

**FACTORS AFFECTING THE CUSTOMS CLEARANCE TIME AT PRIME
CUSTOMS OFFICE TYPE A OF TANJUNG PRIOK**

**Khamami Herusantoso¹
Ardyanto Dwi Saputra²**

¹ General Finance Education and Training Center, Ministry of Finance, Indonesia

² Polytechnic of State Finance STAN, South Tangerang, Indonesia

Email: *khamami.herusantoso@kemenkeu.go.id*

ARTICLE INFORMATION

Received date

[15-09-2020]

Revision

[16-11-2020]

Accepted date

[22-12-2020]

ABSTRACT

In the dwell-time, the customs clearance is considered as the most complex phase, even though its portion is the shortest among other phases, such as pre-clearance and post clearance. In order to improve the efficiency and effectiveness on the services performed in the customs clearance process, the customs authorities must start considering the help of database analysis in identifying obstacles instead of depending on the personal analysis. Useful information is hidden among the importation data set and it is extractable through data mining techniques. This study explores the customs clearance process of import cargo whose document is declared through the red channel at Prime Customs Office Type A of Tanjung Priok (PCO Tanjung Priok), and applies a specific data mining classifier called the decision tree with J48 algorithm to evaluate the process. There are 11 classification models developed using unpruned, online pruning, and post-pruning features. One best model is chosen to extract the hidden knowledge that describes factors affecting the customs clearance process and allows the customs authorities to improve their services performed in the future.

Keywords: *customs clearance, data mining, decision tree, J48 algorithm, red channel.*

INTRODUCTION

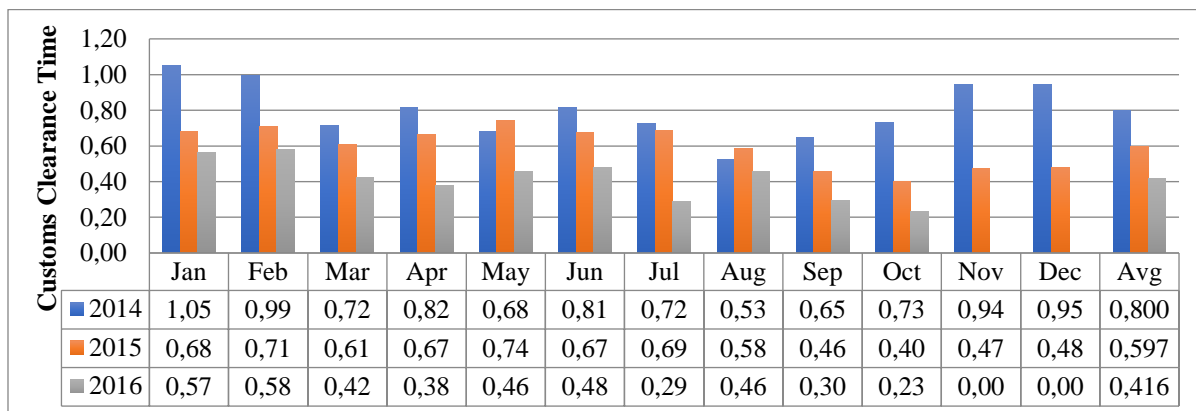
One of the components analyzed by the World Bank in measuring the international logistics performance index (LPI) is the efficiency of customs and border management performance. Based on the latest LPI in 2018, Indonesia has climbed 17 places from the previous position at 63 with a score of 2.98 to 46th position with a score of 3.15. Despite the increase, Indonesia is still behind some ASEAN countries namely Malaysia, Vietnam, Thailand and Singapore. These countries are ranked 41, 39, 32 and 7 respectively.

For many years, the World Bank has urged the Indonesian government to draft a clearer and more comprehensive strategy to improve logistics performance as well as to reduce the dwell-time. As the response, the government including the Directorate General of Customs Excise (DGCE) and other agencies related to this matter, have established many bonded logistics centers across the country and has introduced the Indonesia Single Risk Management program which will center the overall services in export and import under one rood with a single identity. In fact, these actions are still not enough to overcome the obstacles occurred regarding both the inefficiency and the dwell-time.

Dwell-time in international trade indicates the actual time a consignment stays at the port of entry, commencing from the time the transport has discharged the cargo until it exits from the port premises, having completed all relevant formalities (USAID Trade Project, 2014). Dwell-time consists of three stages viz., pre-clearance, customs clearance, and post-clearance. The length of dwell-time depends on how fast both documents and the imported goods are being handled in those stages. The direct role of DGCE, in this case, is to accelerate the process of customs clearance as efficiently and effectively as possible, in turn reducing the dwell-time as well.

