World Customs Organization

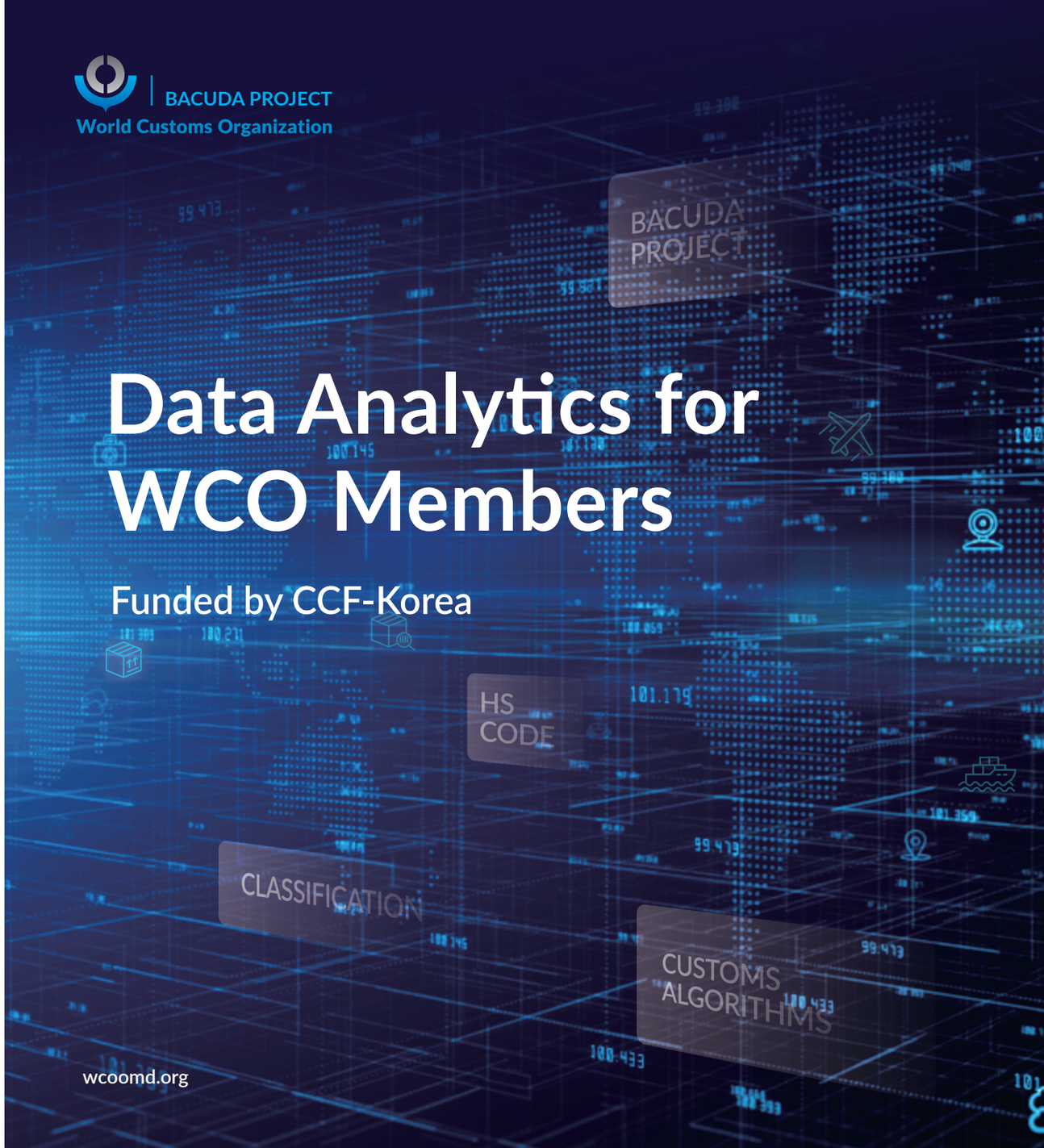
Rue du Marché 30, B-1210
Brussels, Belgium
bacuda@wcoomd.org

#WCOOMD
wcoomd.org

BACUDA PROJECT
World Customs Organization

Data Analytics for WCO Members

Funded by CCF-Korea



wcoomd.org

DATA ANALYTICS FOR CUSTOMS

The rise of Big Data combined with the emergence of powerful and sophisticated analytics tools and technologies has changed the way public and private organizations are operating.

