[2020 Winter/Spring semester URP] KAIST URP Program

Individual Research Project Final Report

Research Participation Project

세관 통관 사기 검출을 위한 액티브 러닝 방법론

Active Learning for Custom Fraud Detection

School of Computing - Mai Tung Duong (20180745)

Contents

Abstract

Customs is an authority responsible for controlling the flows of goods across borders and collecting tariffs. However, traders might commit fraud by manipulating declaration details to avoid ad-valorem duties and taxes. Customs administrations need to detect and duly collect the custom tariffs from these illicit trades without interrupting legitimate trade flows. Many effective and interpretable methods for predicting suspicious trades are devised using various machine learning frameworks. The state-of-the-art model can catch 92.7% of illegal cases and secure 49.3% of revenue by inspecting only 1% of trades.

However, these models usually assume the fully-supervised static setting. The setting is unrealistic since real-life custom declarations coming continually with potentially different trade patterns, a problem known as domain shift. Even well-trained models might not catch these shifts, and continual labeling of training examples is a costly task in supervised learning. To mitigate these problems, we proposed an active learning framework to provide a guideline on how to interactively select and annotate data points to adapt to the new trade patterns. *Our strategy outperforms the state-of-the-art active learning strategy by 3-7% in customs settings.*

In the conventional active learning setting, the model encourages to query data to clear uncertainty and secure diversity. However, in a real-world online prediction scenario, queried data are often subjected to evaluation. This means the system needs to be profitable while it is securing knowledge for the future. So the learning model cannot fully follow the explorative principles of active learning. To describe this exploration-exploitation dilemma, we introduce a human-in-the-loop customs selection scenario in which customs administrations maintain an AI-based selection model to support officers' collecting duties. We show that some exploration is needed to cope with the domain shift through years of customs selection simulation. O*ur hybrid strategy of selecting fraud and uncertain items will eventually outperform the performance of the exploitation strategy.*

Finally, we formulate a multi-armed bandit problem and *adaptively adjust the degree of exploration-exploitation* to achieve both short-term revenue and long-term model performance.

Part 1: Introduction

1. Research background

The customs administration aims to detect fraudulent transactions and maximize the tax revenue from illicit trades. However, as transaction volume is huge, customs administration cannot practically screen all suspicious transactions. While some customs offices had conducted 100% manual inspection, this has become costly and time-consuming with astronomically growing trade volume. Furthermore, most recently, customs offices are compelled to cut down on many intensive operations in light of the health pandemic and social distancing. There is a growing need to develop an efficient selection strategy to identify suspicious trades and increase revenue.

Many machine learning techniques are adopted to devise models for customs selection. However, these models often assume a fully supervised, static setting. This conventional setting is unrealistic as customs data labeling a costly task with proprietary nature. Also, the trade data comes continually with possibly pattern shifts. In the context of customs operations, the list of countries procuring a particular product will change over time, and some importers can declare an unknown product. Even a well-trained machine learning model can fall into the trap of confirmation bias and may not capture these changes. Particularly for situations in which manual labeling is expensive, it can be challenging to make significant changes to the model's working logic.

2. Research objective

Figure 1: Illustration of the customs clearance process

Figure 1 depicts the customs clearance process. To trade goods across borders, importers need to specify the trade items' information in import declaration forms. The officers, using a set of selection rules (possibly advised by an AI-based model) to determine which items should be inspected. Officers determine the authenticity of the items. If the items are fraud, offices levy additional duties accordingly. Due to astronomical trade volume, not all items are subject to inspection. The amount

of additional tariffs that can be secured differs by item. Besides, it is challenging to calculate the potential benefits of examining uncertain items. Once the items are inspected, the selection model's parameters can be updated using newly inspected items as additional training samples. Overall, devising a selection strategy that *can maximize the current performan*ce while *securing its knowledge for the future* is crucial for maintaining the customs selection system in the long run.

3. Related works

Customs fraud detection. Earlier research on customs fraud detection has focused on rule-based or random selection algorithms [13, 19]. However, these classical methods are known to be brittle and are hard to maintain. The application of machine learning in customs administration has been a closed task primarily due to data's proprietary nature. Several recent studies have shown the use of off-the-shelf machine learning techniques like XGBoost, SVM [6, 27]. State-of-the-art is the DATE (Dual Attentive Tree-aware Embedding) model, which employs the transaction-level embeddings in customs fraud detection [17]. This model gives interpretable decisions that can be checked by the customs officers and achieve high revenue in the collected tax.

However, an optimized algorithm is expected to face performance degradation over a longitudinal time period due to limited adaptability to uncertainty, diversity, and the domain shift [22] in customs traffic. This exploitation versus exploration dilemma is what we try to tackle in this paper.

Active learning. Active learning is one of the promising domains in machine learning. It enables an algorithm to elicit ground truth labels for uncertain data instances and enhance its performance [12, 24]. This learning technique is utilized for training the model to deal with high dimensional data [11], to offer long-term benefits [7, 20], to select appropriate data instances to speed the model training [25], or to train the model with limited budgets [29] effectively.

One research proposes a way to measure the informativeness of given samples [10, 15]. This approach tries to collect as much information as possible to boost the training by choosing the samples it finds uncertain [14, 28]. Another line of research focuses on improving diversity by collecting diverse samples representing the overall data distribution. Diversity based algorithms include region-based active

learning [8, 9] and the core-set based approach [23]. Recent research has also focused on concurrent inclusion of both uncertainty and diversity aspects [2, 30].

There are two main differences between existing active learning approaches and our customs selection settings. First, active learning research usually assumes the offline setting [18, 26], including the state-of-the-art BADGE algorithm [2]. This assumption is impractical and far from the real-world customs selection scenario. Even if some of the algorithms set the training dataset label as dynamic, the test dataset in the offline setting remains static. Over time, in the customs selection setting, the evaluation dataset's domain can become outdated.

Secondly, our customs selection setting does not have any explicit test data for evaluation, and the performance evaluation of the selected item is done during the annotation process. Since the queried fraudulent items (i.e., exploitation) and uncertain items for acquiring new information (i.e., exploration) share the same budget for annotations, the balancing between exploration and exploitation becomes crucial. This constrained optimization setting is not common in conventional active learning techniques.

Recent studies have come up with an interpolating exploration and exploitation strategy in the context of active learning, as pure exploration or exploitation is not applicable to test datasets that are prone to domain shifts [16, 21].

Randomized algorithm. In customs selection scenarios, one can also assume the presence of fraudsters that would act as an adaptive adversary of our model. Research literature involving online learning algorithm [4, 5] has shown that randomness in the selection process improves an algorithm's competitiveness under an online setting ski-rental problem, as a randomized algorithm is more robust against adaptive online adversary model. In this light, we additionally introduce randomness to our sampling strategy.

Part 2: Problem definition

1. Customs selection problem formulation

The customs administration's goal is detecting illicit transactions and securing the maximal tax revenue - this is the customs fraud detection problem. Given an import trade flow T, the main goal is to predict both the fraud score y^{cls} and the raised revenue y^{rev} obtainable by inspecting transactions x.

