Dynamic classification of bank clients by the predictability of their transactional behavior*

Alexandra Bezbochina, Elizaveta Stavinova, Anton Kovantsev, and Petr Chunaev

National Center for Cognitive Research, ITMO University
Saint Petersburg 199034 Russia
aabezbochina@itmo.ru, stavinova@itmo.ru, ankovantcev@itmo.ru, chunaev@itmo.ru

Abstract. We propose a method for dynamic classification of bank clients by the predictability of their transactional behavior (with respect to the chosen prediction model, quality metric, and predictability measure). The method adopts incremental learning to perform client segmentation based on their predictability profiles and can be used by banks not only for determining predictable (and thus profitable, in a sense) clients currently but also for analyzing their dynamics during economical periods of different types. Our experiments show that (1) bank clients can be effectively divided into predictability classes dynamically, (2) the quality of prediction and classification models is significantly higher with the proposed incremental approach than without it, (3) clients have different transactional behavior in terms of predictability before and during the COVID-19 pandemics.

Keywords: Predictability · Incremental learning · Transactional data.

1 Introduction

Analyzing client's (especially transactional) behavior is highly demanded nowadays [3,10,23,24,26] and is related to different tasks—from the prediction of client's next purchase [24,26] to that of future client's location [16,23]. These studies are particularly motivated by the purposes of the company's marketing efficiency and risk management. For example, companies are interested in determining profitable clients [27,28] and understanding the general client's behavior with respect e.g. to demographics. This is known as client segmentation [1,4,17] and aims at increasing company's profitability. Furthermore, the COVID-19 pandemics and the related economical processes have emphasized once again the necessity to analyze client's transactional behavior in dynamics [3,10] so that a company can use the results to withstand future possible crises.

In this paper, we face the task of analyzing bank clients' transactional behavior in a dynamic manner in terms of *predictability*. This means that we study

* This research is financially supported by the Russian Science Foundation, Agreement 17-71-30029, with co-financing of Bank Saint Petersburg, Russia.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

how the behavior can be predicted by a chosen *prediction model*, *quality metric* and *predictability measure*, see [24]. However, we do not evaluate the predictability on the whole client set (as is usually performed) but distinguish and analyze *classes* of clients by means of their *predictability profiles*. One can consider predictable clients more profitable as they have stable patterns of behavior and thus a bank can make marketing activities for such clients less risky. The essential novelty of our study consists in that we adopt an *incremental approach* that allows solving the problem of such classification of clients dynamically and can be used by banks not only for determining predictable clients currently but also for analyzing their dynamics during economical periods of different types. To reduce computational time for predictability analysis when the number of bank clients is large, we also use an incremental *classifier* to divide clients by their predictability. One can use the classifier to estimate the client's predictability already without the necessity to perform actual prediction by the prediction model. To summarize, the impact of this study is as follows:

- we propose a new method for dynamic classification of bank clients by the predictability of their transactional behavior (with respect to the bank's chosen prediction model, quality metric, and predictability measure);
- we use the prediction and classification models based on Long Short-Term Memory (LSTM) network [20] that are adaptive in the sense that they exploit the incremental learning principle;
- we show in our experimental study of the proposed method that it is efficient in dividing bank clients into predictability profile classes dynamically and moreover the quality of prediction and classification models is significantly higher with the proposed incremental approach than without it (as e.g. in [24]) for economical periods of different types (in particular, before and during the COVID-19 pandemics).

The code, public datasets, and experimental results are given on GitHub[1].

2 Related work

The task considered in this paper is connected with several scientific topics such as predictability analysis, classification, clients' segmentation and incremental learning. That is why a review covering several related topics is provided below.

Predictability analysis The topic of predictability analysis can be divided into two sub-topics: the topic of *intrinsic* [7] predictability analysis, which is aimed at evaluating the prediction quality using only the data characteristics regardless of the prediction model, and *realized* [12] predictability analysis, that estimates the chosen model's predictive quality on a certain data.

The existing papers on the topic of predictability estimation can be classified according to the object of the research: *univariate time series, multivariate time*

[1] https://github.com/AlgoMathITMO/Dynamic-classifier

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

series, event sequences and *network links*. However, there is no single methodological approach even inside these groups as different researchers look at the predictability phenomenon from the positions of different science areas (information theory, dynamical systems theory, etc.).

Many works in the field are devoted to measuring the predictability of a univariate time series before the actual forecasting. The first realized predictability measure is proposed in [12], and it is further used in [18] as a base for a new measure of similar type. There is a significant group of intrinsic predictability estimation methods that exploit entropy evaluation: in [2] predictability measure based on permutation entropy is proposed, in [5] its weighted modification is presented, in [22] the Shannon entropy is used to estimate the categorical time series predictability, in [14] several time series features representing its predictability is analyzed, and, finally, in [13] time series predictability is explored via its transition graph analysis. As for the event sequences predictability estimation, to the best of our knowledge, there are only three works on this topic. In [8,11] the intrinsic predictability is estimated, while in [24] the realized one.