The use of data by Customs is not new. However, Big Data and algorithms are certainly providing new opportunities for Customs, while also raising new challenges in the implementation of Data Analytics.

Such challenges require specific approaches to be considered, as well as specific policy attention to develop the appropriate organizational and technical capabilities needed to take advantage of the explosion in data, and to gain the insights that will help Customs make better informed decisions. Over the past few years, the Secretariat has made numerous efforts to raise awareness among WCO Members about the potential benefits of data analytics, and to support them in their journey towards a successful implementation of the analytics function.

While these efforts have proven to be beneficial to Members to engage in the analytics journey, there was clearly a need to complement them with measures pertaining to other dimensions of institutional capacity building which are outlined in this leaflet.

WCO BACUDA PROJECT

The BACUDA project was implemented in 2019 with funding from the Customs Cooperation Fund of Korea (CCF-K) to raise awareness and build capacity in Data Analytics among WCO Members. The Project team consists of three officials from the Capacity Building Directorate and closely works with BACUDA expert groups from academia and research institutes to develop methodologies for Customs purposes to be deployed among Member Customs Administrations.

The project conducts regional and national workshops to raise awareness and share Members' examples of good practices in the area of Data Analytics.

The project's main achievements include the development of Capacity Building Framework on Data Analytics, practical algorithms for Customs purposes, Online training courses on CLiKC!, and establishment of the BACUDA Scholarship Programme.

2019

Implementation of the BACUDA Project

Funding

Customs Cooperation Fund of Korea (CCF-K)

Objective

Build WCO Members' knowledge and capacity in Data Analytics

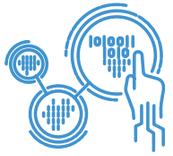
Project team

3 officials from WCO in collaboration with BACUDA expert groups

Instruments

Workshops organised to grow knowledge and share examples in the area of Data Analytics

Outcomes



Algorithms available on the WCO website and shared during WCO meetings

CAPACITY BUILDING FRAMEWORK FOR DATA ANALYTICS

The 82nd Session of the Policy Commission that was held in December 2019, in Seoul, South Korea, recognized that building Organizational Capacity for analytics in Customs required more than just the availability of tools, applications and solutions and thereby endorsed the recommendation to develop a dedicated Capacity Building Framework for Data Analytics which would form part of the WCO Capacity Building tools and serve as a guide to formulate, monitor, and evaluate activities for organizational analytics development in Customs.

The framework uses the results of a survey that was conducted in January 2020 to understand how WCO Members were using data analytics to shape their future and what challenges they were facing in implementing data analytics.

It provides information on how to effectively implement successful analytics initiatives and build the organizational capacities that are needed to optimize the use of data analytics.

Specifically,
it addresses the following questions

1. What types of assessments and planning should occur prior to undertaking an analytics project?
2. What data and analytics processes should be in place?
3. What would be the most appropriate analytics organizational structure?
4. Why is data governance important, and how can a data governance framework be developed?
5. What skillsets and types of staff are needed?
6. Why and how is change management critical?

CUSTOMS ALGORITHMS

DATE ALGORITHM FOR DETECTING UNDEREVALUATION

This model employs a cutting-edge mechanism; "ATTENTION", a rising superstar for AI-based language translation and self-driving cars. Thanks to this up-to-date technology, the model outperforms other traditional machine learning models (such as XGBoost) in fraud detection.

Notably, the model also performs better with relatively small-sized training data (of countries with low trade volumes) and low inspection rates (of countries with huge trade volumes).