As one of the biggest ports in Indonesia, Tanjung Priok port is considered the main gate of the imported goods coming into the country. In 2015, Prime Customs Office Type A of Tanjung Priok (PCO Tanjung Priok) has served around 1.4 millions twenty foot equivalent units (TEU) out of 2.4 millions TEU total. It means approximately 58% of total imported goods are brought into the country through Tanjung Priok port. The dwell-time at PCO Tanjung Priok in 2015 had been targeted at 4.7 days, consisting of 2.7 days (57%) for pre-clearance, 0.5 days (11%) for customs clearance, and 1.5 days (32%) for post-clearance. The actual average customs clearance time, however, was 0.6 days . Although it did not meet the target, the 2015 customs clearance time has been improved quite well comparing to the 2014's. Details of monthly customs clearance time are shown in Figure 1.

Figure 1. Diagram of Monthly Customs Clearance Time at PCO Tanjung Priok in 2014 to 2016



Source: Developed from the performance report of PCO Tanjung Priok

Customs clearance time, particularly for imported goods, starts from when the importer or the customs broker (PPJK) submits the import declaration (PIB) into customs in-house system (CEISA) and end right after a customs clearance approval certificate (SPPB) has been issued. There are practically five categories of the channel in importing goods, namely red, yellow, green, customs main partner (MITA) priority, and MITA non-priority channel. Every channel is treated differently based on the risk associated with both the importer and the goods imported. Among those channels, only the red and yellow ones affect the customs clearance significantly. In both channels, SPPB is issued after the document verification. A physical inspection is conducted as an extra treatment for PIB declared in the red channel only.

Customs clearance process at PCO Tanjung Priok is usually the fastest phase in the dwell-time which is on average approximately 10 to 15% of the total time. When document verification is conducted before issuing SPPB, the customs clearance time becomes 32% to 35%. But when a physical inspection is carried out as an extra treatment, it escalates to 40% to 50% and becomes the longest period compared to both pre-clearance and post-clearance. Detailed monthly dwell-time and its comparison in every phase are shown in Table 1. This condition compels PCO Tanjung Priok to continue improving their services and supervisions to be as efficient and effective as possible, especially in the customs clearance process for documents declared through the red channel.

In order to improve the efficiency of services performed, PCO Tanjung Priok has established the service level agreement (SLA). Unfortunately, it does not cover specifically the completion of PIB declared in the red channel. However, SLAs in PCO Tanjung Priok only state that the customs value and tariff determination as a part of the PIB completion shall be delivered within three workdays. Due to lack of SLAs established in PIB completion, the

compliance counseling and information service section (BKLI) at PCO Tanjung Priok facilitates the service user who does not receive any response within three days. In 2015, at least 1,500 complaints were processed in regard to documents confirmation to the customs specialist (PFPD). It shows that many service users do not receive the services promised regarding the determination of customs value and tariff.

In order to enhance services and supervisions, PCO Tanjung Priok has identified these following problems related to the customs clearance process: the lack of synergy among stakeholders in the port, handing over the hardcopy of PIB for the red and yellow channels still takes more than one day, and the waiting time before the physical inspection can be conducted is too long. Although the customs clearance process mostly relies on the competence of customs officers, it cannot be separated from stakeholders' role as well. Some problems identified depend entirely on the stakeholders, such as in the process of handing over the hardcopy of PIB and preparing the container to be ready for physical inspection.

Table 1. Time Comparison between Every Phase in the 2016 Dwell-time

2016	In General			Yellow Channel			Red Channel		
	Pre	CC	Post	Pre	CC	Post	Pre	CC	Post
January	3.57	0.70	1.47	4.35	2.56	1.26	4.98	4.62	1.14
February	2.72	0.69	1.44	3.52	2.47	1.14	4.56	4.58	1.12
March	2.00	0.54	1.42	3.07	2.33	1.09	2.99	4.11	1.08
April	1.99	0.51	1.41	3.18	2.12	1.09	3.05	3.91	1.07
Mei	2.45	0.58	1.46	3.49	2.57	1.17	3.00	4.65	1.24
June	2.30	0.57	1.47	3.11	2.17	1.10	3.82	4.52	1.13
July	2.31	0.37	1.47	3.38	1.92	1.13	3.59	3.79	1.27
August	2.67	0.59	1.32	3.79	2.34	1.05	3.97	4.28	1.01
September	2.39	0.37	1.44	3.16	2.25	1.06	3.63	3.73	1.00
October	1.95	0.29	1.26	2.81	2.04	0.98	3.33	3.73	0.93
AVG (days)	2.44	0.52	1.42	3.39	2.28	1.11	3.69	4.19	1.10
%	55.70	11.92	32.37	50.03	33.63	16.35	41.09	46.67	12.23

Source: Developed from monthly dwell-time report collected from the technical support section

Regarding those obstacles, PCO Tanjung Priok has fully implemented Director General of Customs and Excise Decree number PER-16/BC/2016 concerning the guidance to release imported goods for local consumption starting from July 1st, 2016. It has simplified the customs clearance procedures and also reduced the time as much as possible, yet still allows for stakeholders to participate during the process. Even so, the same problems still continually persist because the identification risk process is mainly based on personal analysis of the customs staffs. Although the process has been proven useful handling the obstacles identified, it may be improved by considering the help of database analysis. Thus, PCO

Tanjung Priok shall start considering the utilization of importation database in analyzing the whole process of customs clearance. Applying data mining technique to the importation database can be one of the alternatives that provides some useful information for PCO Tanjung Priok to identify factors affecting the customs clearance process thoroughly.