Client's predictability class identification To verify the novelty of the proposed method two papers should be regarded.

The first one is [24], where a method for client's predictability classification is proposed. There are several similarities between [24] and our paper: the experiments are conducted on transactional data, the object of the researches is categorical time series (or event sequences), the method for assessing the client's predictability class in the current is a generalization of the method from [24] on dynamic settings. Nevertheless, our paper considers the problem of estimating the clients predictability classes from a new perspective. Firstly, in our work the information about client's belonging to a predictability class in the sense of a single financial event is supplemented with the similar information corresponding to several events described by different financial categories. Secondly, our method uses incremental learning to dynamically estimate arriving data.

The second paper is [25], where a classification method for network links' predictability is introduced. The idea is that by using link features one can determine link predictability classes, i.e. [25] is rather close ideologically to [24].

Clients' segmentation There is also a group of works aimed at clients' segmentation (an unsupervised task, in opposite to classification considered in our paper) that is usually performed to obtain similar client groups according to some features [1,4,17]. Applying personalized marketing company for these groups makes it possible to better satisfy individual needs of clients. The important type of the clients' segmentation task is determining valuable clients. For example, [28] proposes a method for identification of valuable travellers. Also, purchasing behavior of valuable clients is considered in [27].

Incremental learning In classical scenarios of Machine Learning it is assumed that all the training data is available at the beginning of training but it is

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

not always true. In turn, incremental learning is used when the training data becomes available gradually. The goal of incremental learning is to adapt the learning model to new data (and not to forget the existing knowledge) without entire retraining the model.

The following groups can be distinguished among the neural network approaches to incremental learning. Regularization-based approaches [15] are usually based on imposing additional restrictions on changing the weights of the neural network when training on new data. Approaches based on dynamic architecture [19] are based on "freezing" the already trained neural network weights and adding new neurons for training on new data. Finally, learning with replay [21] involves saving old data in order to use not only new data, but also samples of old data for future training.

3 The description of the proposed method

Pipeline Suppose there is a set of bank clients $\{c_i\}_{i=1}^n$. Every client is represented by a set of transactions. Every transaction is given a MCC, or a Merchant Category Code, which is a 4-digit code representing a certain category of transactions. All MCC are grouped into N categories. Therefore, the spending of a client c_i during a period t can be defined as a vector of dimension N: $S_{i,t} = (m_{1,t}, \ldots, m_{N,t})_i$, where $m_{k,t} \in \{0, 1\}$ is the indicator of spending in category k. In this way we can define the history of ith client's transactions as the sequence of vectors $\{S_{i,t}\}_{t=1}^T$, where T is the total number of periods.

Having a set of bank clients $\{c_i\}_{i=1}^n$ presented by event sequences $\{S_{i,t}\}_{t=1}^T$, our task is to divide this set into predictability classes according to values of a chosen *quality metric* for a chosen *prediction model* and a certain category of transactions (at each time step). In this work, for simplicity we distinguish two classes of client's transactional predictability: namely, high (*Class A*, green plots in figures) and low (*Class B*, red plots in figures) predictability classes. The presence of a client in Class A means that his/her transactional behavior is predicted with higher quality than the behavior of any client from Class B.

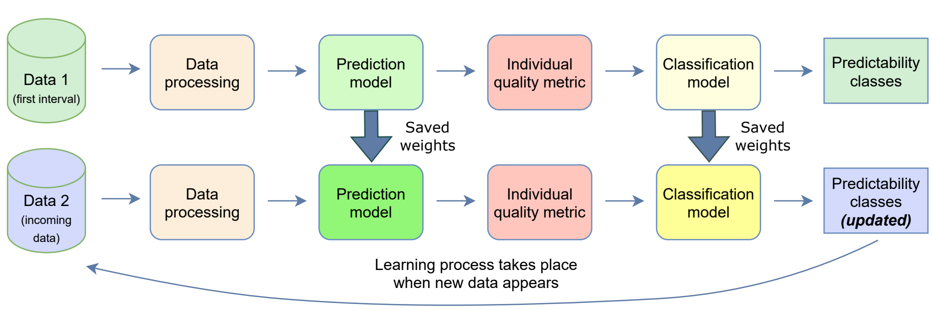


Fig. 1. The pipeline of the proposed method.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

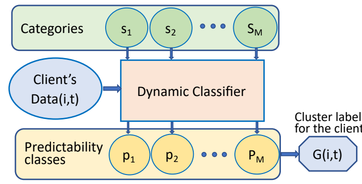


Fig. 2. The formation of client's predictability profiles.

We use the pipeline presented in Fig. 1 to perform the above-mentioned task. At the first stage, the input data is processed. We do some transformations on the history of transactions and extract vectors that act as an input of the prediction model. At the second stage, the prediction model is trained on the processed data and the probability of making a transaction in a particular category is predicted. Then we calculate the values of the quality metric at the next stage. Based on these values two classes are determined. Next, we train a *classification model* that identifies the classes without making the actual predictions.