However, as transaction volume is huge, customs administration cannot practically screen all suspicious transactions. While some customs offices had conducted 100% manual inspection, this has become costly and time-consuming with astronomically growing trade volume. Furthermore, most recently, customs offices are compelled to cut down on many intensive operations in light of the health pandemic and social distancing. There is a growing need to develop an efficient selection strategy to identify suspicious trades and increase revenue. We formulate the customs trade selection problem as follows:

> *Given a trade flow T, can we construct a selection strategy f that maximizes the detection of fraudulent transactions and the associated tax revenue?*

Trade flow T consists of the online stream of trade records, which include importer id and HS code of the goods. With time, the dynamics and distribution of fraud transactions change.

2. Active selection for online setting

Table 1: Notations used throughout the paper.

Symbol	Definition
$\mathcal{T}$	Import trade flow
f	Customs selection strategy
r	Inspection (selection) rate
$\mathcal{B}_t$	Items that is received at timestamp t
$\mathcal{B}_t^S(f)$	Items selected by strategy f at timestamp t, $\mathcal{B}_t^S(f) \subset \mathcal{B}_t$
$\mathcal{B}_t^F(f)$	Items selected as frauds, $\mathcal{B}_t^F(f) \subset \mathcal{B}_t^S(f)$
$\mathcal{B}_t^U(f)$	Items selected as uncertain, $\mathcal{B}_t^U(f) \subset \mathcal{B}_t^S(f)$
$\mathbf{x}_i$	i-th transaction from the selected batch $\mathcal{B}_t^S$
y_i^{cls}	binary variable denoting item $\mathbf{x}_i$ is fraud
y_i^{rev}	non-negative value denoting item $\mathbf{x}_i$'s revenue
X_t	Training data for timestamp t, $X_t = X_0 + \bigcup_{k=1}^{t-1} \mathcal{B}_k^S(f)$
m	Evaluation metric for $\mathcal{B}_t^S(f)$, (e.g., Revenue@k%)

Current research on customs fraud detection mainly concentrates on the static setting, in which a model is trained from large training batches and deployed for fraud detection without updates. This approach assumes that the training and test data in trade flow T would be drawn from the same distribution. However, in an online setting, fraud transactions' distribution changes over time, and traditional

approaches will fail to detect novel frauds.

This project builds a customs selection model that best performs in an online setting, namely *Active Customs Selection.* Unlike the customs fraud detection problem that only requires selecting the maximum fraudulent transactions for every timestamp, the active customs selection problem also requires the selection strategy to help the model update and adapt for new types of fraud. All inspected items can bring additional information, but choosing items that aid the model performance with the selection strategy to improve the model's future performance is crucial in this problem.

Algorithm 1 Active Customs Selection

Input: Inspection rate r%, historical data X_0, unlabeled datastream of new items in each timestamp $\mathcal{B}_t$, number of timestamps T
Output: Items for inspection in each timestamp t
Initialize the training data as the historical data $X_1 = X_0$
for $t = 1, 2, 3, \ldots T$ **do**
  Obtain the batch of new items $\mathcal{B}_t$;
  Get the inspection budget $k_t = |\mathcal{B}_t| \times r\%$;
  Train the strategy f with X_t;
  Based on the trained strategy f, select a set of k_t items $\mathcal{B}_t^S(f)$ for manual inspection;
  Get the ground-truth annotation $(\mathbf{x}_i, y_i^{cls}, y_i^{rev})$ for each item $\mathbf{x}_i \in \mathcal{B}_t^S$ after manual inspection;
  Add the newly annotated items into the training data: $X_{t+1} = X_t \cup \mathcal{B}_t^S$;
end

We can formally define the active customs selection problem as follows: At each time t, given a batch of items B_t from the trade flow T, based on a strategy f trained with X_t, customs officers will select a batch of items B_t^s to inspect and add into the training set manually. The goal is to devise a strategy f_m that maximizes precision and revenue in the long-term future (let say from timestamp t_0 onward).

$$\mathrm{f_m} = \mathrm{argmax_f}\left(\frac{1}{\mathrm{T}-\mathrm{t_o}+1}\sum_{\mathrm{t=t_o}}^{\mathrm{T}}\mathrm{m}\left(\mathrm{B_t^s(f)}\right)\right)$$

where m is the evaluation metric (precision or revenue for fraud transactions). Table 1 lists some frequently used notations throughout the paper, and the main training process for fraud detection with active customs selection is described in Algorithm 1.

3. Data description and evaluation settings

Datasets. For experiments, we employed transaction-level import declarations of three countries in Africa. The import data fields include numeric variables such as item price, weight, quantity, and categorical variables such as tariff code, importer ID, country code, and received office. Each country had slightly different data variables, that were preprocessed by following the previous study [17]. Note that

these three customs were subjected to detailed inspection (i.e., achieving nearly 100% inspection rate). But this practice is not sustainable and the customs offices of these countries plan to reduce the inspection rate in the future. Due to the manual inspection, the transaction labels and charged tariffs are accurately labeled in these logs at the single-goods level. Table 2 and Figure 2 depict the statistics of the data we have utilized.

Long-term simulation setting. The experiment aims to find the best selection strategy to maintain the customs selection model in the long run. Therefore, we simulate an environment where a selection model is deployed and maintained for multiple years[1].

Table 2: Statistics of the datasets

Datasets	Country M	Country N	Country T
Periods	Jan 13–Dec 16	Jan 13–Dec 17	Jan 15–Dec 19
# imports	0.42M	1.93M	4.17M
# importers	41K	165K	133K
# tariff codes	1.9K	6.0K	13.4K
Illicit rate	1.64%	4.12%	8.16%

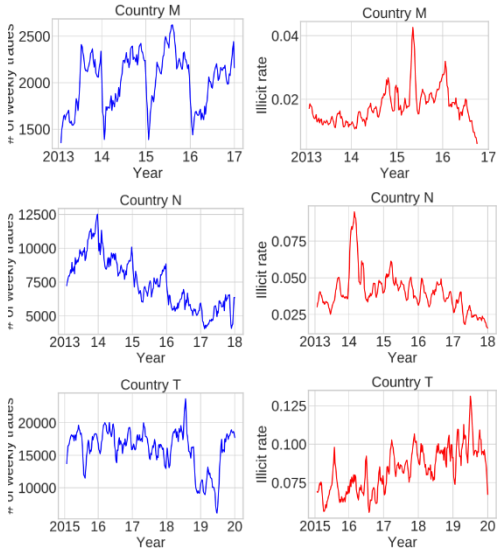


Figure 2: Number of weekly trades and illicit rate

Given that one month's worth of training data is available, the system receives import declarations and selects a batch of items to inspect during the week. A selection model is trained based on a predefined strategy, and the most recent four weeks' worth of data is used for validating the model. By using inspection results, the model is updated every week. To simulate a scenario of data providers who are willing to reduce the inspection rate gradually, we implemented several methods to decay the inspection rate by time. In this experiment, we set the target inspection rate at 10%. Starting with 100% inspection, we used a linear decaying policy by reducing the inspection rate by 10% per week. Once the target inspection rate is reached, the system maintains the inspection rate for the remaining period.[2]

Evaluation metrics We evaluate selection strategies' performance by referring to two metrics used in previous work [17]; Precision@n% and Revenue@n%. Since we

[1] Previous works split data into training and testing sets on the temporal basis and compared the performance of diverse machine learning models [17, 27]. However, the algorithm's performance in static prediction state cannot depict the model's performance in the real setting when the model is deployed.