LITERATURE REVIEW

Artakusuma (2012) studied the import container dwell-time in the Jakarta International Container Terminal (JICT). The conclusion of the study is that containers in the red and yellow channels have the longest dwell-time compared with other channels. Moreover, the customs clearance process as a part of the dwell-time depends mostly on whether the physical inspection is conducted or not.

Han, Kamber, & Pei (2006) as cited in Wang & Song (2011) mentioned that Data mining could be understood as mining implicit, unknown, potential, and useful information and knowledge from mass data including the structured and unstructured. The information extracted from data mining analysis might be potentially valuable to support decision-making for enterprises. Wang & Song applied data mining technology to analyze the efficiency of customs clearance as many studies in customs clearance time always focus on the analysis and comparison of a business process using the qualitative method, instead of using large amount of business data and analyze it with the help of data mining techniques. They then concluded that data mining method could be used for risk revaluation on customs clearance business to improve the customs inspection and monitoring level.

Another study conducted by Geourjon, Laporte, Coundoul, & Gadiaga (2012) showed the use of data mining for risk management by Customs Administration in order to inspect less with better output. Chermiti (2019) and Zhou (2019) has started looking into the potential of machine learning and data mining methods for fraud detection and risk profiling. In 2015, National Academy of Customs Excise and Narcotics (NACEN) published a working paper entitled Data Warehousing and Data Mining for improvement of Customs Administration in India and concluded that customs will be able to do research using large amount of data that are generated. Specifically, it could be used to predict, to anticipate frauds, and to respond in advance so that threats to compliance were effectively mitigated. Therefore, data mining could be a win-win situation for customs administration and trade by requiring less inspection yet providing effective supervision.

RESEARCH METHODS

The object of this study is the importation database extracted from Customs and Excise Information System and Automation (CEISA). The database used is considered as the secondary data and limited to the import declaration documents (PIB) in the red channel only and in the period of January to October 2016. PIB whose inspection is through x-ray container inspection system (Hi-Co Scan) is excluded due to limited timestamps received from the data collection process.

In order to succeed, the researcher conducts informal interviews with customs specialist (PFPD), customs inspectors, and other customs officers who perform services related to the customs clearance process. The interview is carried out to gather relevant information in business understanding as well as to analyze the output. This research uses the Cross Industry Standard Process for Data Mining (CRISP-DM) reference model as it has been suggested by Provost & Fawcett (2013) as one good reference model for the implementation of big data. CRISP-DM consists of six phases, namely business understanding, data understanding, data preparation, modeling, evaluation, and deployment.

ANALYSIS AND DISCUSSION

In order to arrange this study systematically, the researcher chooses to use CRISP-DM reference model that consists of six phases, as follows:

1. Business understanding

Shearer et al. (2000) stated that business understanding is perhaps the most important phase of any data-mining project. It involves several key steps including determining business objectives, assessing the situation, determining the data mining goals, and producing the project plan. These steps are implemented in this study by doing preliminary research related to the customs clearance process at PCO Tanjung Priok. Specifically, the researcher conducts these following activities in this phase: literature reviews, observing the process of custom clearance by interviewing the users and overviewing the performance report of PCO Tanjung Priok. Furthermore, the performance report of several sections related to the process, such as the compliance counseling and information service section (BKLI), the service on customs and excise section (PPC), and the customs specialist (PFPD) is very helpful in understanding the customs clearance process at PCO Tanjung Priok. The data mining goal that the researcher wishes to accomplish is to identify factors affecting the customs clearance time for imported goods whose PIB is declared through the red channel at PCO Tanjung Priok. It may be

achieved by analyzing the importation database using Weka 3.6.14 version as the software analysis and decision tree J48 classifier as the chosen algorithm.

2. Data understanding

In this study, the data is extracted from the Customs-Excise Information System Automation (CEISA) database. Some additional data related to the users, such as the customs specialist and the customs inspector-specific information, are collected from the employee database at PCO Tanjung Priok. The extracted data covers three types of perspective, such as the procedure (activities' timestamp), the user (the identity of the user in every activity), and the object (in this case, a few general data declared through PIB). During the verification, the researcher finds many empty fields in several columns, especially in some specific activities' timestamp. Considering the type of database, if the field is empty, it may be assumed that either the process is not yet done or it is done by other means. For example, when the inspection is conducted through Hi-Co Scan, some fields in the database are empty because the completion of PIB does not pass through certain activities. Thus, the researcher must consider excluding these data from the database used for analysis.

