94), outperforming logistic regression and random forest classification.

The performance gap between the models can be attributed to their respective abilities to handle high-dimensional and nonlinear relationships. Linear regression, while interpretable, assumes a linear relationship between inputs and outputs, limiting its effectiveness in modeling IFFs, which must take multiple inputs into account. While random forest models improved performance through ensemble learning, they still lacked the advanced feature extraction capabilities of ANNs. This limitation affects their effectiveness, as random forests rely on pre-engineered or raw input features without transforming them into higher-level representations. The ANN's ability to automatically detect hidden patterns and interactions, coupled with its robustness to outliers and noisy data, allowed for superior predictive accuracy. Additionally, SHAP analysis revealed that key features such as population, aid percent of GNI, and government effectiveness had a strong impact across all models, reinforcing their importance in understanding IFFs. However, the ANN model's usage of hidden layers and nonlinear transformations allowed it to more effectively discern relationships between these features (16).

These findings highlight the necessity of leveraging deep learning techniques for analyzing financial crime patterns, particularly in contexts where traditional statistical methods fall short. However, while ANNs provided the most accurate predictions, they also pose interpretability challenges compared to regression and random forest models, due their nature of being "black box" models (16). Future research could focus on refining ANN architectures while enhancing their explainability through advanced feature attribution techniques like SHAP to improve transparency in policymaking applications.

We speculate that the distribution of the IFFs, having a large concentration at the first quartile and then decreasing as IFF levels increase, has occurred because of two cases with developing countries (**Figure 2**). The larger concentration is the vast majority of developing countries, which have low IFF levels because the economy of these countries is too small, or the government has cracked down effectively upon these activities, for example Ecuador and Bulgaria (6). Another possibility might be that political officials, through public companies, might transport money overseas. Money transfers by public officials through public companies does not count in IFF statistics, since IFF statistics take into account 'the spirit of the law' when calculating these statistics (1).

The most important feature, population, should not be considered as a strong indicator for the prevalence of corruption. Instead, it serves as a scaling mechanism for the economy of countries (6). The second important feature was the aid percent of GNI (**Figure 1, 6**). The country with the highest rate of this is Yemen, followed by similar war-stricken low-income countries, mostly in Africa and the Middle East. This might be due to increased political instability, terrorist activities, and illegal trade (18). The importance of the amount of aid percent of GNI, which is high in these countries due to humanitarian crises that accompany political instability, terrorist activities, and illegal trade, might be due to the foreign aid being embezzled by the people in charge, likely contributing to the ineffectiveness of foreign aid (18). When aid money is embezzled, it is often sent or invested overseas, which is an IFF (1). This is very prevalent in countries undergoing civil unrest or countries that do not have strong governance (14). For example, for the 2 trillion USD that came into Africa as foreign direct investment or official development assistance, 1 trillion USD left the country in the form of IFFs (19). Therefore, it is likely that there is a connection between the levels of IFFs and the amount of aid that consists of the GNI. This highlights the issue of improper use of funds that foreign direct investment and economic development aid provides to lower-income countries.

https://doi.org/10.59720/24-030

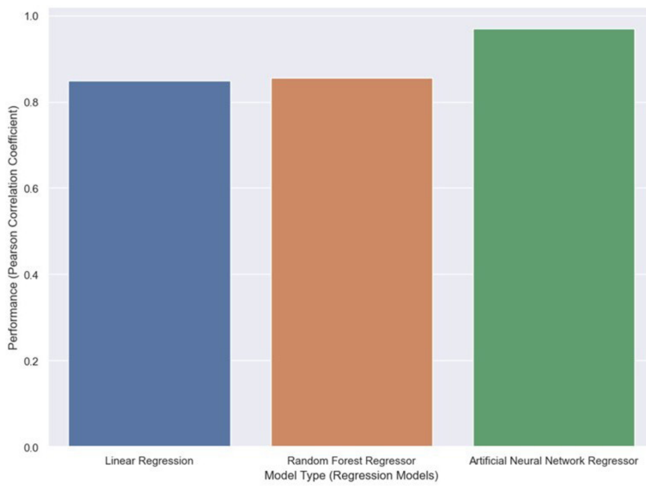

Figure 4: Bar plot of the regression model results. Comparison of the Pearson correlation coefficients of the three regression models—linear regression (blue), random forest (orange), and artificial neural network (ANN, green)—for predicting illicit financial flows (IFFs). The ANN model outperformed the others with a Pearson coefficient of 0.970, indicating its superior ability to capture complex relationships between economic and political indicators and IFFs.