[2] In countries where the daily import declaration's size is larger, it would be possible to update a selection strategy every day, and more reliable results could be obtained even with a shorter period.

are dealing with an online simulation setting where illicit rate changes each week, we divided each metric's value by the oracle's performance.

- Norm-Precision@n%: Pre@n% divided by the performance measured by oracle. Pre@n% explains how many transactions are illicit, among the top n% of transactions.
- Norm-Revenue@n%: Rev@n% divided by the performance measured by oracle. Rev@n% is the total revenue in top n% transactions identified by a model divided by all transactions' total revenue. This metric explains how much customs duties can be generated from the top n% of transactions than the revenue generated by inspecting the entire transactions. In the following sections, we mainly used this metric to report the results.

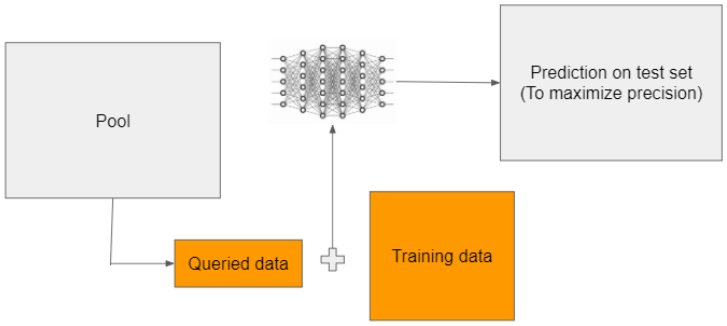
For example, if the system with 5% inspection rate is operating in the environment with 2% illicit rate, Pre@5% and Rev@5% of the oracle would be 0.4 and 1, respectively. Let us consider the deployed selection strategy achieves the Pre@5% value of 0.39. To avoid any potential interpretation bias caused by a small absolute value, we divide 0.39 by the performance upper-bounds of 0.4, which results in 0.975 of Norm-Pre@5%.

Note: Because we are given a fully-labeled dataset, we can measure these metrics with ground truth information. For countries that are already maintaining a low inspection rate, these metrics can be modified by conditioning on their observable goods.

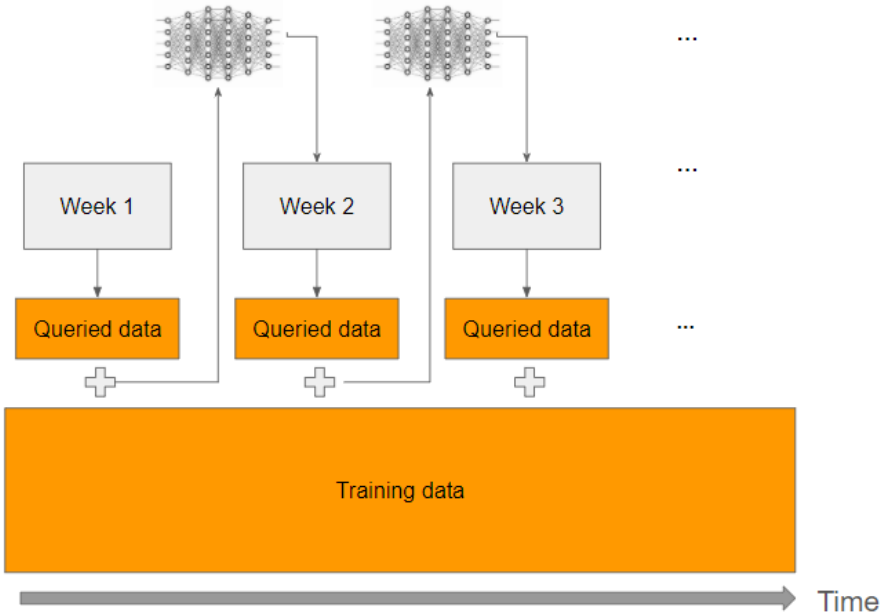
We will show the 13-week moving-average plot of performance over time and report the last 13-week average metrics in all experiment results.

Part 3: Active learning explorative strategies

1. The difference from conventional active learning setting:



(a) Conventional setting



(b) Customs setting

Figure 3: Active learning settings

Figure 3 shows the difference between the realistic custom setting in our project and the conventional active learning:

- The dynamic pool of items: Conventional active learning considers a static pool of inputs and attempts to find a query strategy to pick the most informative items from the set. New trades come continuously in the customs setting, and the items after clearance are released (hence, free from any further inspection). The pool of items changes continuously.
- Revenue consideration: Conventional active learning usually employs

accuracy as the main evaluation. However, both accuracy and revenue must be evaluated in the customs setting, as the customs process aims to detect fraudulent items and secure government revenue.

- The evaluation is done directly on queried items: Conventional active learning adds query items to the training set, and the model is then tested on a held-out validation set. In the customs setting, the queried items are both added to the training set and used for testing the algorithm.

  These are also the obstacles that our method need to overcome

2. Method:

The exploitation strategy selects the more familiar and highly suspicious transactions for inspection; therefore, it tends to underperform over time when trade patterns gradually change. Instead, our hybrid strategy chooses to add a small portion of new and uncertain trades as a learning sample in the training data, which gradually affects the model's future prediction performance. Since fraud types are always evolving, the model performance might drop over time. We propose an exploration strategy to select uncertain trade items, with additional consideration on diversity and revenue, to resolve these issues.

Exploration in light of uncertainty and diversity. One option for detecting new types of fraud is to use the network's uncertainty as a query strategy. Selecting items for which the model is least confident can provide more information on similar new observations. However, this strategy can create an unfavorable scenario where newer labeled data does not include diverse transaction types and labels for identical types of transactions keep on accumulating. In this light, we include the concept of diversity along with uncertainty in our selection strategy, i.e., choosing as diverse samples as possible for stable and fast exploration [2, 18]. By considering uncertainty and diversity concepts, we adopt gradient embedding and k-means++ initialization from BADGE [2] in our exploitation model DATE to figure out which trades should be queried for inspection. Detailed implementation of each concept is described in turn below.