The length of service in every procedure is obtained by reducing a timestamp in one phase to another's. The timestamp in every phase has to be continual. If a timestamp in the second phase of the procedure is less than the first one, then the length service is a minus. For instance, when a customs specialist has started looking at the detailed information in PIB whose inspection is not yet finished. The researcher will keep this type of data because the anomaly might form any pattern that affects the process.

3. Data preparation

The data preparation phase covers all activities to construct the final dataset or the data that will be fed into the modeling tools from initial raw data (Shearer et al., 2000). They further mention that there are five steps in data preparation as follows:

a. Selecting data

In selecting data, the researcher has considered the availability of database regarding confidential information shared in PIB. Some details, such as data related to the importer or PPJK cannot be shared except for investigation purposes. However, by extracting data from CEISA, there are three different types of influential factors proposed, namely importation-specific, user-specific, and procedure-specific factors. Importation-specific factors represent general information of PIB. From the user perspective, there are at least two different officers who have an active role in the customs clearance process, which are the customs inspector and PFPD. The last yet probably the most important factor comes from the procedure. The

researcher distinguishes every activity during the customs clearance process. Every PIB may be treated differently in every activity. Thus, the length of service performed is considered one of the most influencing factors. Details of proposed categorical factors, type of information collected, and attribute code used for this research, can be seen in the fourth step of data preparation; integrating data.

Certain data, such as the length of document verification, are not included in this study, simply because the researcher finds it difficult to measure the exact time for verifying one document. However, there are six possible responses that can be issued in every PIB during its verification. Any responses except for SPPB involve external stakeholders to give feedbacks. For example, when a confirmation note, INP, or SPTNP is issued, the duration of document verification should be stopped until the importer or PPJK fulfills any requirements stated in the response. It is unfair for customs specialists to burden the entire process themselves.

b. Cleaning data

Based on data extracted from CEISA over a period of 10 months starting from January 2016, there are originally 21,565 PIB declared in the red channel which means those PIB must go through a physical inspection. Among this number, 1,115 PIB are inspected through Hi-Co Scan. Due to different treatment during the inspection and limited timestamps received from the data collection process, these PIB should be excluded from this study. Moreover, 158 PIB whose timestamp in SPPB is empty are eliminated as well. No record on SPPB indicates that the customs clearance process is not yet finished. Altogether, the researcher uses 20,292 rows of data that represent 20,292 PIB in total for the study.

c. Constructing data

New records are constructed regarding user-specific factors. It is obtained from the employee database at PCO Tanjung Priok. New attributes developed include the level of grading, years of services, and tenure for both PFPD and the customs inspector. In addition, the last education of PFPD and the participation of a customs inspector in technical inspection training are added as the new attributes. Altogether, these new attributes are expected to represent users' professional judgment in performing their duties especially related to the customs clearance process.

d. Integrating data

In this stage, the researcher combines the new attributes developed relating to user-specific factors into the primary dataset which consists of 20,292 data rows. This process produces a complete dataset that is ready to use in classification. There are 34 attributes in

total including one variable target. Detailed proposed categorical factors, types of information collected, and attribute code utilized for this research are shown in Table 2.

Table 2. Proposed Attributes Used for the Study

Proposed categorical factors	Type of information collected as proposed attributes and the code used	
Importation-specific factors	<ul style="list-style-type: none"> – Detailed channel (CHN); – Category of import (IMP); – Number of item of imported goods (ITM); – Number of container used (CTN); – Type of container used (CTN_LOAD); – Country of origin (COUNTRY); – Harmonized system codes (HSCODE); – Day of week of issuing SPJM (DAY_SPJM); – Time of the day of issuing SPJM (TR_SPJM). 	
User-specific factors	Customs specialist-related: <ul style="list-style-type: none"> – Code of customs specialist (CS); – Level of grading (GRD_CS); – Years of service (YOS_CS); – Tenure (TNR_CS); – Last education (EDC_CS). 	Customs Inspector-related: <ul style="list-style-type: none"> – Code of customs inspector (CI); – Level of grading (GRD_CI); – Years of service (YOS_CI); – Tenure (TNR_CI); – Technical inspection training (TRN_CI).
Procedure-specific factors	<ul style="list-style-type: none"> – Time needed for collecting hardcopy of PIB (HC); – Time needed until the container is ready to be examined (READY); – Time needed to get a customs inspector (ASSIGN); – The length of conducting physical inspection (INSPECT); – The length of uploading inspection report (LHP); – Time lag before a customs specialist starts PIB verification (LAG); – Kind of response issued; a confirmation note (CONFIRM), a consultation note (CONSUL), a laboratory test (LAB), a customs value information (INP), a customs value declaration (DNP), a prohibition and restricted goods assessment form (SPBL), a tariff and customs value assessment form (SPTNP), and an audit marking (AUDIT). 	

e. Formatting data

There are two types of software used: Microsoft Excel 2011 to build the dataset, and Weka 3.6.14 version to create models prediction. The type of file that can be read in both software is *.csv (comma delimited) file.