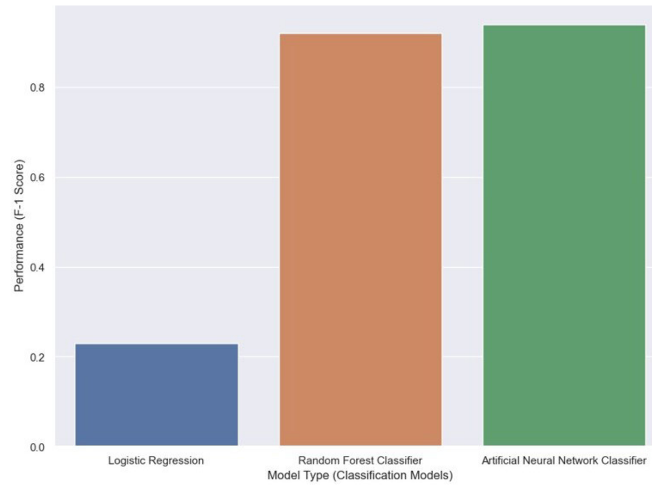

Figure 5: Bar plot of the classification model results. Weighted average F1-scores for the logistic regression (blue), random forest (orange), and artificial neural network (ANN, green) classifiers in predicting quartile-based categories of IFF levels. The ANN classifier achieved the highest F1-score (0.94), indicating its effectiveness in classifying IFF magnitudes despite data imbalance and outliers.

We found that population was the most important feature among all models (**Figure 1, 6**). However, there is no intuitive way to describe the connection between population and a high rate of IFFs. The relationship between population and IFFs might be simply because there are more opportunities to take money overseas with a higher population. In the ANN, a high population indicated high levels of IFFs, while a high gross domestic product (GDP) per capita indicated a low level of IFFs (**Figure 1**). We propose that, in the ANN and other models, the population was used as a method to adjust for the size of the economy, per country. This is because, in the developing world, larger populations usually mean a larger economy, in terms of GDP, which offers more chances to transport funds overseas as IFFs, due to more trade connections (3). The GDP per capita can indicate that most IFFs are taken out from petty corruption in developing countries, which have low GDP per capita levels.

Higher levels of government effectiveness led to lower rates of IFFs because effective government jurisdiction prevents money from being used for non-state purposes (**Figure 1, 6**) (6). Most of the IFFs happen due to improper use of government funds and the inability of governments to properly track their expenditures (6). From the effectiveness of the government effectiveness feature, we can assume that the study that formulated this feature effectively captures the levels of government control on corruption (16). In the study of IFFs, trade mis-invoicing, the practice of incorrectly invoicing trade to illegally shift capital over orders, has largely not been used to track IFFs (1). Yet, a connection can be made between government effectiveness and trade mis-invoicing, because the lack of government controls of trade can lead to trade mis-invoicing and even IFFs being noted as legitimate trade, especially by multinationals (20). Evidence can be found in the example of Côte d'Ivoire coffee exports, as Côte d'Ivoire has a low government effectiveness rating and a high level of IFFs (20).

In parallel, the rule of law variable is a predictive feature of high IFFs (**Figure 1, 6**). The rule of law has a clear

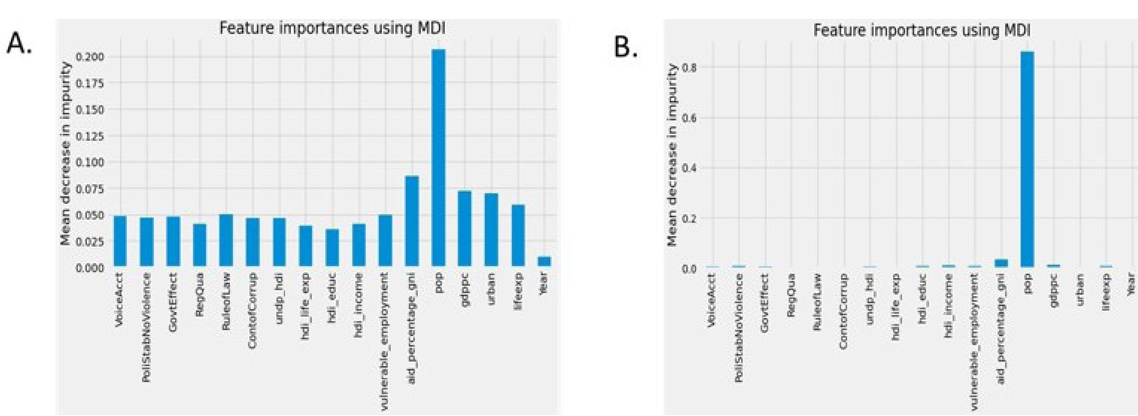

Figure 6: Feature importances for the random forest classifier and regressor. A) Feature importances using permutation on full model for the random forest classifier. **B)** Feature importances using MDI for the random forest regressor. These were extracted by obtaining the feature loadings from the models.