Uncertainty: If a sample generates a large gradient loss and, consequently, a large parameter update, the item potentially contains useful information. This means the magnitude of gradient embedding reflects the uncertainty of the model on samples. With this motivation, we aim to choose trade

flows with uncertainty using magnitude of gradient embedding. At time t, for each trade item x_i in B_t, the illicitness classifier h_θ from the DATE model return its fraud score y_i^{cls} which reflects the illicit class of y_i^{cls}.

$$h_\theta(\mathbf{x}_i) = \sigma(W \cdot z_\phi(\mathbf{x}_i)) \quad (1)$$

where W is a weight matrix that projects the output z_ϕ of DATE illicitness class space.

The gradient embedding g_{x_i} is the gradient of the loss function with respect to W and sample x_i. Since the received data points are unlabeled, we predict the pseudo label $\hat{c}_i$ by the fraud score with threshold 0.5 (i.e. $\hat{c} = \mathbf{1}(\hat{y}_i^{cls} \geq 0.5)$). This pseudo label is used to calculate loss, resulting in the gradient embedding described in [2]:

$$g_{x_i}^c = p_i^c - \mathbf{1}(\hat{c}_i = c)) \cdot \ z_\phi(x_i) \quad (2)$$

where $c \in \{0, 1\}$ corresponding to 2 classes and p_i^c is the predicted probability for class c; $p_i^{c=0} = 1 - \hat{y}_i^{cls}$ and $p_i^{c=1} = \hat{y}_i^{cls}$.

Diversity. We create a batch of query items based on gradient embedding with k-means++ algorithm [1] in connection with the diversity. We obtain the set B_t^s of k centroids that are sampled with probability proportional to the nearest sets' distance to take diversity into account. Samples with small gradients are also unlikely to be chosen, as the distance between them is small. Thus, gradient embedding with *k*-means++ seeding tends to choose a large gradient sample diversely.

Scale the uncertainty and revenue effect. Furthermore, to induce the algorithm to select more uncertain and high-revenue items, we introduce extra weights to amplify the effect of uncertainty and revenue. The weights called uncertainty scale and revenue scale, adjust the probability of samples to be chosen by resizing their gradient embedding vectors.

Uncertainty scale. We magnify the impact of uncertain items by quantifying the model's ability to calibrate an item. We give each item an uncertainty score[3] (Eq. 3) such that the score implies the magnitude of model's uncertainty about the item. In this paper, we use a fraud score $\hat{y}_i^{cls}$ from the DATE model as a measure of item uncertainty. We give a higher scale of uncertainty unc_i if a fraud score $\hat{y}_i^{cls}$ is close to the midpoint 0.5, the uncertainty score unc_i is defined as:

$$unc_i = -1.8 \times |\hat{y}_i^{cls} - 0.5| + 1. \quad (3)$$

Revenue scale. Active learning in customs operation requires additional consideration, as revenue needs to be collected as customs duty. Maximizing customs duty is one of top priorities of customs authorities. Therefore, we further amplify the gradient embedding by the DATE model's predicted revenue $\hat{y}_i^{rev}$. The distribution of the amount of customs duty is right-skewed, so we take the log of the predicted revenue (Eq. 4). We can define a final scale factor S_i of $\mathbf{x}_i$ as

$$S_i = unc_i \cdot \log(\hat{y}_i^{rev} + k). \quad (4)$$

[3] A metric ranging in from 0.1 to 1. The smallest value is 0.1 (instead of 0) for computation stability.

As a result, gradient embeddings $g_{x_i}^c$ becomes

$$g_{x_i}^c = |S_i| \cdot (p_i^c - 1(\hat{c} = c)) \cdot \ z_\phi(x_i) \qquad (5)$$

k is a constant for computation stability. We name the algorithm covered so far as bATE, denoting the fusion of BADGE strategy [2] and DATE model [17].

Gatekeeping. In practice, some importers might commit fraud by analyzing and reverse engineering the model's prediction patterns. We can call them adaptive adversaries of the model. In this situation, randomness is known to improve the online algorithm's robustness and competitiveness [4, 5]. With this motivation, we additionally introduce randomness to our sampling strategy. Using a validation performance of the DATE model, we operate a gatekeeper. If the Rev@n% is higher than a predefined value of θ, the bATE exploration algorithm is used. Otherwise, if the DATE models' outputs are highly unreliable, those inputs can be considered an attack, thereby facilitating the random selection of items for inspection. To alleviate these discussed issues, we propose the final exploration strategy gATE. The gATE algorithm can be formally represented in Algorithm 2.

Algorithm 2 Exploring unknown items by gATE

Input: Current training set X_t, items received $\mathcal{B}_t$, inspection budget k_t
Output: A batch of selected items $\mathcal{B}_t^S$
Train the DATE model using training set X_t;
Get the Rev@n% from validation set;
if *Rev@n% > θ* **then**
  Perform prediction on $\mathcal{B}_t$, get the predicted annotation $(\mathbf{x}_i, \hat{y}_i^{cls}, \hat{y}_i^{rev})$ for each item $\mathbf{x}_i \in \mathcal{B}_t$;
  Calculate the gradient embedding g_{x_i} (Eq. 5);
  Obtain the set $\mathcal{B}_t^S$ of k_t items by k-means++ initialization;
else
  Obtain the set $\mathcal{B}_t^S$ of k_t items by random sampling;
end

3. Result and discussion:

We measure the performance of our proposed algorithm (bATE and gATE), along with that of other pure exploration strategies. This experimental setting is widely used in the active learning community [2, 18], for comparing performances between the static pool-based active learning algorithms. We performed experiments with four exploration strategies, including our proposed model designed in Section 2.

- Random [13]: Known to be used as an exploration strategy to detect novel frauds in some countries' production system.
- BADGE [2]: State-of-the-art active learning approach by selecting items considering uncertainty and diversity.
- bATE: Exploring by considering predicted revenue as well as item uncertainty and item diversity.

• gATE: Strategically decide the exploration strategy between random and bATE, depending on the base model's performance.

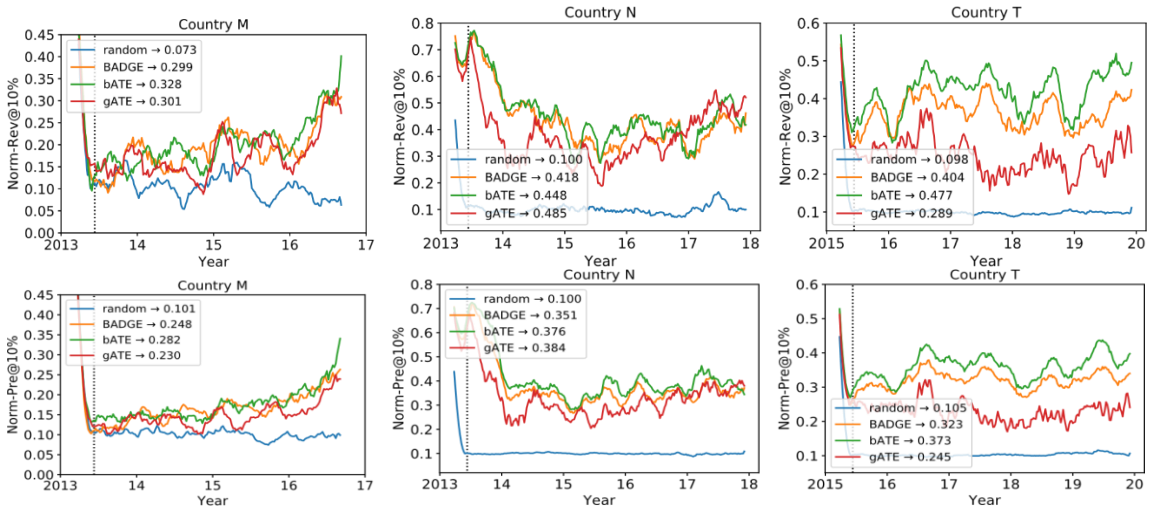


Figure 4: Performance of the advanced exploration strategy outperforms the random selection when only exploration strategy is used. Note that the random selection is widely used in many customs offices.