The proposed pipeline is constructed in a dynamical manner and serves as a tool for maintaining arriving transactional data according to the principle of incremental learning. In this work we use a simple but efficient approach based on updating the model every time after new data arrival. Our approach can be classified as learning with replay because updating is performed using not only the new data, but also a sample of the already used data.

In the first step in Fig. 1, the prediction as well as the classification models are trained on the data that we have at a certain point in time. The model trained with the initial data is called a *base model* below while the dynamic one is called an *incremental model*. After training the base model, it is saved. At the second step, the saved model is loaded and updated using the new data that has arrived. The test and training samples are shifted by some interval every step.

Predictability profile analysis For a set of clients $\{c_i\}_{i=1}^n$, we further define a set of M categories $s_1, \ldots, s_M$, in terms of which we want to analyze the clients' behavior. Having the history of the ith client's transactions $\{S_{i,t}\}_{t=1}^T$, we are able to obtain the labels of predictability classes $(p_1, \ldots, p_M)_t$, $p_m \in \{0, 1\}$ for this client according to chosen categories using the predictability classifier.

We call the vector of predictability class labels $(p_1, \ldots, p_M)_t$, where $p_m \in \{0, 1\}$, a *predictability profile*. In fact, the vector can be interpreted as a predictability cluster label $G(i, t)$ for ith client in time period t, see Fig. 2. These clusters, their population and clients' transitions from one cluster to another characterize the clients' preferences along the time in a quite explainable way.

4 Experimental study

Now we aim to show the efficiency of the proposed method. It is important to check that the quality of prediction and classification models is higher within

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

the method than for non-updated models (as in [24]). We use transactional data for economical periods of different nature for this purpose.

Data description and processing For our experimental study we use two datasets: the first one (called $D1$ below) is public and provides the reproducibility of our results, while the second one (called $D2$ below) cannot be made public as provided by our commercial partner. Recall that the code, the public dataset $D1$ and the results for both $D1$ and $D2$ are available on GitHub, see Section 1.

The dataset $D1$ is from the Kaggle competition by the Raiffeisen bank (was publicly available at the time of competition). It represents the transaction history of 10,000 bank clients from January to December 2017. The data contains information for one year about the country, city, transaction address, date, amount of money spent and other values. MCC codes are categorized into $N = 87$ categories (in order to reduce the number of categories in the data).

The other dataset $D2$ contains client transactions for the period of one year from October 2019 till September 2020. Thus it contains the data of the first period caused by COVID-19 restrictions which were proclaimed between the 30th of March and the 11th of May 2020 in Russia. There are 11,166,746 records about 7,287 clients. The structure of $D2$ is the same as for $D1$.

Data preprocessing is identical for both datasets. First of all, corrupted or missing values are removed. Then the data are aggregated into categories by MCC codes and resampled to the weekly frequency with sum of transaction numbers for each client per week. Zero value is set if a client has no transaction during a certain week. After that the clients with less than one transaction in the period are excluded from the set. Thus we obtain a table with the columns of client's identifier, week number and amount of transaction in each category.

Prediction and classification models and their quality metrics The first model in the proposed pipeline (Fig. 1) is the prediction model which is aimed at the predicting the fact of transaction at a certain category. In this paper LSTM network [9] is used as the prediction model because of its advantage in remembering time dependencies [6,23,24]. Our predictive model consists of two layers of 64 LSTM-cells and one dense layer for output with the dimension of the number of categories. It takes the data according to the length of the training period and the number of categories in batches of 64 and returns the predicted probabilities of transaction for every client in each category in a certain period.

For this task, we form the following input vectors from the processed data: $A_{i,t} = (b_{1,t}, \ldots, b_{N,t})$, where $b_{k,t}$ is the number of transactions made by a client c_i during a period t ($t \in \{1, \ldots, T\}$). Here we choose the time step for t equal to one week. In this terms, the input for the prediction model can be formulated as $\{A_{i,t}\}_{t=1}^{\widetilde{T}}$, where $\widetilde{T}$ is a chosen length of the input sequence. The desired output for the prediction model is the indicator of the transaction made by a client c_i in category k at the next time step $(\widetilde{T}+1)$: $Q^i_{k,\widetilde{T}+1}$. The LSTM network used as a prediction model outputs the estimated probability of this event $\hat{Q}^i_{k,\widetilde{T}+1}$. For