https://doi.org/10.59720/24-030

connection to corruption, since enforcement of existing laws is the first deterrent to IFFs. The rule of law encompasses the government's ability to enforce its rules, while government effectiveness is the ability of government institutions to pass new regulations. Less respect for the rule of law promotes more IFFs and domestic tax avoidance. This feature has a lot of overlap with the government effectiveness feature, but they each predict opposite levels of IFFs, due to government efficiency being the difference between an issue being formed and the issue being solved, because if there are longer times to address issues, more IFFs can leave the country. Rule of law is the ability for governments to enforce their laws, so low levels of rule of law lead to high levels of IFFs. The rule of law feature has a similar center of distribution and overall distribution to the government effectiveness feature (**Figure 6**). The rest of the features had a relatively minimal impact on the rate of IFFs.

We also observed progress in reducing IFFs (the 'year' feature), since as time goes on, the predicted number of IFFs decreases in the model (**Figure 1**). The time's effect on IFFs is a relatively small factor in the model, which is to be expected, since these anti-IFF efforts take quite a bit of time to start showing observable results in the model. IFFs have been identified as a major issue by the international community during the early 2000s, and this dataset encompasses the years 2009—2018 (6). Therefore, the results of this dataset may indicate that efforts such as the UNCAC and OECD Anti Bribery Convention have had some success.

Some limitations of our study include not having a complete understanding of the population's effect on the level of IFFs: population was suggested to lead to a higher level of IFFs, but we do not know whether other confounding factors are also causing this shift. Moreover, the common understanding is that trade mis-invoicing in commodity-exporting countries contributes greatly to IFFs. Considering that most of the commodity-exporting countries have small populations, the population's positive trend with IFF levels might contradict existing studies (16). Additionally, we did not extensively cover the effectiveness of the effect of global attempts to prevent IFFs, since the analysis of the 'year' column in our dataset was not a significant feature in the ML models. Studying this could entail increasing the dataset to encompass years before and after the time range for the dataset used in this study and also studying countries that have gone through wars and recovered. Additionally, IFFs in countries that have signed the UN agreement could be compared against countries that have not signed it. The effect of the OECD agreement could be understood by seeing how many IFFs came into the OECD countries before and after the agreement was signed and how that trend evolved.

Further work could improve upon our findings by understanding the features that lead to IFFs in more detail, such as the population, which was unclear in this study, although it was one of the most important features. A more detailed investigation needs to be performed on foreign aid's relationship with IFFs. Additionally, due to many IFFs being due to trade mis-invoicing, further broad metrics could be developed to assess the effect of trade mis-invoicing on IFFs from developing countries, possibly focusing on commodity exports.

The technology generated from this study could help risk assessment in various countries for businesses and other organizations and help truly show the scale of the impact of IFFs on these countries. This is the first accurate predictor of IFFs for developing countries. Additionally, using the various features that have been found to significantly impact IFFs, policy makers and government officials could effectively target and minimize IFFs.

MATERIALS AND METHODS

There are 120 countries in the dataset, with each country appearing 10 times (one column for each year), resulting in the dataset having 1200 distinct data inputs, from the years 2009 to 2018. The year has been set as a feature, to analyze if that might contribute to predicting the level of IFFs. The null data has been replaced with averages of the features for the country, so that we could run analysis and train ML models on the data. In this study, 75% of the 1190 data points are used for training ML models and 25% are used for treating the models, a ratio commonly used in machine learning regression and classification tasks (17). The data was then scaled using a Mix—Max normalizer to set all the values between 0 and 1. Both foreign direct investment and economic development aid are counted as a part of the aid percent of GNI in the present study.