Figure 4 shows the 13-week moving average performance of exploration strategies. The result shows that three advanced strategies, BADGE, bATE, and gATE outperform random strategy in a large margin. The proposed models (bATE or gATE) outperformed the state-of-the-art active learning strategies BADGE by considerable margins (3-7pp improvement over BADGE). bATE is the top-performing strategy in country M and T, and gATE performs the best in country N.

Note that in this setting, evaluation is performed directly on the queried items. Explorative active learning strategies aim to find uncertain and informative items, while evaluation favors fraudulent and high-revenue items. Hence, we expect that the fully-exploration strategy's performance would not be impressive, suggesting the exploration-exploitation dilemma that will be discussed in part 5. This experimental setting may not be very realistic for advanced countries that maintain clean enough histories to train a model and would like to exploit their knowledge. However, this experiment is necessary for customs administration where there are not enough import histories available, so they want to fit the model as quickly as possible.

Part 4: Exploitative strategy and domain adaptation issue

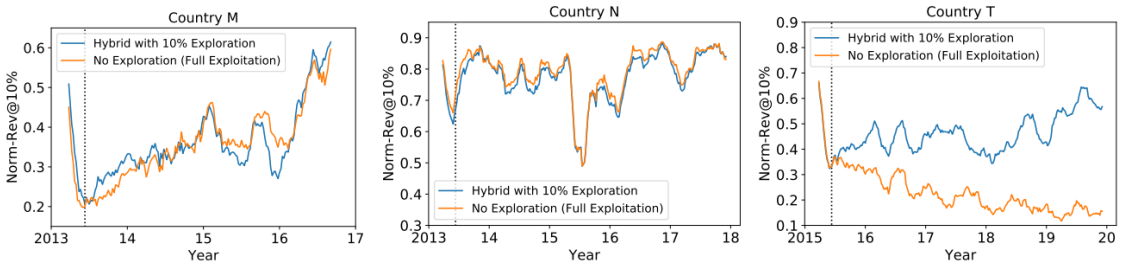
1. Exploitative strategy:

Algorithm 3 Exploiting suspicious items by DATE

Input: Training set X_t, items received $\mathcal{B}_t$, inspection budget k_t
Output: A batch of selected items $\mathcal{B}_t^S$
Train the DATE model using training set X_t;
Perform prediction on $\mathcal{B}_t$, get the predicted annotation $(\mathbf{x}_i, \hat{y}_i^{cls}, \hat{y}_i^{rev})$ for each item $\mathbf{x}_i \in \mathcal{B}_t$;
Obtain the set $\mathcal{B}_t^S$ of k_t items with highest fraud score $\hat{y}_i^{cls}$;

The state-of-the-art algorithm to detect illicit transactions and predict the raised revenue in a customs setting is the DATE model [17]. It is a tree-enhanced dual-attentive model to optimize dual objectives (illicit transaction classification and revenue prediction). We leverage the predicted fraud score of DATE for our exploitation strategy. We update the DATE model at each timestamp and select the most suspicious items as per the inspection budget (see Algorithm 3).

2. Result and discussion



(a) In country M, the performances of both strategies increase over time.

(b) In country N, and the performances of both strategies high and stable.

(c) In country T, the exploitation strategy failed, unlike a hybrid model.

Figure 5: For some cases, the performance of an exploitation strategy DATE drops as time goes, but the performance of hybrid strategies remains stable even when the exploitation strategy fails. This shows that exploration is necessary for maintaining a selection system in the long run.

We observe that a pure-exploitation strategy can crash sometimes. Figure 5(c) illustrates the customs simulation in country T. The performance of the state-of-the-art DATE exploitation strategy unexpectedly drops as time goes. It is worth noting that the performance drops significantly over time despite the training data's increasing size. This confirms that the items chosen for inspection are uninformative and evidence a domain shift in country T's trade pattern.

To back up the claim, we explored whether there exists a domain shift in the trading pattern of country T. Figure 6 describes the ratio of each import country for an item with commodity code starting with 620 in 2015, 2017, and 2019. There are

significant trade rates and domain shifts observed in the top three countries that imported items. Country A, B used to be where the item is imported the most, but starting from 2017, the shift in import countries sharply changes and country C becomes a dominant source country for imported goods.

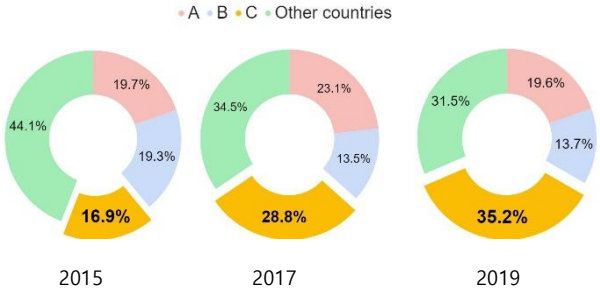


Figure 6: An example of domain shift in country T: The source country for commodity X (HS-code starting with 620) is rapidly changing over the years.

We check again to see if these behaviors are common across all datasets. Reassuringly, we also observe that the fully-exploitation strategy does not always fail. In Figure 4(a)–(b), we can see the results obtained from country M and country N. For those countries, maintaining the strategy of screening the most fraudulent items is still valid.

As suggested from the result, the exploitative model might fail, and exploration is needed to adapt to the domain shift. However, will exploration strategy helps, and in the case that exploitative works reasonably well (country M and N), will exploitation harm the performance? This will be discussed in part 5.

