Manifest Monitoring Model as Support for Customs Risk Management: Evidence from Taiwan

Yen-Hui Kuo and Shu-Ching Chou

Abstract

This study constructs a feature prediction model in border smuggling inspection using the Artificial Neural Networks technique. We extracted six input variables relevant to smuggling through an iterated learning process on 176,869 import manifests. By adopting logistic regression, we also revealed that the six identified factors are significantly related to the occurrence of smuggling. We further used the six seizure factors to predict smuggling in 2019 and obtained an average accuracy rate of nearly 83 per cent. The accuracy rate was much higher around Fridays and holidays. The results will help provide cost-effective screening during customs clearance inspection and effectively manage border risks.

1. Introduction

Efficient enforcement of border inspection has become a necessary response to the increasingly diverse modes of international trade and flows of travellers, goods and money. As global economic interactions intensify, smugglers have increasingly utilised cross-national information and modern technologies to smuggle goods and money. Traditional manual screening of manifests, which depends heavily on the inspectors' experience, might not detect illicit activities of importers during the brief inspection process. Developing an efficient prediction model to curtail smuggling in a timely manner is crucial in border security control.

Traditional border inspection uses arduous and time-consuming processes that involve routine investigation and strategic inspections on specific targets such as seasonal routine adjustments or policy requests. Facing increasing international trading and complex smuggling behaviour, Customs therefore could consider technology-based inspection to promote trading security at borders. With rapid advances in artificial intelligence (AI) technology, AI-based prediction models have been widely applied in different fields given their considerable potential in logic deduction and massive data processing. A technology-based border inspection and administrative approach is expected to trigger an 'alarm' to potentially high-risk consignments, thus enhancing inspection accuracy and curbing smuggling.

This study utilises artificial neural networks (ANNs) optimisation methodology to build a smuggling prediction model. Using 16 input variables and training these variables with 176,869 import manifests at Taichung Port in 2018, the ANNs identified the following six input variables as seizure factors: firm age, firm capital, firm risk level, container ship deadweight, cargo volume and loading port. We used the six identified factors as independent variables to predict the occurrence of smuggling cases using logistic regression. The results showed that all six factors are significantly related to smuggling. We further used the six seizure factors on manifests at Taichung Port in 2019 to predict the smuggling

cases and obtained an average accuracy rate of nearly 83 per cent. An interesting finding was that the accuracy rate of prediction was substantially higher around Fridays and holidays, days which often have more declared cargo and usually have a higher probability of smuggling.

The existing literature discusses general policies regarding cross-border crime and border security management (Tertereanu et al., 2020; Tjiptabudi et al., 2018). The contribution of the study described here is the construction of a smuggling-detection model using real manifest data through the ANNs technique. As reported by Tan et al. (2016), the amount of noise in high-dimensional datasets makes it difficult to understand the information therein and to construct prediction models through data mining. Selecting a subset of relevant factors could reduce the dimensionality and help construct the prediction model (Zhou, 2019). The results show that ANNs could be applicable in screening large amounts of manifest information and determining relevant seizure factors on smuggling. The smugglers in modern society tend to adopt flexible smuggling methods. Our study confirms that a real-time AI-based border inspection system provides Customs with timely assistance on border risk management.

2. Literature review

2.1 Decision-making in global border risk management

Governments worldwide have adopted various border control strategies and implemented strict investigations on goods and people for clearance. Recent practice indicates that the patterns of traditional criminal activities such as drug smuggling and trafficking in migrants have gradually changed. For example, the possibility of international terrorist organisations infiltrating multinational enterprises and financial institutions through trade fraud and money laundering has greatly increased with increasing seemingly legitimate activities (Tessa, 2020). In 2020, the spread of the coronavirus COVID-19 restricted the area of activities and profitability of international criminal groups. To consolidate their power and interests, the criminal groups might increase their activities through illegal but seemingly legal transactions. Experts on Italian mafiahave urged governments to act quickly to fight the rising activities of these groups during the COVID-19 crisis (Savio, 2020).

In particular, in cooperation with high-performance cross-border logistics, border management authorities must enhance the inspection and monitoring of the data pertaining to cargo flow such as using big data analysis to establish regional monitoring models (Lee et al., 2009).When customs clearance data are mined, the data must be highly uniform and strictly adhere to international law and agreements (Garrie & Gelb, 2012). Therefore, the conversion of trade information into applicable data is necessary for administrative decision-making, which allows for economically related evaluations to improve current deployment strategies (Burinskienė & Burinskas, 2011). For example, administrative and legislative agencies must establish measures to verify the authenticity of online user accounts involved in e-commerce trade (such as using two-step authentication) and continually ensure that users of the data adhere to and are protected by regulations and the law (Raymond, 2011).