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

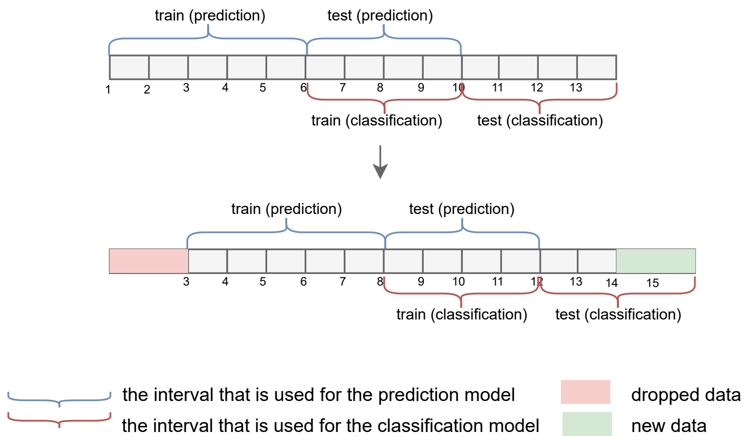


Fig. 3. Intervals for the chosen prediction and classification models.

our purposes, the category called Restaurants is selected as a target category in our experiments. This category represents visits to catering places (restaurants, cafes, fast food places, coffee houses, etc.)

Then the data is split into train and test periods according to the chosen threshold corresponding to a certain week. The data splitting scheme is presented in Fig. 3. Note that the prediction model allows to update its weights dynamically on data streams. With the new data arrival, the train and test periods are shifted to the number of weeks presented in the new data and the LSTM's weights are updated using the remaining old data and the new data. To estimate the quality of the proposed prediction model we use the Precision-Recall curves.

In our experiments we use a Bidirectional LSTM network [20] as a classification model. The *input* of used BiLSTM network is a set of categorical sequences consisting of the event indicators corresponding to a client c_i with the step of one week: $D_i = (d_1, \ldots, d_K)_i$, where d_l is the number of transactions in a chosen category, K is a chosen window length. The network consists of two layers: bidirectional LSTM of 20 cells and one output dense with two output cells for two predictability classes. The network *outputs* the predictability class p for each sequence from the input. The train period for the classification model coincides with the test period for the prediction model, and the test period for the classification model is the next K months. The details of the data splitting can be found in Fig. 3. The classification model also has the ability of dynamical learning on data streams, whose mechanism is the same as the prediction model has. To evaluate the quality of the classification model we use the ROC-AUC curves.

Predictability measure We use the sample *predictability rate* of an event in the next period from [24] as the *predictability measure*:

$$C(L, Q, m, i, k) = 1 - \frac{1}{L} \sum\nolimits_{j=1}^{L} |Q^i_{k,j} - \hat{Q}^i_{k,j}| \in [0, 1], \tag{1}$$

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

where L is the test period size, $Q^i_{k,j}$ is the actual event indicator, $\hat{Q}^i_{k,j}$ is the predicted probability of the event, m is a forecasting model, i corresponds to the ith client in a set $\{c_i\}_{i=1}^n$, k is an index of a chosen transaction category.

We compute the values of (1) for every client in the set $\{c_i\}_{i=1}^n$ to distinguish clients by predictability, thus, forming the set $\{C(L,Q,m,i,k)\}_{i=1}^n$. Namely, those clients for whom the values of C belong to $[0, median(\{C(L,Q,m,i,k)\}_{i=1}^n)]$ are in the class with low predictability (Class B), and vice versa (Class A).

Using the chosen prediction model, we perform the prediction of a transaction in the Restaurants category. Then, using the chosen quality metric we divide all bank clients into the predictability classes. Finally, using the trained classification model, we solve the problem of identifying the client's predictability class skipping the stage of using the prediction model. In the Fig. 4 (a)-(b) one can see Precision-Recall curves for predictability classes after the first step of training: classes obtained after the prediction model and classes obtained after the classification. In the case of the classification model's perfect quality, the left and right figures will be the same. But we can see that the current predictability classes obtained by the classification model have more similar quality between each other than in the case of the division by the prediction model. But still the division by the classification model saves the classes quality hierarchy.

Dynamic classifier analysis The length of our first dataset allows us to simulate the appearance of the new data and to train the model in nine steps. At every step the test and training samples are shifted by 2 weeks. Using this data, we can update the model's weights and the predictability classes labels.

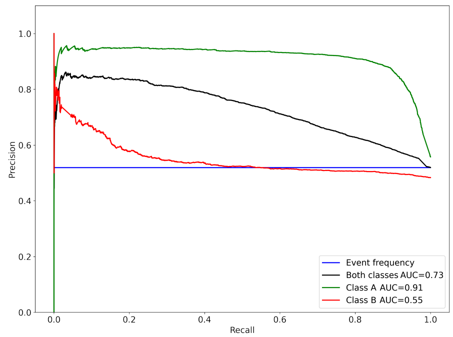
Since the proposed method assumes constant updating, we have the opportunity to evaluate the forecasting accuracy within several steps. Fig. 5 (a) shows the median of coefficient (1) dynamics of prediction model. When the base model is applied to new data, the prediction error stays approximately the same. On the contrary, with dynamic relearning, the error reduces sharply, then changes insignificantly. This shows that the dynamic classifier can distinguish the changes of arriving data distribution better than the model trained once.