4. Modeling

a. Building the model

In order to achieve the best output, parameters of the model need to be calibrated to optimal values. There are 11 classification models developed in this study using a decision tree with J48 algorithm. Specifically, one model built using unpruned, 5 (five) models using online pruning, and 5 (five) models using post-pruning. Details of parameters setting used to build the classification models are shown in Table 3.

Table 3. Summary of Parameter Setting Details Used to Build Classification Models

Parameter	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
	<i>Unpruned</i>	Online Pruning					Post-pruning				
<i>confidenceFactor</i>	1	1	1	1	1	1	0.75	0.5	0.25	0.05	0.001
<i>MinNumObj</i>	2	50	150	250	350	450	2	2	2	2	2
<i>numFolds</i>	10	10	10	10	10	10	10	10	10	10	10
<i>seed</i>	1	1	1	1	1	1	1	1	1	1	1
<i>unpruned</i>	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
<i>testmode (cross folds)</i>	10	10	10	10	10	10	10	10	10	10	10

b. Assessing the model

In this step, the researcher needs to interpret the models using these following indicators; number of leaves, size of the tree, correctly classified instances, kappa statistic, ROC area and sensitivity. The interpretation in every model is required to select one or several models with the best classification. The best model is expected to produce a simple yet understandable tree to identify factors affecting the customs clearance process. Summary of indicators resulted in every model built in this study are presented in Table 4.

Table 4. Summary of Models' Indicator Value

Indicator	Model										
	1	2	3	4	5	6	7	8	9	10	11
<i>Leaves</i>	9669	416	194	18	13	11	9669	1161	119	91	42
<i>Tree</i>	10065	471	218	35	25	21	10065	1342	204	159	83
<i>CCI</i>	89.05	90.39	88.66	87.87	86.66	86.11	89.06	90.33	90.76	90.7	90.45
<i>Kappa</i>	0.696	0.731	0.683	0.65	0.625	0.63	0.697	0.731	0.741	0.739	0.72
<i>ROC</i>	0.884	0.945	0.921	0.901	0.886	0.881	0.89	0.912	0.948	0.948	0.937
<i>TP Rate "on time"</i>	0.935	0.947	0.935	0.945	0.925	0.907	0.935	0.944	0.95	0.951	0.953
<i>TP Rate "late"</i>	0.751	0.767	0.733	0.671	0.681	0.726	0.752	0.776	0.773	0.769	0.752
<i>WA TP Rate (sensitivity)</i>	0.89	0.904	0.887	0.879	0.867	0.864	0.891	0.903	0.908	0.907	0.905

Source: Developed from classifier output in Model 1 to Model 11

Table 4 shows that there are two models whose value of number of leaves and size of the tree are the biggest; model 1 and model 7. Model 1 is built unpruned while model 7 is with post-pruning in which the confidence factor value is held at 0.75. Both models are eliminated simply because they have too high value in both number of leaves and size of the tree. Kappa statistic, in this case, cannot be used to interpret the models because the value in every model is in the same level category, which is substantial agreement.

The value of ROC Area in the eight models remaining varies in two categories. Based on its ROC Area value, these following models: model 2, model 3, model 4, model 8, model 10, and model 11 are built with excellent classification, better than model 5 and model 6. Thus, the last two models are removed from the available options for the next assessment.

To choose the best among the remaining seven models in which its value of kappa statistic and ROC Area are all in the same category, the researcher considers the sensitivity of the models. It can be shown from the value of true positive (TP rate) specifically for instances classified as “on time”. The researcher separates the remaining seven models based on its pruning featured used in building a decision tree: online pruning and post-pruning, and chooses one best model in both categories. In online pruning category, model 4 is considered as one best model because it has quite high value of TP rate for “on time” yet establishes the simplest decision tree.

In the second category, model 11 has the greatest value of TP rate “on time”. Favorably, it has the smallest value of number of leaves and size of the tree as well. Based on these indicators, model 11 may be considered one of the best models under the category of post-pruning. Moreover, one best model is selected by taking the number of leaves and size of the tree produced in both model 4 and model 11. Using model 4, the researcher may be able to develop a tree of size 35 with 18 leaves while the tree built through model 11 produces a bigger tree of size 83 with 42 leaves. Although model 11 has greater value statistically, it is still in the same category of model 4. Hence, model 4 is the one considered the best model to identify factors affecting the customs clearance time at PCO Tanjung Priok.