2.2 Administrative measures to reinforce customs border enforcement in Taiwan

Globalisation and the rapid evolution of information technology have resulted in goods, services and technology flowing quickly and freely across borders, which has, in turn, made customs monitoring more difficult. The duties and concurrent matters of Customs have become increasingly complicated in Taiwan over the years. On the one hand, multinational companies and participants in the global supply chain desperately require the government to establish a new operating model in border management in addition to requiring Customs to simplify procedures and speed up the release of goods from customs clearance procedures. On the other hand, the Executive Yuan of Taiwan requires the Ministry of Legal

Affairs, the Ministry of Finance, and the Financial Supervision and Administration Departments to fight against speculative international finance companies without hindering the free flow of international trade funds, aiming to establish a fair international business environment.

In response to the ever-changing economic tendencies and diverse modes of international transactions, apart from the changing administrative systems of customs operators, Customs has actively optimised technology-based enforcement devices, strengthened border inspection capabilities, and applied the Internet of Things (IOT) technology to monitoring to enhance cargo transportation safety and administration plans, which make customs management more innovative and customs clearance more efficient (Ministry of Finance, 2020).

3. Methodology

3.1 Sample used in this study

This study uses the import manifest data of Taichung Port in Taiwan for the period 1 January 2018 to 31 December 2018. Taichung Port is located in central Taiwan, and is administrated by Taichung Customs, which is governed by the Customs Administration of the Ministry of Finance. Being the largest port for international trading in the central area, Taichung Port receives data on cargo manifests of 500–600 consignments daily and 1200–1500 consignments on weekends and holidays. We collected the information of 176,869 import manifests through the PORTAL system of the Ministry of Finance. The PORTAL system contains information on cargo content, shipping terms, the importer's background, customs clearance status and smuggling-detection record, and is utilised as an important databank by the customs administration. Because smugglers often deliberately conceal their activities and related information, this large and real data sample is suitable for both manifest screening and the extraction of factors significantly related to smuggling (Khunkitti & Chongsujjatham, 2019).

3.2 Artificial neural networks (ANNs)

ANNs are machine learning techniques and have been frequently utilised recently to build prediction models. ANNs not only calculate and recognise patterns of data but also learn and generate optimised estimation through cognitive functions (Chen et al., 2009). Because ANNs develop models in a manner like the human nervous system with deep learning and the ability to adapt, ANNs are well suited for an environment with free constraints in which the data investigated does not require the assumption of a normal distribution, interactive factors and nonlinear relations (Warner & Misra, 1996). Due to these advantages, many studies have suggested that ANNs outperform conventional statistical or other classification methods in classification and prediction.

For example, Wray et al. (1994) compared ANNs to regression with respect to predictive performance (specifically, accuracy), and found ANNs to perform better. Tam and Kiang (1992) compared several prediction methods, namely ANNs, discriminated analysis, k-nearest neighbour algorithms and decision tree algorithms, with respect to how well they predicted a company's bankruptcy, and ANNs were found to be the best. Dutta and Shekhar (1988) applied ANNs to bond evaluation, where their ANNs outperformed multiple regression models in predictions. Refenes et al. (1997) suggest that ANNs are suitable for time-series prediction because ANNs can process interactive data and evolve with trends through the input of various types of data. In addition, ANNs have been found to be suitable for determining the best combinations of input variables that affect output variables, as mentioned by Niazian et al. (2018).

This study adopts the ANNs technique to construct a smuggling-detection model. We utilise the simple feed forward neural network model, with one input layer, three hidden layers and one output layer (multilayer perception, MLP). The data are partitioned as 60 per cent for training, 20 per cent for validation and the remaining 20 per cent for testing. There are 13 input variables used to train the algorithms iteratively. The cargo manifests contained these 13 input elements: firm age, loading port, firm log capital, firm risk level, transportation method, cargo volume, container ship deadweight, tariff number, container transportation time, container slot, mode of container transport services and manifest submission time. The output variable is the seizure case record. The training and deep learning procedure by ANNs helps to screen for the major factors that relate to smuggling at the Port of Taichung.

3.3 Binary selection analysis

After the screening process and identification of influential factors by ANNs, which strongly relate to smuggling, we then used a logistic regression model to inspect the relation between the identified input variables and the output variable (seizure record of smuggling). Logistic models have been commonly used for prediction, particularly for the binary choice as the dependent variable. This study adopts a binary selection model, as follows:

$$\text{Smuggling} = \alpha_0 + \sum_{i=1}^{i=n} X_i + \varepsilon \qquad (1)$$

In model 1, α_0 is the intercept. *Smuggling* is the binary dependent variable. It is coded as 1 when firms have a smuggling record; otherwise it is coded as zero. The independent variables, *X*, are the influential factors identified by the ANNs described above. The number of identified input variables ($i = 1 - n$) is determined based on the results of the ANNs, with a maximum of 13. ε is the residual of this regression.