In Fig. 4 (c)-(d) Precision-Recall curves for different predictability classes after the ninth step of training are presented: classes obtained after prediction model and classes obtained after classification. Comparing this figure with Fig. 4 (a)-(b), one can note that the quality of the prediction model (the left figures) has increased (from 0.73 to 0.77 in terms of Precision-Recall AUC). After the ninth training step the division into the predictability classes obtained by the prediction model is more contrast than it was earlier: two classes are further from each other. As for the classification model, it catches the division better after the whole training process as seen from the plots.

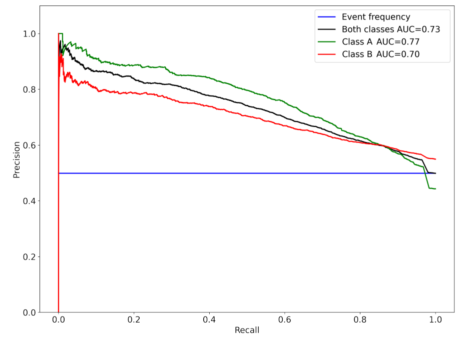
Fig. 5 (c) shows the comparison of the base and incremental models. While the accuracy of the base model decreases, incremental training allows to achieve a higher level of accuracy by constant updating.

We now apply our method to the dataset $D2$ that contains data for the period of COVID-19 restrictions. The quality scores for incremental learning and base

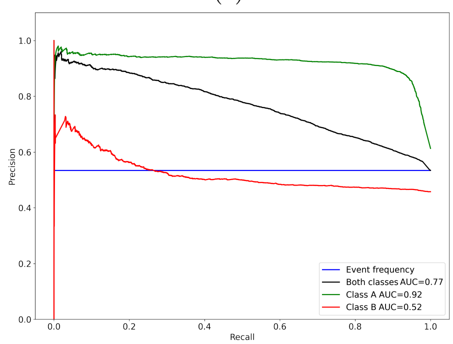
ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36



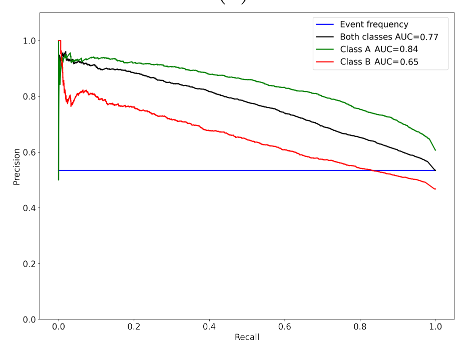
(a)



(b)



(c)



(d)

Fig. 4. Precision-recall curves for predictability classes on $D1$: (a) true classes (by the prediction model) after the first training, (b) the classes obtained by the classifier after the first training, (c) true classes (by the prediction model) after the ninth training, (d) the classes obtained by the classifier after the ninth training.

model are calculated according to (1). Their median values for each step are shown in Fig. 5 (b). One can notice that decreasing prediction quality happens after a time delay of three of four weeks after a critical transition has occurred. It was the last week of March 2020 when the restrictions were proclaimed, while the predictive quality fell by the end of April. But nevertheless, the incremental training model not only provides higher quality for each step of the process, but it can recover faster when the crisis is over. Furthermore, it is shown in Fig. 5 (d) how our models manage to overcome the difficulties caused by data volatility during Christmas holidays and COVID-19 restrictions.

Client predictability profile analysis In order to have the most versatile evaluation of client's predictability in our study we choose five transaction categories in different spheres of interest: restaurants and cafes, food stores, hairdressers and beauty salons, cosmetic stores, medical care. Fig. 6 shows an example of a predictability profile for one bank client. The first column of the