5. Evaluation

a. Model evaluation

Model 4 is an online-pruning decision tree which is built by altering the value of minimum instances (minNumObj) per node at 250. Statistically speaking, it has the accuracy (shown by the value of correctly classified instances) as much as 87.87%. This number represents the model ability to classify 18,385 instances into the correct category. The value of kappa statistic produced in this model, which is at 0.65 is considered in the substantial agreement level. Moreover, model 4 is supported by a high value of ROC Area at 0.901 that represents an excellent classification used in building the decision tree. The confusion matrix of model 4 shows that there are 15,002 instances in “on time” category correctly identified, while 877 other instances that should have been classified as “on time” are misclassified into the “late” class. For total 5,043 instances in the “late” category, 2,383 instances are accurately identified in its original class and the remaining 1,660 “late” instances are incorrectly classified as “on time”.

b. Pattern evaluation

This step is more focused on how to extract essential information from the decision tree produced by the model. From 34 attributes proposed including the variable target, model 4 only uses five attributes in building a decision tree, as follows:

- 1) Time required until the container is ready to be examined (READY);
- 2) Time lag before the customs specialist starts PIB verification (LAG);
- 3) The issuance of a tariff and customs value assessment form (SPTNP);
- 4) Time needed for collecting hardcopy of PIB (HC);
- 5) The issuance of customs value information (INP).

All attributes employed to develop the tree come from procedure-specific factors. In developing a decision tree, model 4 eliminates 28 other attributes. However, there are at least ten conditions in which a “late” instance is most likely to occur at PCO Tanjung Priok, namely:

- 1) If the value of READY is greater than 118.6 hours, there is a 97% chance that PIB is served exceeding 5 days;
- 2) If the value of HC is less than 24.52 hours and the value of READY is in the range of 92.98 to 118.6 hours, the opportunity of late service appeared is at 91%;
- 3) If the value of HC is less than 24.52 hours and the value of READY ranges from 92.98 to 118.6 hours, but the value of LAG is more than 20.49 hours, then there is a 78% chance of PIB being served late;
- 4) If the value of READY is 92.98 hours or less but the value of LAG is more than 114.47 hours, the chance of late service performed is 97%;
- 5) If the value of READY in the range of 33.12 to 92.98 hours and the value of LAG is more than 74.74 hours, there is a 91% chance of late service being given;
- 6) If there is a combination of these condition: SPTNP is issued, the value of READY is 92.98 hours or less, the value of LAG is less than 74.74 hours, and the value of HC is more than 25.11 hours, then the probability of a late service being performed is at 83%;
- 7) If the combination of these conditions appears: SPTNP is delivered, the value of READY is between of 38.05 and 92.98 hours, the value of LAG is less than 74.74 hours, and the value of HC is 25.11 hours or less, it is most likely that the chance of late service being performed is at 74%;
- 8) If these following conditions occur in one PIB: the value of READY ranges from 21.56 to 92.98 hours, the value of LAG is less than 74.74 hours, the value of HC is

more than 72.64 hours, but SPTNP is not issued, then there is an 83% chance of late service being performed to that PIB;

9) If the value of READY is in ranging from 39 to 92.98 hours, the value of LAG is less than 74.74 hours, the value of HC within range of 17.792 to 72.64, SPTNP is not issued yet INP is delivered, and they all happen in one PIB, then a late service is most likely performed by 74.85% chance;

10) If these following condition occur altogether: the value of READY ranges from 48.3 to 92.98 hours, the value of LAG is between 46.69 to 74.74 hours, and both SPTNP and INP are not issued, there is a 69% chance of late service being performed, even though the value of HC is less than 72.64 hours.

6. Deployment

The average time to prepare for containers to be ready for inspection over period of 10 months is approximately 41 hours. Using the decision tree built by model 4, the maximum value in which a PIB most likely to receive on time service is 21.56 hours. This amount is obtained from the range of amount in READY attribute that is excluded from the ten conditions to predict a late service performed to a PIB. There are in fact 12,162 out of 20,922 instances whose value in READY attribute is greater than 21.56 hours. This fact may be a reason for it to be chosen as the root node. Considering this amount, the researcher agrees on the implementation of DGCE Decree number PER-12/BC/2016 concerning the simplification of physical inspection in which it forces the participation of all stakeholders in burdening the responsibility in the dwell-time process starting from August 10th, 2016. By fully implementing this regulation, it is expected that every PIB would have a value of READY attribute of less than 24 hours.

The time required for the importer or PPJK submitting the hardcopy of PIB on average is 21.3 hours, while the baseline determined by model 4 is at a value of 17.92 hours. In fact, the maximum time set for submitting the hardcopy is 24 hours which is still reasonable for the moment. Thus, PCO Tanjung Priok shall consider another way to encourage the importer and PPJK to submit the hardcopy earlier rather than reducing the limitation time, such as counseling or holding a workshop concerning how important their role as stakeholders in reducing the dwell-time which customs clearance time is included.