4. Results and discussion

4.1 ANNs results

The ANNs iterated learning process extracted six input variables which feature smuggling at Taichung Port. The six input variables are firm age, firm capital, firm risk level, container ship deadweight, cargo volume and loading port. The prediction performance in the 35,373 tested manifests (20 per cent of the total sample) reached an accuracy of nearly 70 per cent. Table 1 lists the results of feature screening on test manifests by ANNs.

Table 1: ANNs results

Tested sample size: 35,373 manifests Output variable: seizure case record Selected input variables:	
Firm age	Container ship deadweight
Firm capital	Cargo volume
Firm risk level	Loading port

To understand how they relate to smuggling, descriptions of the six input variables are listed below:

(1)   Firm age: The age of the importing business is calculated based on the tax registration database of the Finance Centre of the Ministry of Finance and the company registration database of the Department of Commerce of the Ministry of Economic Affairs, starting from the time the importer is registered.

(2)   Firm capital: The amount of capital of the importer is based on the company registration database of the Department of Commerce of the Ministry of Economic Affairs. Larger capital presents larger firm size.

(3)   Firm risk level: Customs evaluate the importer's risk based on (a) their record of violation of regulations; or (b) the customs clearance records of the importer; or (c) governmental policy requirements; or (d) seasonal or other adjustments for specific inspections. Based on these evaluations, importers bearing higher risk will be evaluated with a higher firm risk level.

(4)   Container ship deadweight: The container ship deadweight is the container weight declared by the importer based on the loading plan for container ships (Bay Plan).

(5)   Cargo volume: The cargo volume is based on the declaration by the importer, calculated by using cargo length multiplied by width and height.

(6)   Loading port: Based on investigation experience, Taichung Customs has listed some export ports and trans-shipment ports as high-risk routes for smuggling.[1]

4.2 Binary selection regression analysis

We use these six input variables as independent variables to predict smuggling at Taichung Port using logistic regression with the 2018 manifests. We took the log value of capital and created a dummy variable equal to one for the loading port being recorded as the target for reinforced investigation. As shown in Table 2, all six independent variables are significantly related with seizure cases. The negative coefficients of *firm age* and *firm capital* show that older and larger firms are less likely to engage in smuggling.

Table 2: Logistic regression (LR) results

	Coefficient	**z-Statistic**	**Prob.**
Firm age	-0.027***	-4.148	0.000
Firm capital	-0.045***	-5.400	0.000
Firm risk level	0.046***	4.207	0.000
Container ship deadweight	0.005***	3.284	0.001
Cargo volume	0.010***	3.598	0.000
Loading port	0.012**	2.474	0.013
Intercept	0.356***	4.226	0.000
LR statistic = 145.019 Prob. (LR statistic) = 0.000			
*, ** and *** denote significance at 0.1, 0.05 and 0.01 level or better, respectively.			

In contrast, the coefficients of *firm risk level*, *container ship deadweight*, *cargo volume* and *loading port* are all significantly positive. The positive coefficient of *firm risk level* indicates that an importer's risk record is valuable in port investigation. For shipments, larger and heavier containers are more likely to be involved in smuggling. This may be because prohibited items are more likely to be carried in containers with more cargo. In addition, one of the routine investigations of Customs is to identify tax evasion where importers declare less cargo than is actually shifted. Therefore, larger containers are more likely to be linked to tax-evasion cases. Finally, the regression results reveal that some loading ports and trans-shipment ports have higher risks relating to smuggling (see Note 1).

4.3 Additional tests and findings − Friday events

To verify the accuracy of our smuggling prediction model, we further utilised the six input variables to conduct testing on 2019 smuggling events at Taichung Port. Figure 1 shows the average prediction accuracy rate was nearly 83 per cent. Furthermore, we found that there were seven dates with an accuracy rate close to 99 per cent: January 18 (Friday), March 6 (Wednesday), April 26 (Friday), June 21 (Friday), August 19 (Monday), October 11 (Friday) and December 6 (Friday). For example, on 18 January 2019 (see Figure 1), Taichung Customs seized a total of 962 packs of ketamine, classified as the third category of illegal narcotic, which were hidden in a batch of curtains rods imported from South-East Asia. The gross weight was 305.118 kg. In another case on 19 August 2019 (see Figure 1), 79 packs of ketamine were found in stainless steel moulds, imported from the Middle East. These empirical cases of seizure correspond to predictions by our ANNs and logistic model.