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

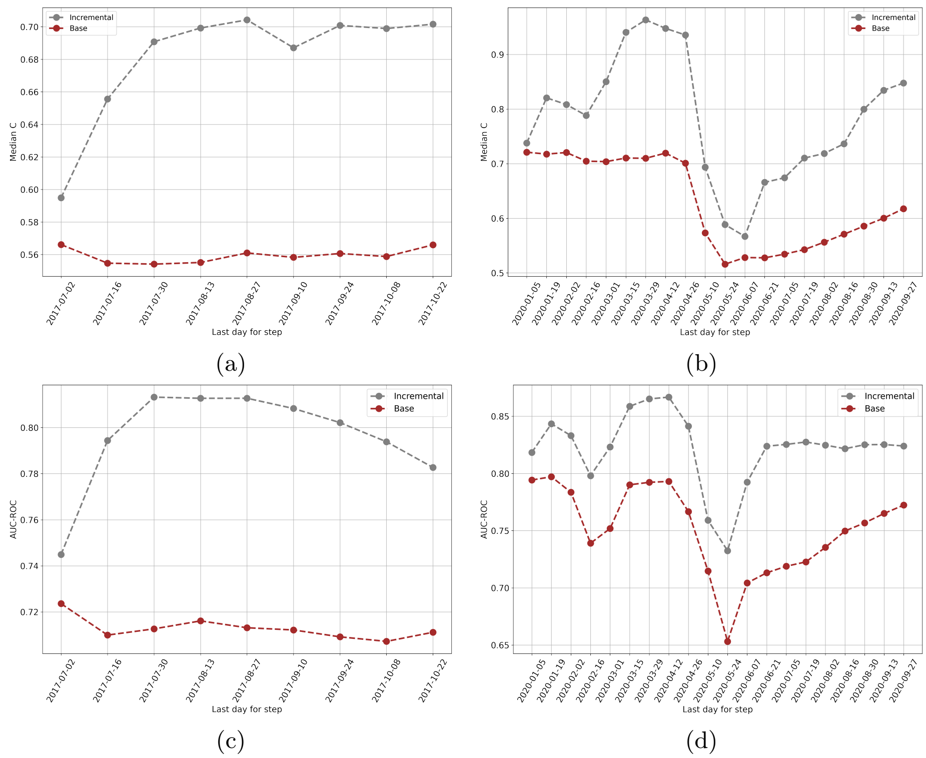


Fig. 5. The median C values for the prediction models: (a) $D1$, (b) $D2$; the ROC AUC values for the classification models: (c) $D1$, (d) $D2$.

predictability profile indicates the step of model training, while the other five columns represent the five chosen transaction categories. The values inside these five columns are binary; they indicate the predictability class at which a client belongs to (at a certain step of model training and a certain transaction category). For example, at the first step of training the model (Fig. 6) we can say that for two categories out of five, we can define the client's behaviour as "predictable". Over time, the predictability profile changes and by tracking it we can analyze client's behavior.

Let us say that the five client predictability classes together represent a binary number. When this number is converted to a decimal number system for each of the clients, we get 32 segments, or predictability clusters from 0 which means "00000" to 31 which means "11111". Then we can trace the changes in clients' behavior and their transit from one cluster to the other from step to step during the incremental learning process.

So, we can analyze the dynamic of group clients' behavior along the time and see how its predictability changes from step to step of our incremental process. In Fig. 7 we can see what was happening with predictability clusters in $D1$ with

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

step number	# 1	# 2	# 3	# 4	# 5
step 1	1	0	1	0	0
step 2	1	0	1	0	0
step 3	1	1	1	1	0
...					
step 9	1	1	1	1	0

Fig. 6. An example of a dynamic predictability profile for a client.

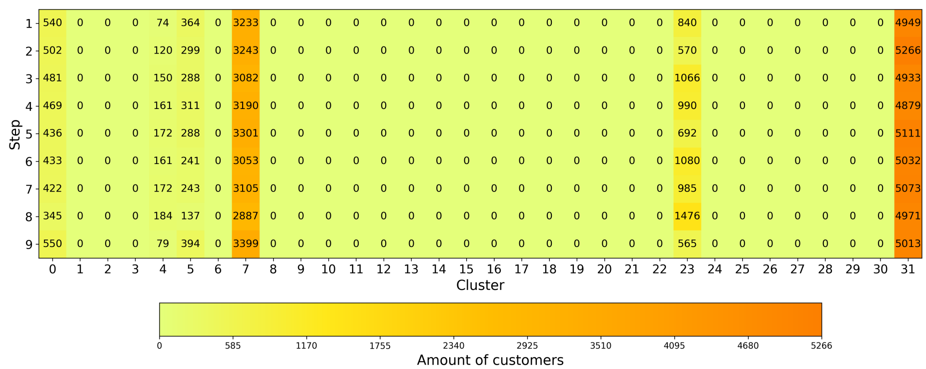


Fig. 7. Clients' transitions between predictability clusters for the $D1$ dataset.

more or less stationary data. The x-axis represents the number of cluster and the y-axis shows the course of time in terms of the model training step. The value inside the cells indicate the number of clients belonging to a cluster at a certain time step. Moreover, the cell colour highlights the most represented clusters. Analyzing Fig. 7, we can conclude that all clients are concentrated in clusters 0 ("00000"), 4 ("00100"), 5 ("00101"), 7 ("00111"), 23 ("10111"), 31 ("11111"). The most populated is the cluster where clients have good predictability, because it includes those who use their cards very rarely. Obviously, lack of transactions for a long time causes good predictions of no transaction in future.

The experiment with $D2$ set shows a bit different distribution of clients in predictability profiles. Most of them perform good predictability in every trade category so they belong to cluster 31. Many clusters are empty. Nonetheless, transitions between clusters happen on each step and the population of clusters never stays unaltered, as is seen in Fig. 8.

