**Strategic Recommendations for Actionable Steps in Customs Administration**

**WCO Regional Pre-Accreditation Workshop for TOAs on Data Analytics**

Prepared by: Diyouva Christa Novith

Date: 16 January 2024

Role: Data Analyst

Contact: diyouva.christa@customs.go.id

**INTRODUCTION**

Customs administrations play a pivotal role in safeguarding trade integrity while ensuring the efficient collection of tariffs and taxes. This report leverages advanced data analytics techniques to examine a sample import dataset, aiming to uncover actionable insights for improving fraud detection and revenue optimization. The dataset includes critical fields such as transaction details, importer information, tariff codes, and binary indicators of fraudulent activity (“illicit”). Additionally, it provides insights into revenue generation from flagged cases, offering an opportunity to assess and enhance Customs operations.

Through an in-depth analysis of patterns, trends, and anomalies, this report highlights key findings and presents practical recommendations supported by visualizations. These insights aim to empower Customs administrations to streamline processes, mitigate risks, and maximize operational effectiveness.

**KEY FINDINGS**

1. **Exploratory Data Analysis**

**Boxplot of Numerical Features**

Significant outliers across numerical features, such as tariff.code, suggest possible anomalies in trade declarations.

**A graph with lines and dots

Description automatically generated**

**Correlation Heatmap**

Strong correlations between features, such as fob.value and cif.value, suggest patterns that could be used for anomaly detection.

**A screenshot of a graph

Description automatically generated**

**Skewness of Numerical Features**

**Gross Weight** displays an extremely high skewness, indicating that the data is heavily concentrated around lower values, with the presence of significant outliers or extreme values. **FOB Value, CIF Value, and Total Taxes** exhibit moderate skewness, suggesting that their distributions are somewhat asymmetric and may benefit from transformations to normalize them (e.g., log transformation). **Tariff Code** shows minimal skewness, suggesting a relatively balanced distribution. **Illicit and Revenue** have low skewness, indicating that their distributions are closer to symmetrical.

**A graph with blue bars

Description automatically generated**

**Distribution of Target Variables (**illicit **and** revenue**)**

Data imbalance is evident, with far fewer illicit trade instances (illicit = 1). Revenue distributions are heavily skewed, highlighting the presence of extreme values.

**A blue rectangular object with white text

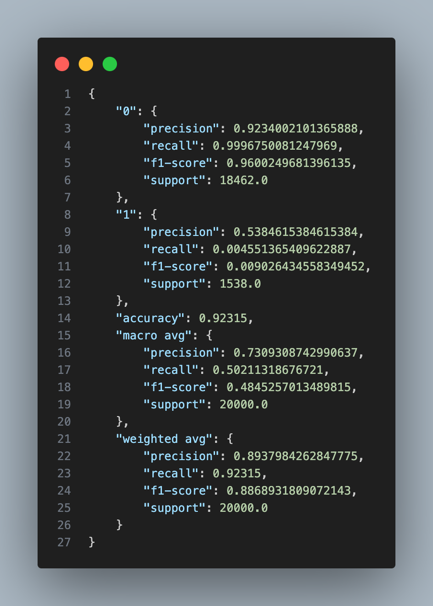
Description automatically generatedA graph with numbers and lines

Description automatically generated**

1. **Model Metrics:**

**Classification Metrics:**

The classifier for detecting illicit trade demonstrates high accuracy (92.3%) and precision for the non-illicit class (0.92). However, it struggles to recall illicit trade cases (recall = 0.004), indicating that most illicit instances are missed.