Unlike the previous factors, the time lag from when inspection report (LHP) is uploaded and starting time of the document verification by a customs specialist, is surprisingly a new factor identified affecting the customs clearance process. Although the average duration of this time lag is 24.42 hours, but the baseline obtained from model 4's decision tree is 46.69

hours. This number shows that in the “late” class, it is least likely to have LAG attribute whose value is less than 46.69 hours. Moreover, there are 3,506 out of 20,922 instances in which its value of LAG attribute is greater than the baseline.

Practically, there are no regulations or policies that place a limitation on how long the time lag should be. Thus, PCO Tanjung Priok shall take this factor into account by establishing new policies to continue to urge the customs specialist to prioritize in serving PIB declared through the red channel. Proposing the policy concerning the new calculation of number of credits for the customs specialist. The existing policy does not consider the type of channel. There should be a distinction of the weight of score in every document based on the type of channel for which the completion of PIB declared through the red channel may earn more score. This proposal may be one of the alternatives to motivate the customs specialist in performing PIB completion for the red channel.

Regarding another factor identified related the responses issued, both SPTNP and INP involve the participation of third parties. When INP is issued and the importer does not submit DNP within three days, the customs specialist may not consider any other option other than his/her assessment. This condition usually leads to the issuance of SPTNP. Based on all data collected, there are only 1,647 DNP submitted out of 3,697 INP issued, yet 1,630 INP issued lead to the issuance of SPTNP. Another fact revealed is that the average time required by the importer to collect proofs or any information required from INP is one day. Although this amount of time is considered to be good, but it does not represent all DNP collected. Due to manual input, it is possible that DNP has been submitted yet the customs specialist does not mark it on CEISA. Thus, in control of this manual marking activity, PCO Tanjung Priok shall consider giving the authorization to the first officer who receives DNP from the importer or its representative.

Regarding the issuance of SPTNP, from the customs’ perspective, the fulfillment is the most simple. Once the underpayment is settled, SPPB will be issued. In fact, the average time of the underpayment settlement (the time required from when SPTNP is published to SPPB issued) is 2.3 days, which is considered quite long comparing to the length of other activities performed. This amount of time should have been reduced, at least, by half. One way to achieve this condition is by presenting more consequences and forcing the importer or PPJK to settle the underpayment faster, such as blocking the importer temporarily from the system, for those who have not yet settled SPTNP within two days (with consideration of the average time required in SPTNP settlement).

When SPTNP is issued, the importer or PPJK may now know the reason behind the underpayment. Before making any payment, they are most likely to seek the information regarding the underpayment from a general client coordinator. The client coordinator provides an answer directly if possible. Otherwise, they have to pass the complaints to the customs specialist and wait for a response. If the importer is still not yet satisfied with the response regarding the customs specialist's assessment, they can file an objection. A guarantee, as much as the underpayment, is required to settle when an objection is submitted. When the guarantee is paid through a bank, it is difficult to get a refund, in the case of an objection being approved. Thus, many importers consider paying the guarantee in cash through the customs office rather than a bank. Unfavorably, it has the consequence of making the process of SPTNP settlement even longer.

Using a different perspective, another aspect relating to the length of time in issuing SPTNP may affect the customs clearance time. This length of time is calculated from when a customs specialist commences document verification until SPTNP is issued. There are at least 430 out of 2,546 SPTNP issued after 3 days of documents being verified. If the average time of the SPTNP settlement is approximately 2.3 days, then PIB whose SPTNP issued after 3 days is presumably being served late. Thus, one alternative the researcher may offer is to encourage the customs specialist to verify the import documents efficiently. It is an especially sensitive matter for the customs specialist. In fact, there are three out of five factors identified come from the verification of document process in which a customs specialist's professional judgment is irrelevant.

CONCLUSION

These following information are extracted as the conclusion of this research:

1. Applying data mining techniques using a decision tree J48 algorithm to the importation database has been successfully identified factors affecting the customs clearance process at PCO Tanjung Priok. One model, labeled as Model 4, is chosen as the best model in this study;
2. Model 4 which is built using online pruning with the altered parameter of minimum instances per leaf at 250 meets the criteria of model simplicity in which it develops a decision tree of size 18 with 35 leaves. Statistically, the model accuracy is considered good as that is shown by these following indicators: the percentage of correctly classified instances at 87.87%, moderate agreement category of kappa statistic with the value of 0.65, and in the category of excellent classification in which represented by ROC Area at

0.901. Based on the decision tree, there are at least 10 conditions in which a late service being given to PIB can be predicted. Thus, in order to get services on time, these following requirements are the least minimum amount of the factors identified (baseline) that needs to be fulfilled:

- a. The time required for the importer or PPJK submitting hardcopy of PIB should be less than 17.92 hours
 - b. The length of time to prepare the container for a physical inspection shall be less than 21.56 hours;
 - c. The time lag from the uploading of LHP to starting the documents verification by a customs specialist is less than 46.69 hours.
3. There are at least five factors affecting the customs clearance time at PCO Tanjung Priok, namely the length of time to prepare the container for a physical inspection, the time lag from the uploading of LHP to starting the document verification by a customs specialist, the time required for the importer or PPJK to submit hardcopy of PIB, and the issuance of a response by the customs specialist regarding customs duties and taxes underpayment (SPTNP) and a test of reasonableness of customs value set forth in PIB (INP).