How the number of transitions from one cluster to the other changes in unstable situation we can see in Fig. 9 (b). It is on the increase when the critical period begins, then it falls down and recovers on the lower level after this period is finished. During the stationary period this amount keeps on a more or less sustainable level as Fig. 9 (a) illustrates. The analysis conducted in this section can be used as a tool for client profiling.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

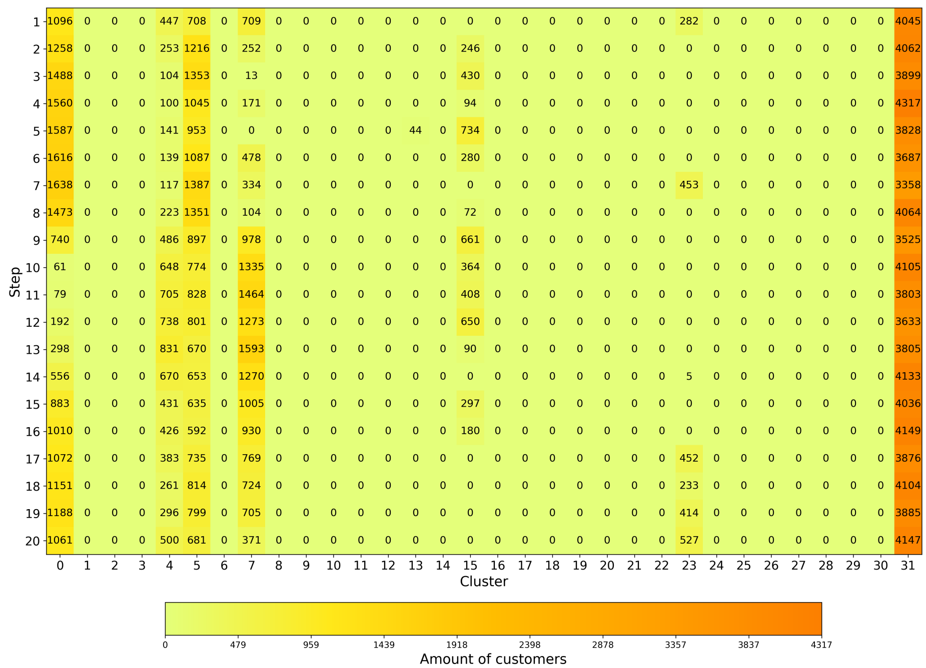
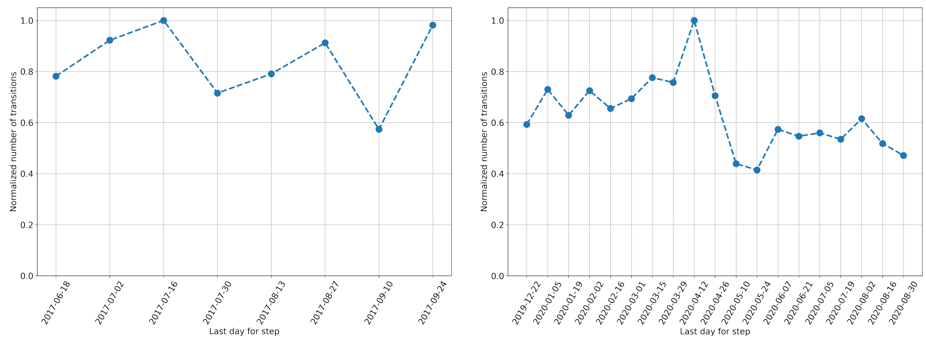


Fig. 8. Clients' transitions between predictability clusters for the $D2$ dataset.

5 Conclusions

We have proposed a method for dynamic classification of bank clients by the predictability of their transactional behavior under a certain choice of the prediction model, quality metric, and predictability measure. For the prediction and classification models, we used the LSTM network and its modification. Using the principle of incremental learning, we made the models dynamically updating on arriving data. After that, we conducted an experimental study of the method that showed the method's efficiency in dividing bank clients into predictability classes dynamically as the method's quality had been increased due to the model's updating on arriving data (ROC-AUC values 0.74 at the beginning of the learning process to 0.78 after the ninth step). Moreover, the proposed dynamic method has better classification quality than the non-adaptive models in the period of changes in the data distribution (for instance, caused by the lockdown due to the COVID-19 pandemic), since the ROC-AUC of the dynamic classifier is always higher than that of the non-adaptive model (Fig. 5). Finally, we formed a bank client's dynamic predictability profile showing the client's predictability in several categories. With the help of the profiles, we have got a tool that can be useful for dynamic analysis of clients' behavior in different spheres of interest. This tool allowed us to demonstrate noticeable changes in the client's transactional behavior during social and economic instability (Fig. 8).

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36



(a) (b)

Fig. 9. Normalized number of transitions between classes: (a) $D1$, (b) $D2$.