Part 5: Exploitation-exploration dilemma - Hybrid strategy

1. Hybrid strategy

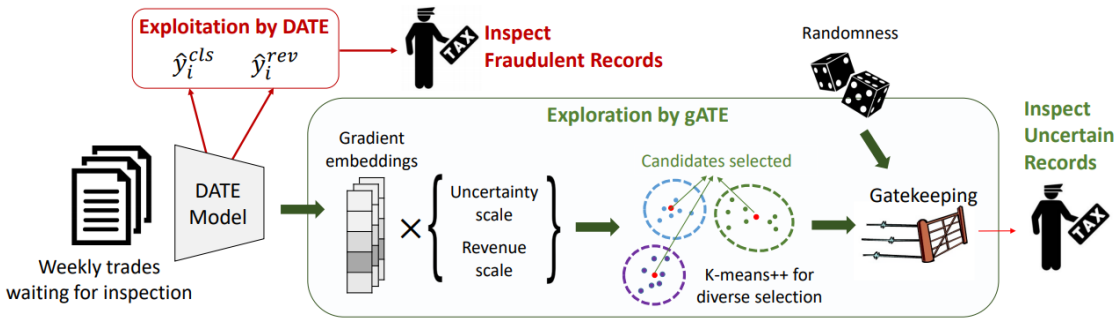


Figure 7: Illustration of the hybrid selection framework. The state-of-the-art exploitation model DATE first computes whether the input trade records are fraud or not. Then, the bATE algorithm computes back-propagation with pseudo-labels to generate a gradient embedding and re-scale it with its uncertainty and revenue. Then, *k*-means++ algorithm is applied to select diverse yet uncertain samples for inspection. Considering the DATE performance, a gating unit decides to use which items to explore.

The exploitation-only model can lead to confirmation bias. With a model trained only on the historical data and given the domain shift in customs datasets, the model tends to be unreliable from outliers. However, a pure exploration strategy cannot secure customs revenue and is unrealistic in the customs setting. Hence, we consider a balance between the two to achieve both short-term and long-term performance. We propose a hybrid selection strategy under the online active learning setting that includes two main approaches: exploitation and exploration by utilizing DATE and gATE.

Algorithm 4 Hybrid Selection using DATE and gATE

Input: Current training set X_t, items received $\mathcal{B}_t$, inspection budget k_t
Output: A batch of selected items $\mathcal{B}_t^S$
Train the DATE model using training set X_t;
Calculate the inspection budget $k_1 = r_1 \cdot |k_t|$, $k_2 = r_2 \cdot |k_t|$ for each strategy based on a predefined (normalized) ratio $r_1 : r_2$;
Obtain the set $\mathcal{B}_t^F$ of k_1 items by exploitation (DATE) strategy;
Obtain the set $\mathcal{B}_t^U$ of k_2 items by exploration (gATE) strategy;
Return the set $\mathcal{B}_t^S = \mathcal{B}_t^F \cup \mathcal{B}_t^U$

The exploitation approach selects the most likely fraudulent and highly profitable items to secure the short-term revenue for customs administration. On the other hand, the exploration approach selects uncertain items at the risk of an instant revenue loss, potentially detecting novel fraud patterns in the future. The algorithm operates by mixing these two components and aims to produce performance improvements in the long run on highly imbalanced customs

datasets with a small initial training set. Figure 7 illustrates the overall framework of the proposed model. This algorithm can be formally represented in Algorithm 4.

2. Result and discussion

To see the effect of exploration, we first compare the performance between the pure-exploitation strategy DATE and the partial-exploitation strategy with some random exploration; we can call it a naive hybrid strategy. The naive hybrid strategy uses DATE and random on a scale of 9 to 1. The result is shown in Figure 5 (see previous part). Surprisingly, in country T (where DATE performance drops over time), the hybrid strategy's performance remains stable. We deduce that the exploration strategy items significantly boost the performance of the exploitation strategy. Considering that the randomly selected items may only affect 1% of the total revenue on average, the performance boost arises from inspecting unknown items.

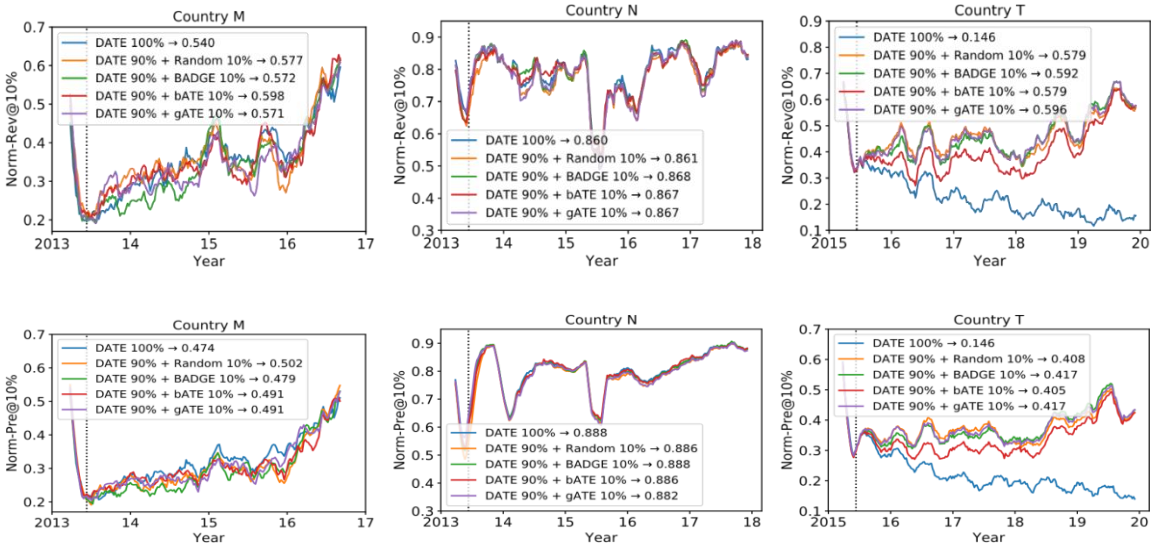


Figure 8: Combined with DATE, there is no significant difference between the performance of hybrid strategies.

However, when we compare the averaged performance between the exploitation strategy and the hybrid strategy, we can also find that the former does not beat the latter (Norm-Rev@10%, Exploitation vs. Hybrid: 54.0% vs. 57.7% for country M, 86.0% vs. 86.1% for country N; a moving average of the last 13 weeks). It is interesting to see that inspecting a set of random items is even better than inspecting reasonably fraudulent items with high $\hat{y}_i^{cls}$ values (top 9-10%) for maintaining a customs selection system in the long run.

So, how much will performance improve if the better exploration strategy is used rather than random? We measured the performance of hybrid with several exploration strategies. Each hybrid strategies select 90% of the item by DATE, and four exploration strategies select the remaining 10% of the item. We also compare them with DATE to show the long-term sustainability of hybrid strategies.

Figure 8 shows the performance of hybrid models with different exploration strategies. First, we can see that all hybrid strategies outperform a fully-exploitation strategy with some margin. For the T dataset, where a staggering decline of exploitation strategy performance is recorded, our hybrid strategy performs exceptionally stable, and the model is getting better towards the end. For the other two datasets, even though the DATE model for exploitation remains effective, the 10% trade-off for exploration does not hurt the overall performance but slightly outperforms the exploitation algorithm. This proves our initial claim that even if we give up some inspection of suspicious items, we can guarantee similar performance by learning new patterns from the unknown. Second, the hybrid model's performance with a random exploration strategy is comparable to the hybrid model with advanced exploration strategies.