Figure 1: Screening of import manifests of the Port of Taichung in 2019
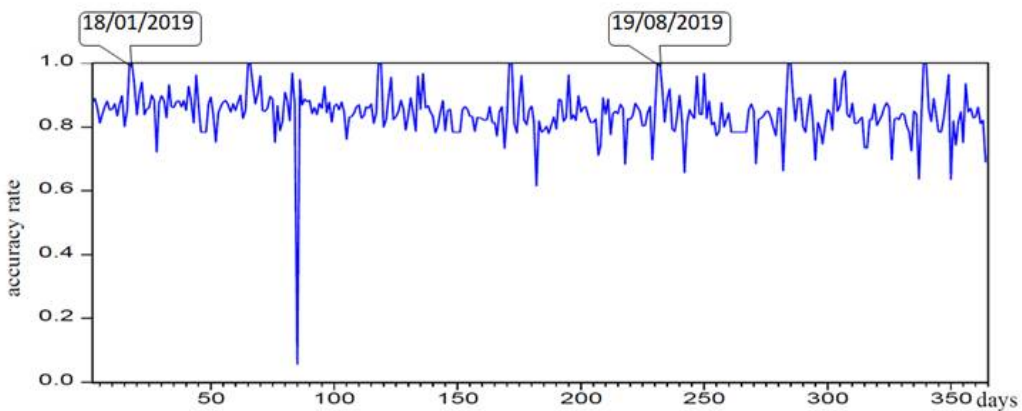
Notably, among the seven dates which had high instances of smuggling, five dates were Fridays and one was a Monday. Some of these dates were close to consecutive holidays, such as January 18 (close to Lunar new year), October 11 (close to Double Tenth Day, a Taiwanese national holiday,) and December 6 (close to Christmas and New Year holidays). The results show that, in addition to the overall accuracy rate of approximately 83 per cent, the six input variables are suitable to help identify smuggling around weekends and holidays when there is usually a higher rate of smuggling.

5. Conclusion and suggestions

This study utilises ANNs techniques and manifest data from Taichung Port to build a smuggling prediction model. The iterated data training process determined six input variables as relevant to smuggling seizure records: firm age, firm capital, firm risk level, container ship deadweight, cargo volume and loading port. The logistic regression further shows that larger, older import firms are less likely to smuggle, while importers with a higher rated risk, with higher weight and volume cargo, and consignments shipped from specific ports more likely to involved in smuggling. We used the six seizure factors to predict the occurrence of smuggling at Taichung Port in 2019. We found that the prediction accuracy nears 83 per cent and is even higher for manifest declared around Fridays and holidays.

Customs today faces multiple challenges requiring a great deal of adaptability. This is visible in all modes of transport. Offenders attempt to abuse vulnerable supply chains to smuggle numerous commodities and money, which pose threats to the safety and security of citizens (World Customs Organization [WCO], 2018). The findings presented herein will appeal to customs enforcers through providing critical factors in predicting smuggling among large quantities of manifest information. For future research and practice, depending on the region and time, and even changes in the method of smuggling, the prediction techniques and the six input variables may need adjustments before application.

Disclaimer

The findings and the views in this paper are those of the authors and do not necessarily present the views and policies of Taichung Customs or the WCO.

References

Burinskienė, A., & Burinskas, A. (2011). The perception of market and e-commerce. *European Integration Studies*, *5*, 179−85.

Chen, W., Lu, H., Wang, M., & Fang, C. (2009, November 7−8). Gene expression data classification using artificial neural network ensembles based on samples filtering. In *2009 International Conference on Artificial Intelligence and Computational Intelligence* (pp. 626−628). IEEE. DOI: 10.1109/AICI.2009.441.

Dutta, S., & Shekhar, S. (1988, December). Bond rating: a non-conservative application of neural networks. In *IEEE Int Conf on Neural Networks* (pp. 443−450). IEEE.

Garrie, D. B., & Gelb, D. K. (2012). An argument for uniform E-discovery practice in cross-border civil litigation. *Journal of Business & Technology Law*, *7*(2), 341−359.

Khunkitti, A., & Chongsujjatham, P. (2019, November 2−4). A Rule-Based Training for Artificial Neural Network Packet Filtering Firewall. In *2019 6th International Conference on Systems and Informatics (ICSAI)* (pp. 1010−1014). IEEE.