RECOMMENDATIONS

The researcher may offer several recommendations as mentioned below:

1. The information extracted from this study especially regarding factors identified may be applied so that it is useful within the organization's decision-making processes. The least minimum amount of the factors identified (baseline) to have an import declaration served on time may be taken into account in establishing some policies regarding the customs clearance process;
2. Considering that PIB through the red channel requires the longest time in the customs clearance process, PCO Tanjung Priok shall continue to urge the customs specialist to prioritize in serving PIB declared through the red channel. One of the alternatives that the researcher may offer is by proposing the new policy that distinguishes the weight of score in every document based on the type of channel for which the completion of PIB declared through red channel will earn more score;
3. User-specific factors that cover the customs specialist and the customs inspector-related information shall be expanded to prove that professional judgment is significantly affecting either part of or the whole customs clearance process;

4. Although the use of data mining identifying factors affecting the customs clearance process has been proven useful, this study still requires improvements, especially more focused on the document verification process simply because three out of five factors identified appear in the document verification process yet there is no certain understanding when the process shall be finished.

REFERENCES

- Artakusuma, A. (2012). Analysis of Import Container Dwelling Time in Jakarta International Container Terminal (JICT) Tanjung Priok, 1–4. Retrieved from <http://www.ftsl.itb.ac.id/wp-content/uploads/2012/11/15008081-Afif-Artakusuma.pdf>
- Chermiti, B. (2019). Establishing risk and targeting profiles using data mining: Decision trees. *World Customs Journal*, Volume 13, Number 2, 39-57. Retrieved from [https://worldcustomsjournal.org/Archives/Volume%2013%2C%20Number%202%20\(Sep%202019\)/1886%2002%20WCJ%20v13n2%20Chermeti.pdf?_t=1569889901](https://worldcustomsjournal.org/Archives/Volume%2013%2C%20Number%202%20(Sep%202019)/1886%2002%20WCJ%20v13n2%20Chermeti.pdf?_t=1569889901)
- Geourjon, A., Laporte, B., Coundoul, O., & Gadiaga, M. (2012). Inspecting less to inspect better : The use of data mining for risk management by customs administrations.
- Han, J., Kamber, M., & Pei, J. (2006). *Data Mining: Concepts and Techniques*. San Francisco, CA, itd: Morgan Kaufmann. <http://doi.org/10.1016/B978-0-12-381479-1.00001-0>
- National Academy of Customs Excise and Narcotics. (2015). Data Warehousing and Data Mining for improvement of Custom Administration in India – Lessons learnt overseas for implementation in India: Data Warehousing and Data Mining for improvement of Custom Administration in India – Lessons learnt overseas for imp, 1–18.
- Provost, F., & Fawcett, T. (2013). *Data Science for Business*. *Book*. <http://doi.org/10.1007/s13398-014-0173-7.2>
- Shearer, C., Watson, H. J., Grecich, D. G., Moss, L., Adelman, S., Hammer, K., & Herdlein, S. a. (2000). The CRISP-DM model: The New Blueprint for Data Mining. *Journal of Data Warehousing*14, 5(4), 13–22. Retrieved from www.spss.com%5Cnwww.dw-institute.com
- USAID Trade Project. (2014). *Dwell Time Study*.
- Wang, Y. Q., & Song, Y. M. (2011). Appling Data Mining Technology to Analasis Clearance Efficiency in the Port Logistics. *Key Engineering Materials*, 480–481, 1144–1149. <http://doi.org/10.4028/www.scientific.net/KEM.480-481.1144>
- Zhou, X., (2019). Data mining in customs risk detection with cost-sensitive classification. *World Customs Journal*, Volume 13, Number 2, 115-130. Retrieved from [https://worldcustomsjournal.org/Archives/Volume%2013%2C%20Number%202%20\(Sep%202019\)/1886%2002%20WCJ%20v13n2%20Xin%20Zhou.pdf?_t=1569889901](https://worldcustomsjournal.org/Archives/Volume%2013%2C%20Number%202%20(Sep%202019)/1886%2002%20WCJ%20v13n2%20Xin%20Zhou.pdf?_t=1569889901)