References

1. Bach, M.P., Jukovic, S., Dumicic, K., Sarlija, N.: Business client segmentation in banking using self-organizing maps. South East European Journal of Economics and Business (Online) **8**(2), 32 (2013)
2. Bandt, C., Pompe, B.: Permutation entropy: a natural complexity measure for time series. Physical review letters **88**(17), 174102 (2002)
3. Carvalho, V.M., Hansen, S., Ortiz, A., Garcia, J.R., Rodrigo, T., Rodriguez Mora, S., Ruiz de Aguirre, P.: Tracking the covid-19 crisis with high-resolution transaction data (2020)
4. Cuadros, A.J., Domínguez, V.E.: Customer segmentation model based on value generation for marketing strategies formulation. Estudios Gerenciales **30**(130), 25–30 (2014)
5. Fadlallah, B., Chen, B., Keil, A., Principe, J.: Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information. Physical Review E **87**(2), 022911 (2013)
6. Gers, F.A., Schraudolph, N.N., Schmidhuber, J.: Learning precise timing with lstm recurrent networks. Journal of machine learning research **3**(Aug), 115–143 (2002)
7. Granger, C., Newbold, P.: Forecasting Economic Time Series. Elsevier (1986)
8. Guo, J., Zhang, S., Zhu, J., Ni, R.: Measuring the gap between the maximum predictability and prediction accuracy of human mobility. IEEE Access **8**, 131859–131869 (2020)
9. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation **9**(8), 1735–1780 (1997)
10. Horvath, A., Kay, B.S., Wix, C.: The covid-19 shock and consumer credit: Evidence from credit card data. Available at SSRN 3613408 (2021)
11. Järv, P.: Predictability limits in session-based next item recommendation. In: Proceedings of the 13th ACM Conference on Recommender Systems. pp. 146–150 (2019)
12. Kaboudan, M.: A measure of time series' predictability using genetic programming applied to stock returns. Journal of Forecasting **18**(5), 345–357 (1999)
13. Kovantsev, A., Chunaev, P., Bochenina, K.: Evaluating time series predictability via transition graph analysis. In: 2021 International Conference on Data Mining Workshops (ICDMW). pp. 1039–1046 (2021)

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36

14. Kovantsev, A., Gladilin, P.: Analysis of multivariate time series predictability based on their features. In: 2020 International Conference on Data Mining Workshops (ICDMW). pp. 348–355. IEEE (2020)
15. Li, Z., Hoiem, D.: Learning without forgetting. IEEE transactions on pattern analysis and machine intelligence **40**(12), 2935–2947 (2017)
16. Moon, G., Hamm, J.: A large-scale study in predictability of daily activities and places. In: MobiCASE. pp. 86–97 (2016)
17. Panuš, J., Jonášová, H., Kantorová, K., Doležalová, M., Horáčková, K.: Customer segmentation utilization for differentiated approach. In: 2016 International Conference on Information and Digital Technologies (IDT). pp. 227–233. IEEE (2016)
18. Prelipcean, G., Popoviciu, N., Boscoianu, M.: The role of predictability of financial series in emerging market applications. In: Proceedings of the 9th WSEAS International Conference on Mathematics & Computers in Business and Economics (MCBE'80). pp. 203–208 (2008)
19. Rusu, A.A., Rabinowitz, N.C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R., Hadsell, R.: Progressive neural networks. arXiv preprint arXiv:1606.04671 (2016)
20. Schuster, M., Paliwal, K.K.: Bidirectional recurrent neural networks. IEEE transactions on Signal Processing **45**(11), 2673–2681 (1997)
21. Shin, H., Lee, J.K., Kim, J., Kim, J.: Continual learning with deep generative replay. arXiv preprint arXiv:1705.08690 (2017)
22. Song, C., Qu, Z., Blumm, N., Barabási, A.L.: Limits of predictability in human mobility. Science **327**(5968), 1018–1021 (2010)
23. Stavinova, E., Bochenina, K.: Forecasting of foreign trips by transactional data: a comparative study. Procedia Computer Science **156**, 225–234 (2019)
24. Stavinova, E., Bochenina, K., Chunaev, P.: Predictability classes for forecasting clients behavior by transactional data. In: International Conference on Computational Science. pp. 187–199. Springer (2021)
25. Stavinova, E., Evmenova, E., Antonov, A., Chunaev, P.: Link predictability classes in complex networks. In: International Conference on Complex Networks and Their Applications. pp. 376–387. Springer (2021)
26. Vaganov, D., Funkner, A., Kovalchuk, S., Guleva, V., Bochenina, K.: Forecasting purchase categories with transition graphs using financial and social data. In: International Conference on Social Informatics. pp. 439–454. Springer (2018)
27. Wong, E., Wei, Y.: Customer online shopping experience data analytics: Integrated customer segmentation and customised services prediction model. International Journal of Retail & Distribution Management (2018)
28. Wong, J.Y., Chen, H.J., Chung, P.H., Kao, N.C.: Identifying valuable travelers and their next foreign destination by the application of data mining techniques. Asia Pacific Journal of Tourism Research **11**(4), 355–373 (2006)

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_36