Lee, J., Kang, S., Lee, Y., Lee, S. J., & Song, J. (2009). BMQ-Processor: a high-performance border-crossing event detection framework for large-scale monitoring applications. *IEEE Transactions on Knowledge and Data Engineering*, *21*(2), 234−252.

Ministry of Finance. (2020). 109 Annual Policy Plan. Republic of China Ministry of Finance (in Chinese).

Niazian, M., Sadat-Noori, S. A., Abdipour, M., Tohidfar, M., & Mortazavian, S. M. M. (2018). Image processing and artificial neural network-based models to measure and predict physical properties of embryogenic callus and number of somatic embryos in ajowan (Trachyspermum ammi (L.) Sprague). *In Vitro Cellular & Developmental Biology-Plant*, *54*(1), 54−68.

Raymond, A. (2011). Improving Confidence in Cross Border Electronic Commerce: Communication, Signatures and Authentication Devices. *Journal of Internet Law*, *14*(7), 25−34.

Refenes, A. P., Burgess, A. N., & Bentz, Y. (1997). Neural networks in financial engineering: A study in methodology. *IEEE transactions on Neural networks*, *8*(6), 1222−1267.

Savio, I. (2020, April 13). Italia y la hoja de ruta del crimen: cómo las mafias del mundosacarántajada del COVID [Italy and the crime roadmap: how the world's mafias will get their share of Covid]. *El Confidencial*. https://www.elconfidencial.com/mundo/europa/2020-04-13/mafia-italiana-crisis-coronavirus_2543271/

Tam, K. Y., & Kiang, M. Y. (1992). Managerial applications of neural networks: the case of bank failure predictions. *Management science*, *38*(7), 926−947.

Tan, P. N., Steinbach, M., & Kumar, V. (2016). *Introduction to data mining*. Pearson Education India.

Tertereanu, P., Gîngu, A. I., & Țîțu, A. M. (2020). Aspects regarding the possibilities of implementing integrated border management as a global strategy promoted by the European Union. *Review of General Management*, *32*(2), 29–42.

Tessa, M. (2020, May 20). Imprese: sale pericolo di infiltrazioni criminali, servono nuovi indicatori di anomalia [Businesses: danger of criminal infiltration rises, new anomaly indicators are needed]. *Wall Street Italia (WSI)*. https://www.wallstreetitalia.com/imprese-sale-pericolo-di-infiltrazioni-criminali-servono-nuovi-indicatori-di-anomalia/

Tjiptabudi, M. F., Igon, S. S., Bernardino, R., & Muharram, A. T. (2018). Secure and effective reengineering information system and business processes of cross-border control between the Republic of Indonesia and the Republic Democratic of Timor-Leste. In *2018 6th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1−7). IEEE.

Warner, B., & Misra, M. (1996). Understanding neural networks as statistical tools. *The American Statistician*, *50*(4), 284−293.

World Customs Organization (WCO). (2018, June 6−8). WCO IT conference and exhibition: Building a reliable digital landscape to boost cross-border trade. *WCO IT Conference*. http://www.wcoomd.org/en/events/event-history/2018/2018-wco-it-conference-and-exhibition.aspx.

Wray, B., Palmer, A., & Bejou, D. (1994). Using neural network analysis to evaluate buyer-seller relationships. *European journal of Marketing*, *28*(10), 32−48.

Zhou, X. (2019). Data mining in customs risk detection with cost-sensitive classification. *World Customs Journal*, *13*(2), 115−130.

Notes

1   Based on the confidentiality of customs investigations, variable descriptions are expressed in general terms. For example, the standards for the classification on import's risk level and the name of loading ports are not suitable for detailed descriptions.

Shu-Ching Chou

Professor Shu-Ching Chou teaches in the Department of Finance at National Yunlin University of Science and Technology (NYUST), Taiwan. She completed her MBA degree at Pennsylvania State University and obtained a PhD in Business Administration from National Sun Yat-sen University, Taiwan. Her research topics include corporate governance, corporate finance, financial decision making and government policy. She has been the Dean of the Doctoral Program in Industrial Management at NYUST, Dean of the Finance Department at NYUST and senior supervisor auditor of KPMG, Taiwan. Professor Chou is also a Certified Public Accountant (US).

Yen-Hui Kuo

Yen-Hui Kuo works for Taichung Customs with experience in customs operations, including valuation, cargo examination, bonding and seizure. In 2016, Yen-Hui Kuo represented Taichung Customs in a risk management taskforce organised by the Customs Administration. In 2019, he was transferred to Taichung International Airport to participate in analysing advance passenger information. He completed his master in finance at National Yunlin University of Science and Technology (NYUST), Taiwan, in 2007. Yen-Hui Kuo is a PhD candidate in finance at NYUST.