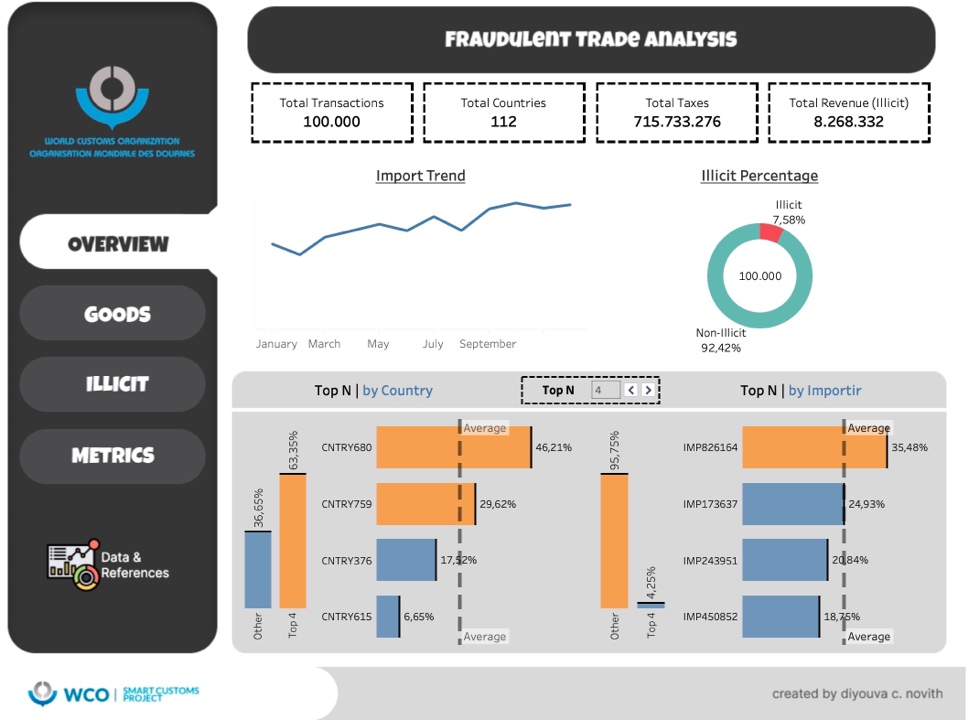
**Regression Metrics:**

The regressor for predicting potential revenue on illicit trades has a negative R² (-0.855) and a high mean squared error (22.60), demonstrating poor predictive performance.

A computer screen shot of a black and white error

Description automatically generated

1. **Descriptive Analysis**

****

**A screenshot of a computer

Description automatically generated**

1. **Conclusion**

**Invest in Training Programs:**

Equip Customs officers with skills in data analysis and interpretation. Understanding patterns in data can empower them to identify irregularities more effectively.

**Develop a Data-Driven Culture:**

Encourage a mindset that integrates data analytics into daily operations. Utilize dashboards and real-time monitoring systems for better oversight.

**Continuous Learning:**

Support officers with access to online courses, certifications, or workshops on advanced analytics and machine learning to adapt to evolving technologies.

**Improving Data Quality and Feature Engineering:**

Address **outliers** and missing values in critical fields (e.g., tariff.code, fob.value, quantity). This will improve model robustness and ensure reliable predictions.  
Enforce stricter data validation protocols at the point of data collection to reduce inconsistencies.

**Enhancing Model Performance:**

**Classification Models:**  
Use oversampling (SMOTE) or undersampling techniques to balance the dataset and improve recall for illicit trade detection.  
Experiment with alternative models like Gradient Boosting or XGBoost for enhanced classification performance.

**Regression Models:**  
Apply transformations (e.g., log transformations) to handle skewness in revenue data.  
Perform hyperparameter tuning and ensemble modeling for better accuracy.

**Automation and Real-Time Analytics:**  
Implement anomaly detection systems that can automatically flag unusual patterns in trade declarations (e.g., unusually high gross.weight or discrepancies in tariff.code).  
Leverage real-time analytics tools to monitor trade flows and detect risks promptly.

**Efficiency and Effectiveness Gains Through Analytics:**

By integrating data-driven approaches, Customs administrations can increase their operational efficiency by:

* Reducing manual inspections with automated detection systems.
* Prioritizing high-risk shipments for inspection, saving time and resources.
* Generating insights to adjust strategies dynamically.
* Effectiveness can be enhanced through better resource allocation, focusing efforts where illicit trade is most likely.

**Policy and Resource Allocation:**  
Use analytics to guide decisions on resource deployment (e.g., stationing officers in high-risk regions).  
Develop policies that emphasize the integration of advanced technologies in trade monitoring.

**Regular Model Updates and Re-Evaluation:**  
Continuously monitor model performance and retrain using updated data to adapt to changing trade dynamics.

By implementing these steps, Customs administrations can strengthen their capabilities in detecting illicit trade and recovering potential revenue losses effectively. Integrating data analytics into operations promises not only improved accuracy but also greater efficiency and adaptability in tackling trade-related challenges.