To carry out the analysis, we used several machine learning models. The problem was approached by framing it as both a classification task as well as a regression task. For the classification task, the dataset was divided into four quartiles based on the IFF values, transforming the problem into a multi-class classification task. Three models were implemented: linear regression, random forest classifier, and an ANN classifier. Linear regression was primarily used as a baseline, mapping the data linearly to predict quartile membership, despite its limitations in capturing complex relationships. The random forest classifier, a tree-based ensemble model, was then applied to leverage non-linear interactions between features, given its success in similar classification problems. Finally, an ANN classifier was constructed to capture more intricate patterns within the data, utilizing multiple layers with ReLU activation functions and Dropout regularization to enhance generalization and prevent overfitting. The performance of these models was evaluated to determine their effectiveness in predicting the quartile-based classification of IFF values.

For the regression task, we first used linear regression, which is the simplest regression model. This was done as a preliminary experiment, to gauge the feasibility of the project. This model simply tries to fit the variables onto a line of best fit and predicts the values from that line of best fit. However, the model has limited ability to comprehend the features.

After this, we decided to use the random forest model to perform regression, because it had very promising results in the quartile classification. Due to having promising results, like the random forest classification, the feature importance was extracted for the results section below. The regression and classification random forest models have been optimized, however the optimization did not yield a better performance. This was used to pinpoint which statistics contributed to predicting the IFF values. After this, an ANN was developed for the regression task. It is the most complex model that was developed for predicting the IFFs. The model comprises several layers to extract and process features. We utilized the Keras API for developing the ANN (18). Beginning with an input layer consisting of 162 nodes, each employing the Rectified Linear Unit (ReLU) activation function, the model incorporates Dropout regularization to prevent overfitting by randomly deactivating 30% of input units during training. After the input layer, multiple hidden layers, each with ReLU activation and Dropout regularization (dropout rate of 0.3), are included. These hidden layers consist of one Dense layer with 81 nodes, one with 40 nodes, one with 20 nodes, and one with 10 nodes. The output layer, without an explicit activation function, outputs predicted values. Compiled with the Adam optimizer and employing mean squared error (MSE) as the loss function, the model is trained using the fit method on training data alongside validation data. With a batch size of 128, the model is trained for 1600 epochs, facilitating robust learning of underlying data patterns. We optimized for the batch size and number of epochs. This ANN architecture adeptly captures nonlinear relationships within the data while mitigating overfitting through dropout layers and leverages the ReLU activation function and Adam optimizer for efficient learning.

To visualize the ANN model, SHAP were used, because it is an effective way to understand ANNs (13). The SHAP values come from game theory and are used to assign contributions to the features in the ML model. We provide a github repository: https://github.com/akshithpt109/Predicting_IFF.