Here is a brief description of the DATE algorithm in human language:

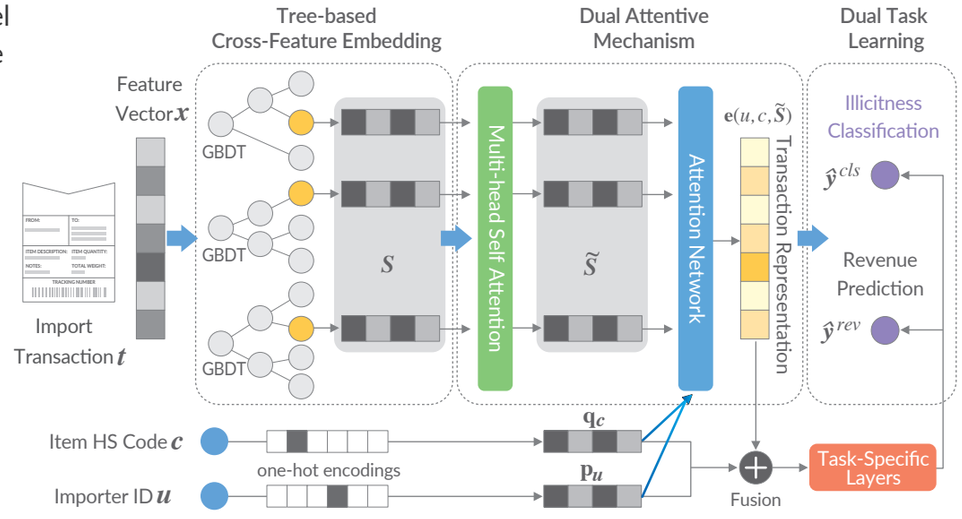
- Imagine that you ("neural networks") are the head of a Customs targeting centre composed of 100 risk analysts ("decision trees"). For a given import, you task the analysts with reporting the probability of undervaluation and the estimate of additional revenue from the inspection ("dual-task").
- **How would you put 100 different reports together in making your final decision?**

  Simply averaging their predictions may neglect some valuable information hidden in 100 reports. The DATE model helps you keep all the information while paying more ATTENTION to more important pieces of information.

- **How to use and learn the DATE algorithm?**

  The DATE is an open-source algorithm using Python code. It is illustrated in the intermediate training course on CLiKC! where Members can learn how to run the algorithm and test with their own data through course exercises.

DATE model architecture



Source: https://github.com/Roytsai27/Dual-Attentive-Tree-aware-Embedding

Firstly,

If there are a majority group of reports significantly similar to each other, you may pay more ATTENTION to those reports.

Secondly,

If you have analysts specialized in the specific HS code and importer of the given import, you may pay more ATTENTION to their reports.

In the end,

Your final decision reflects the reports attracting your highest attention.

HS CODE RECOMMENDATION ALGORITHM

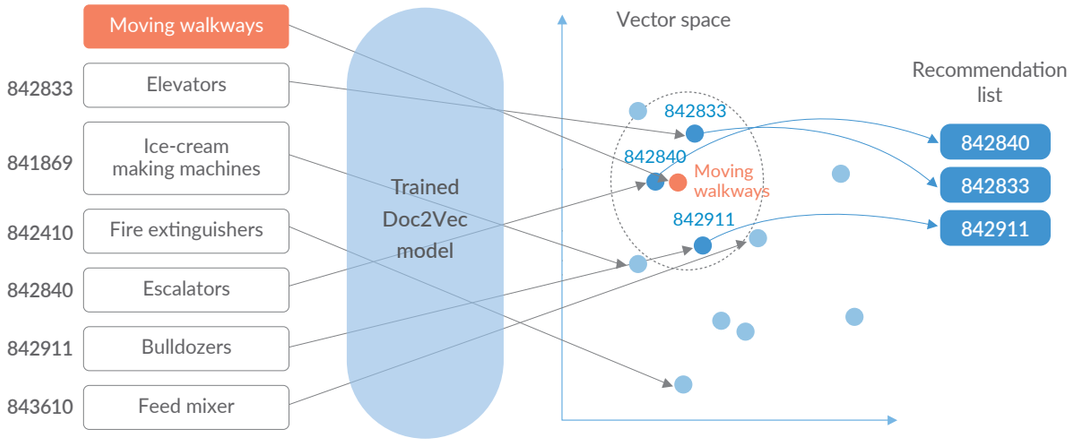
HS misclassification is very common, and the most difficult task in international trade; 1 out of 3 customs entries is misclassified, and tens of billions in duties is incorrectly paid. Moreover, HS classification remains a manual and time-consuming process, requiring deep domain expertise and years of experience.