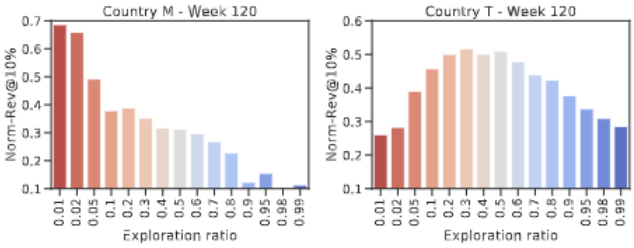
In practice, we encourage customs administration to start with a random strategy since it has almost no computation cost. The customs selection model will be improved even more robustly without using additional computing power. Third, advanced exploration strategies can help to raise the whole hybrid model to the best. The summary of model performance is shown in Table 3.

Table 3: Summary of the overall performance. The first number denotes Norm-Rev@10%, and the second number denotes Norm-Pre@10% of the model. The performance of the best performing model is emphasized in bold.

Model	Country M		Country N		Country T	
	Without DATE (Fully exploration)	With DATE (Hybrid)	Without DATE (Fully exploration)	With DATE (Hybrid)	Without DATE (Fully exploration)	With DATE (Hybrid)
gATE	0.301 / 0.230	0.571 / 0.491	0.485 / 0.384	0.867 / 0.882	0.289 / 0.245	**0.596 / 0.417**
bATE	0.328 / 0.282	**0.598 / 0.491**	0.448 / 0.376	0.867 / 0.886	0.477 / 0.373	0.579 / 0.405
BADGE	0.299 / 0.248	0.572 / 0.479	0.418 / 0.351	**0.868 / 0.888**	0.404 / 0.323	0.592 / 0.417
Random	0.073 / 0.101	0.577 / 0.502	0.100 / 0.100	0.861 / 0.886	0.098 / 0.105	0.579 / 0.408
DATE	0.540 / 0.474		0.860 / 0.888		0.146 / 0.146	

Part 6: Control the degree of exploitation-exploration

1. Weights sensitiveness



(a) The model performs the best with 1% of exploration.

(b) The model performs the best with 30% of exploration.

Figure 9: Best performing exploration ratio differs by data. If the exploitation strategy does not work well, increasing an exploration ratio helps (Country T).

The ratio between exploration and exploitation is set empirically as 0.1/0.9. The model performance is sensitive to this ratio, and the performance numbers vary depending on the dataset (Figure 9).

2. Adaptively determine the weights between exploration and exploitation

An adaptive algorithm to select this ratio will manage this trade-off better. RP1 algorithm [3] leverages an online learning mechanism with the exponential weight framework to dynamically tune this ratio, using Exponential-Weighted Framework. In this framework, we maintain a set of actions (arms) $A = (a_1, a_2, \dots, a_n)$, which essentially determines the weights of exploration. At the beginning of each timestep, the learner must choose an action a_t to minimize the cumulative regret. The regret of each timestep is defined as the loss difference between the actual action taken and the optimal action.

$$R = \sum_{t=0}^{T} l(a_t, B_t\ ) - \min_{a \in A} l(a, B_t)$$

Exponential-Weighted Framework maintains a probability distribution p_t over the set of actions A and use this distribution as a guide to sample the action. We will receive some feedback l(preferably for each action), and the distribution is updated by decreasing the weights of 'bad' actions and raises the weights of 'good' actions exponentially. If the loss is bounded by 1, it is proven that the regret will converge to 0 after a sufficiently large number of rounds [31].

Since we only get feedback for the taken action, we use an unbiased estimator [31]:

$$\hat{l}_t(a_i) = \frac{l_t(a_t)\mathbf{1}(a_i = a_t)}{p_{t-1}(a_t)}$$

The distribution is updated by

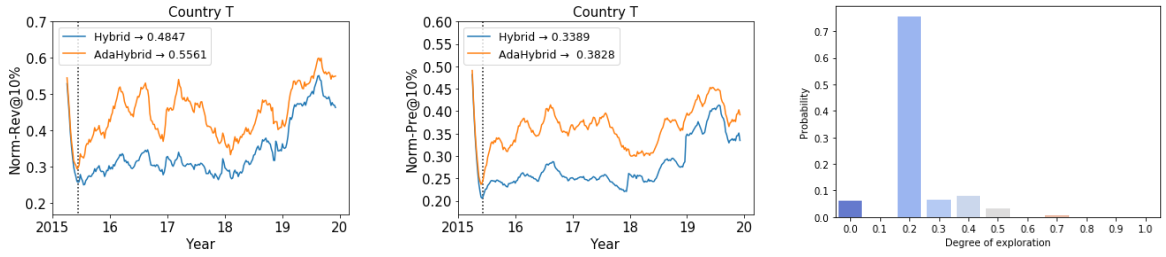
$$p_t(a_i) = p_{t-1}(a_i)e^{-\eta \hat{l}_t(a_i)}$$

where η is the learning rate.

To adapt to our setting, we use a 11-arm setting of weights from 0 to 1 (step size 0.1). The loss is the error rate (1 – accuracy) compared to the latest five rounds. The algorithm is called AdaHybrid algorithm (**hybrid** with **ada**ptively-adjusted weights)

3. Preliminary results:

We perform a simplified light-weight experiment to confirm that AdaHybrid works effectively. We use the hybrid algorithm on country T data, with exploitation by xgboost algorithm (partial part of DATE, faster training process) and exploration by random. The AdaHybrid reaches 56% on and 38% on precision, significantly outperform its non-adaptive counterpart by 7% on revenue and 4% on precision. We investigate the final guiding probability in a run. The distribution concentrates around the 0.2 – 0.5 range, consistent with the pattern shown in Figure 10. We confirm that the AdaHybird successfully guides the weight adaptation.



(a) Revenue performance (averaged) **(b) Precision performance (averaged)** **(c) Final guiding probability**

Figure 10: AdaHybrid outperforms its non-adaptive counterpart. The final guiding probability highly concentrates around the 0.2-0.5 region

Part 7: Conclusion and future work

This project discovers and analyzes the exploitation-exploration dilemma in the customs selection problem, where the indicators of annotated samples are key criteria for evaluation. One such example can be found in customs tax fraud detection, where customs officers need to decide which new cargo to inspect (i.e., an exploration strategy) while keeping the history of existing illicit trades (i.e., an exploitation strategy). We herein present an active learning-based model that efficiently combines the exploration and exploitation strategies. Our numerical evaluation, based on multi-year transaction logs, brings insights for practical guidelines for setting model parameters in the context of customs screening systems.

To facilitate the proposed approach in the customs administrations, the model code will be open-sourced. Currently, it supports diverse exploitation and exploration strategies with a variety of tunable parameters ranging from models to simulation settings. Users would be able to confirm whether our proposed work is well-suited for their data. There are some directions that this work can be extended.