Received: January 15, 2024
Accepted: June 16, 2024
Published: August 24, 2025

REFERENCES

1. "Illicit Financial Flows." Global Financial Integrity. https://gfintegrity.org/issue/illicit-financial-flows/. Accessed 13 June 2023.
2. Dharmapala, D., and J. R. Hines Jr, "Which countries become tax havens?" *Journal of Public Economics*, vol. 93, no. 9-10, 2009, pp. 1058-1068. https://doi.org/10.1016/j.jpubeco.2009.07.005.
3. Brandt, Kasper. "Illicit Financial Flows and Developing Countries: A Review of Methods and Evidence." *Journal of Economic Surveys*, vol. 37, no. 3, 2022, pp. 789–820. https://doi.org/10.1111/joes.12518.
4. Alstadsæter, Annette, et al. "Tax Evasion and Inequality." *American Economic Review*, vol. 109, no. 6, June 2019, pp. 2073–103. https://doi.org/10.1257/aer.20172043.
5. Alm, James, et al. "When You Know Your Neighbour Pays Taxes: Information, Peer Effects and Tax Compliance." *Fiscal Studies*, vol. 38, no. 4, 2017, pp. 587–613. https://doi.org/10.1111/1475-5890.12111.
6. Olken, Benjamin A., and Rohini Pande. "Corruption in Developing Countries." *Annual Review of Economics*, vol. 4, 2012, pp. 479–509. https://doi.org/10.1146/annurev-economics-080511-110917.
7. Colonnelli, Emanuele, et al. "What Predicts Corruption?" *Social Science Research Network* , 23 Dec. 2020., https://doi.org/10.2139/ssrn.3330651.
8. Lima, Marcio Salles Melo, and Dursun Delen. "Predicting and Explaining Corruption across Countries: A Machine Learning Approach." *Government Information Quarterly*, vol. 37, no. 1, Jan. 2020, p. 101407. https://doi.org/10.1016/j.giq.2019.101407.
9. López-Iturriaga, Félix J., and Iván Pastor Sanz. "Predicting Public Corruption with Neural Networks: An Analysis of Spanish Provinces." *Social Indicators Research*, vol. 140, no. 3, Dec. 2018, pp. 975–998. https://doi.org/10.1007/s11205-017-1802-2.
10. Lépissier, Alice. "Machine Learning for Missing Data on Illicit Trade." 2023, https://alicelepissier.com/docs/illicitAI.pdf.
11. Yang, Guangyi, Liu, Xiaoxing, & Li, Beixin. " Anti-money laundering supervision by intelligent algorithm" *Computers & Security*, vol. 131, 2023, p. 103254. https://doi.org/10.1016/j.cose.2023.103344.
12. Charoenwong, Ben, and Prateek Reddy. "Using Forensic Analytics and Machine Learning to Detect Bribe Payments in Regime-Switching Environments: Evidence from the India Demonetization." *PLOS ONE*, vol. 17, no. 6, 2022, e0268965. https://doi.org/10.1371/journal.pone.0268965.
13. Kaufmann, Daniel, et al. "The Worldwide Governance Indicators." *World Bank Policy Research*, Working Paper No. 5430, Sept. 2010. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1682130.
14. "Trade-Related Illicit Financial Flows: Data by Country." Global Financial Integrity. https://gfintegrity.org/iif-data-by-country/. Accessed 13 June 2023.
15. Mungiu-Pippidi, Alina. "Transparency and Corruption: Measuring Real Transparency by a New Index." *Regulation & Governance*, vol. 17, no.4, Sept. 2022, pp. 1094-1113. https://doi.org/10.1111/rego.12502.
16. Rozemberczki, Benedek, et al. "The Shapley Value in Machine Learning." *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, July 2022, pp. 5572-5579. https://doi.org/10.24963/ijcai.2022/778.
17. In'airat, Mohammad. "Aid Allocation, Selectivity, and the Quality of Governance." *Journal of Economics Finance and Administrative Science*, vol. 19, no. 36, June 2014, pp. 63–68. https://doi.org/10.1016/j.jefas.2014.03.002.
18. Madden, Payce. "New Trends in Illicit Financial Flows from Africa." Brookings. www.brookings.edu/articles/new-trends-in-illicit-financial-flows-from-africa/. Accessed 7 September 2023.
19. Forstater, Maya. "Illicit Financial Flows, Trade Mis-invoicing, and Multinational Tax Avoidance: The Same or Different?" Center For Global Development. www.cgdev.org/publication/illicit-financial-flows-trade-mis-invoicing-and-multinational-tax-avoidance. Accessed 8 September 2023.
20. James, Gareth, et al. An Introduction to Statistical Learning: With Applications in R. Springer, 2013.
21. Chollet, François, et al. Keras. 2015, https://keras.io.

https://doi.org/10.5972/24-030

APPENDIX

Feature Overview			
Feature	Description	Numerical Range	Reference
Voice and Accountability	Perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media.	-2.173 to 1.296	10
Political Stability and Absence of Violence	Perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including politically-motivated violence and terrorism.	-2.810 to 1.427	10
Government Effectiveness	Perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.	-1.887 to 1.505	10
Regulatory Quality	Perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.	-2.244 to 1.539	10
Rule of Law	Perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.	-1.922 to 1.426	10
Control of Corruption	Perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.	-1.673 to 1.718	10
UNDP HDI	Summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and having a decent standard of living.	0 to 0.877	12
HDI Life Expectation	HDI measure of Life Expectation.	0.371 to 0.925	12

https://doi.org/10.5972/24-030

HDI Education	HDI measure of Education.	0.22 to 0.866	12
HDI Income	HDI measure of Income.	0.307 to 1	12
Vulnerable Employment	Percentage of that is not stable and does not have any formal severance policy.	0.14 to 91.05	12
Aid Percent of GNI	The percent of foreign aid that makes up a country's Gross National Income	0 to 34.02	12
Population	The population of a country.	101,484 to 1,352,642,283	12
GDP Per Capita	A country's GDP divided by its total population.	400 to 99,147	12
% of population urbanized	The percentage of population of a country that lives in cities.	0 to 100	12
Life Expectancy	The average life expectancy of a person in that country.	44.146 to 82	12
Year	The year that the data was measured. 0 represents 2009, 9 represents 2018.	0 to 9	NA

Table A1 - Feature overview table. This table summarizes the features used in our dataset, including brief descriptions, value ranges, and the source papers from which they were derived.