Using text embedding techniques, the item/product descriptions that are semantically similar can be grouped together. Basically, this analytic model converts the given product description into vectors, a numeric representation of import declarations, then calculates the distances between the declarations and provides the predictions of the appropriate HS codes as a recommendation.

The algorithmic model consists of five stages: data preparation, data preprocessing, data sampling, modeling and training. The core stage is the modeling Doc2Vec classifier with the given product descriptions and the HS codes. The trained model can make recommendations based on its similarity calculations.

This model is expected to serve in the import and export declaration process and can be utilized for better risk assessment and prevention of misclassification fraud.

By providing appropriate HS code options to the field officials, the model is expected to reduce the declaration processing time and reduce HS code misclassification, leading to a reduction in duty underpayment.



DO YOU WANT TO BE INVOLVED IN A PILOT PROJECT?

WCO Members can contact the BACUDA Project team to organize a pilot for joint algorithm testing with their data. As part of a pilot project, volunteering Members will provide past import and export data which will be used to test our algorithms online.

In the second phase, these algorithms may then be deployed, based on the Member's needs, on real-time data at a designated port. It is recommended that Members visit the Data Analytics advanced level E-Learning course to learn how these algorithms work.

E-LEARNING COURSES ON DATA ANALYTICS

1 BEGINNER

As part of the BACUDA project, the online training course on Data Analytics for Beginners in English, French, Spanish, Arabic and Russian has been developed and made available on CLiKC!, the WCO E-Learning Platform.

The course features :

Basic training on Data Analytics and machine learning;	Short term introductory curriculum for knowledge building on data analysis and a session for upskilling Customs analysts on machine learning;	Video lectures, online textbooks and practical exercises.

2 INTERMEDIATE

The intermediate course on Data Analytics has been developed in English and French, targeting entry-level data analysts in Customs with basic knowledge and officials who have completed the beginner course and wish to broaden their skills.

The intermediate course features :

Hands-on lessons on machine learning algorithms and its applications;	Training on how Customs can make the best use of machine learning techniques including the deep learning model;	Exercises on the DATE algorithm – one of the first outcomes of the BACUDA project.

3 ADVANCED

The data analysis application course has been developed for Customs officials to acquire practical skills that can be applied to daily operations. The course follows on from the beginner and intermediate courses on CLiKC!, which officials are recommended to have completed beforehand.

This application course, available in English and French, offers Customs officials :

Training procedures on AI models used in the application;	An introduction to basic theory and preliminary work on HS classification and the DATE algorithm for detecting fraud;	Step by step practice in algorithmic HS classification and recommendation.

HOW TO REGISTER AND BEGIN LEARNING ONLINE

Without any limit, every Customs official can enjoy the E-Learning courses available on CLiKC!.

To register,
follow the simple steps below:

1

Go to the CLiKC! Home page
https://clikc.wcoomd.org

2

Select the "**Get registered**" tab at the bottom of the page

3

Complete and submit the registration form, and wait for your account request to be approved by your country's National Coordinator

THE BACUDA SCHOLARSHIP PROGRAMME

The WCO has decided to implement a Scholarship Programme to boost its Members' capacities in the area of Data Analytics.

The Programme will select 12 Customs officials from developing countries to provide an opportunity for participants to pursue 5 months of in- depth studies and training in Customs-related subjects at the Sungkyunkwan University (SKKU, www.skku.edu/eng/index.do) in Seoul, South Korea.

Students who successfully complete the Programme will be awarded an official certificate. The Scholarship Programme will provide intensive training on Data Analytics-related issues for Customs officials, focusing on Customs operations.

It will be taught in cooperation with the Korean Customs Service and Customs training institute, and will offer classes not only on practical knowledge, but also academic theory through lectures, workshops and seminars, as well as visits to regional Customs offices. Those who wish to benefit from the Scholarship Programme are required to complete two Data Analytics online courses on CLiKC!.

The Programme will be financed by CCF-Korea. The Fund will cover relevant costs for delivering the Scholarship Programme, including travel costs, admission fees, tuition fees, institutional costs, accommodation costs, subsistence allowance, and other approved incidental expenses to enable participants to complete the Programme at the SKKU.