- Reinforcement learning: The AdaHybrid algorithm preliminary results suggest that we can formulate an exploration-exploitation dilemma as a multi-armed bandit problem and apply reinforcement learning methods.
- Anomaly detection: We can consider the legit trade as normal data, and illicit trade as an anomaly. Although there are many anomaly detection techniques, more study is needed for adapting to non-image domains.
- Semi-supervised approach: Higher performance can be achieved by using richer information from a set of uninspected imports by incorporating a semi-supervised learning strategy in our framework. Building a set of augmented customs data and learning from it would be a key challenge for devising a semi-supervised learning model.

Part 8: Acknowledgement

This study is done under the Undergraduate Research Program at Korea Advanced Institute of Science and Technology, advised by Prof. Meeyoung Cha and researcher Sungwon Han. The study is part of the BACUDA project, cooperation for World Customs Organization and Institute of Basic Science, led by Prof. Meeyoung Cha and Dr. Sundong Kim. I would like to express my deepest gratitude towards all the collaborators in the project. I would like to thank all members in the research team who contribute greatly to this study's success, including Prof. Meeyoung Cha, Dr. Sundong Kim, Sungwon Han, Sungwon Park, Duc-KT-Nguyen, Jaechan So and Dr. Karandeep Singh. Part of this study is submitted as the paper "Take a Chance: Managing the Exploitation-Exploration Dilemma in Customs Fraud Detection via Online Active Learning."

Part 9: References

[1] David Arthur and Sergei Vassilvitskii. 2007. K-means++: The advantages of careful seeding. In SODA. 1027–1035.

[2] Jordan T. Ash, Chicheng Zhang, Akshay Krishnamurthy, John Langford, and Alekh Agarwal. 2020. Deep batch active learning by diverse, uncertain gradient lower bounds. In ICLR.

[3] Philip Ball, Jack Parker-Holder, Aldo Pacchiano, Krzysztof Choromanski, and Stephen Roberts. 2020. Ready Policy One: World Building Through Active Learning. In ICML.

[4] Niv Buchbinder, Kamal Jain, and Joseph Seffi Naor. 2007. Online primal-dual algorithms for maximizing ad-auctions revenue. In European Symposium on Algorithms. Springer, 253–264.

[5] Niv Buchbinder and Joseph Naor. 2009. The design of competitive online algorithms via a primal-dual approach. Now Publishers Inc.

[6] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A scalable tree boosting system. In KDD. 785–794.

[7] Yining Chen, Haipeng Luo, Tengyu Ma, and Chicheng Zhang. 2020. Active online domain adaptation. In ICML.

[8] Corinna Cortes, Giulia DeSalvo, Claudio Gentile, Mehryar Mohri, and Ningshan Zhang. 2019. Region-based active learning. In ICML.

[9] Corinna Cortes, Giulia DeSalvo, Claudio Gentile, Mehryar Mohri, and Ningshan Zhang. 2020. Adaptive region-based active learning. In ICML.

[10] Yarin Gal. 2016. Uncertainty in deep learning. Ph.D. Dissertation. University of Cambridge.

[11] Yarin Gal, Riashat Islam, and Zoubin Ghahramani. 2017. Deep bayesian active learning with image data. In ICML.

[12] Jacob Gildenblat. 2020. Overview of active learning for deep learning. https://jacobgil.github.io/deeplearning/activelearning. Accessed: 2020-08-12.

[13] C. Han and R. Ireland. 2014. Performance measurement of the KCS customs selectivity system. Risk Management 16, 8 (2014), 25–43.
[14] Neil Houlsby, Ferenc Huszár, Zoubin Ghahramani, and Máté Lengyel. 2011. Bayesian active learning for classification and preference learning. arXiv:1112.5745
[15] Sheng-Jun Huang, Rong Jin, and Zhi-Hua Zhou. 2014. Active learning by querying informative and representative examples. IEEE Transactions on Pattern Analysis and Machine Intelligence 36, 10 (2014), 1936–1949.
[16] Abbas Kazerouni. 2020. Active Learning for Skewed Data Sets. arXiv:2005.11442
[17] Sundong Kim, Yu-Che Tsai, Karandeep Sigh, Yeonsoo Choi, Etim Ibok, Cheng-Te Li, and Meeyoung Cha. 2020. DATE: Dual attentive tree-aware embedding for customs fraud detection. In KDD.
[18] Andreas Kirsch, Joost van Amersfoort, and Yarin Gal. 2019. BatchBALD: Efficient and diverse batch acquisition for deep bayesian active learning. In NeurIPS. [19] Yiğit Kültür and Mehmet Ufuk Çağlayan. 2017. Hybrid approaches for detecting credit card fraud. Expert Systems 34, 2 (2017), e12191.
[20] J.H. Moon, Debasmit Das, and Chun-Sing George Lee. 2020. Multi-step online unsupervised domain adaptation. In ICASSP.
[21] Michael R. Berthold Nicolas Cebron. 2008. Active learning for object classification: from exploration to exploitation. Data Mining and Knowledge Discovery 18 (2008), 283–299.
[22] Ievgen Redko, Emilie Morvant, Amaury Habrard, Marc Sebban, and Younès Bennani. 2019. Advances in Domain Adaptation Theory. Elsevier.
[23] Ozan Sener and Silvio Savarese. 2018. Active learning for convolutional neural networks: A core-Set approach. In ICLR. [24] Burr Settles. 2009. Active learning literature survey. Computer Sciences Technical Report 1648. University of Wisconsin–Madison.
[25] Hwanjun Song, Minseok Kim, Sundong Kim, and Jae-Gil Lee. 2020. Carpe diem, seize the samples uncertain "at the moment" for adaptive batch selection. In CIKM.
[26] Hwanjun Song, Sundong Kim, Minseok Kim, and Jae-Gil Lee. 2020. Ada-boundary: accelerating DNN training via adaptive boundary batch selection. Machine Learning (2020), 1–17.
[27] Jellis Vanhoeyveld, David Martens, and Bruno Peeters. 2020. Customs fraud detection: Assessing the value of behavioural and high-cardinality data under the imbalanced learning issue. Pattern Analysis and Applications 23 (2020), 1457– 1477.
[28] Donggeun Yoo and In So Kweon. 2019. Learning loss for active learning. In CVPR. 93–102.
[29] Yifan Zhang, Peilin Zhao, Jiezhang Cao, Wenye Ma, Junzhou Huang, Qingyao Wu, and Mingkui Tan. 2018. Online adaptive asymmetric active learning for budgeted imbalanced data. In KDD. 2768–2777.
[30] Fedor Zhdanov. 2019. Diverse mini-batch active learning. arXiv:1901.05954
[31] G´abor Lugosi. 2015. Online learning with structured experts–a biased survey. https://ocw.mit.edu/courses/mathematics/18-657-mathematics-of-machine-learning-fall-2015/lecture-notes/MIT18_657F15_L17.pdf. Accessed: 2020-11-12.