ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН
MINISTRY OF SCIENCE AND HIGHER EDUCATION OF THE REPUBLIC OF KAZAKHSTAN

ХАЛЫҚАРАЛЫҚ УНИВЕРСИТЕТІ
IT
МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ
INTERNATIONAL UNIVERSITY

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

2022 (3) 4

Қазан-желтоқсан

ISSN 2708–2032 (print)
ISSN 2708–2040 (online)

Халықаралық ақпараттық және коммуникациялық технологиялар журналы
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Меншіктенуші: «Халықаралық ақпараттық технологиялар университеті» АҚ (Алматы қ.).
Қазақстан Республикасы Ақпарат және әлеуметтік даму министрлігінің Ақпарат комитетінде – **20.02.2020** жылы берілген.
№ KZ82VPY00020475 мерзімдік басылым тіркеуіне қойылу туралы куәлік.
Тақырыптық бағыты: ақпараттық технологиялар, әлеуметтік-экономикалық жүйелерді дамытудағы цифрлық технологиялар, ақпараттық қауіпсіздік және коммуникациялық технологияларға арналған.
Мерзімділігі: жылына 4 рет.
Тиражы: 100 дана
Редакцияның мекенжайы: 050040, Алматы қ-сы, Манас к-сі, 34/1, 709-кабинет, тел: +7 (727) 244-51-09).
E-mail: ijict@iitu.edu.kz Журнал сайты: https://journal.iitu.edu.kz

Международный журнал информационных и коммуникационных технологий
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)

Собственник: АО «Международный университет информационных технологий» (г. Алматы).

Свидетельство о постановке на учет периодического печатного издания в Министерство информации и общественного развития Республики Казахстан **№ KZ82VPY00020475,** выданное от **20.02.2020 г.**

Тематическая направленность: информационные технологии, информационная безопасность и коммуникационные технологии, цифровые технологии в развитии социо-экономических систем**.**

Периодичность: 4 раза в год.

Тираж: 100 экземпляров.

Адрес редакции: 050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09).

E-mail: ijict@iitu.edu.kz
Сайт журнала: https://journal.iitu.edu.kz

«International Journal of Information and Communication Technologies»
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Owner: International Information Technology University JSC (Almaty).
The certificate of registration of a periodical printed publication in the Ministry of Information and Social Development of the Republic of Kazakhstan, Information Committee No. KZ82VPY00020475, issued on 20.02.2020.
Thematic focus: information technology, digital technologies in the development of socio-economic systems, information security and communication technologies
Periodicity: 4 times a year.
Circulation: 100 copies.
Editorial address: 050040. Manas st. 34/1, Almaty. +7 (727) 244-51-09). E-mail: ijict@iitu.edu.kz
Journal website: https://journal.iitu.edu.kz

МАЗМҰНЫ

БАҒДАРЛАМАЛЫҚ ҚАМТАМАСЫЗ ЕТУДІ ӘЗІРЛЕУ ЖӘНЕ БІЛІМ ИНЖЕНЕРИЯСЫ

ИНФОКОММУНИКАЦИЯЛЫҚ ЖЕЛІЛЕР ЖӘНЕ КИБЕРҚАУІПСІЗДІК

ЗИЯТКЕРЛІК ЖҮЙЕЛЕР

ЭКОНОМИКА ЖӘНЕ МЕНЕДЖМЕНТТЕГІ ЦИФРЛЫҚ ТЕХНОЛОГИЯЛАР

СОДЕРЖАНИЕ

РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНЖЕНЕРИЯ ЗНАНИЙ

ИНФОКОММУНИКАЦИОННЫЕ СЕТИ И КИБЕРБЕЗОПАСНОСТЬ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

ЦИФРОВЫЕ ТЕХНОЛОГИИ В РАЗВИТИИ СОЦИО-ЭКОНОМИЧЕСКИХ СИСТЕМ

CONTENTS

SOFTWARE DEVELOPMENT AND KNOWLEDGE ENGINEERING

INFOCOMMUNICATION NETWORKS AND CYBERSECURITY

INTELLIGENT SYSTEMS

DIGITAL TECHNOLOGIES IN THE DEVELOPMENT OF SOCIO-ECONOMIC SYSTEMS

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Vol. 3. Is. 4. Number 12 (2022). Pp. 57–69
Journal homepage: https://journal.iitu.edu.kz
https://doi.org/*10.54309/IJICT.2022.12.4.005*

UDC 004.42, 519.85

DETECTING CREDIT CARD FRAUD USING MACHINE LEARNING

R.Ye. Baitiles*, B.S. Omarov

Omarov Batyrkhan Sultanovich — PhD, assistant-professor of the «Mathematical Computer Modelling» department, International Information Technologies University
ORCID: 0000-0002-8341-7113. E-mail: rabinurye@gmail.com;
Baitiles Rabinur Yerkinkyzy — student/master student of the «Mathematical Computer Modelling» department, International Information Technologies University.

Abstract. Bank fraud is "The unauthorized use of an individual's confidential information to make purchases or withdraw funds from a user's account." E-commerce is growing rapidly, and the world is moving towards digitization, cashless transactions, the use of credit cards, the number of users is rapidly increasing, and with it the number of frauds associated with it. Due to the development of technology and the increase in the number of online transactions, fraud is also increasing, leading to huge financial losses. Therefore, effective methods to reduce losses are needed. In addition, scammers find ways to steal the user's credit card information by sending fake SMS and calls, as well as by masquerade attacks, phishing attacks, and so on. This article aims to use several machine learning algorithms such as Support Vector Machine (SVM), Decision Tree, Bayesian Belief Networks, Logistic Regression, k-Nearest Neighbor (Knn), and Artificial Neural Network (ANN) to predict the occurrence of fraud. In addition, we differentiate between the implemented supervised machine learning and deep learning methods to distinguish between fraudulent and non-fraudulent transactions.

Keywords: SVM, k-Nearest Neighbor, ANN, fraud detection, credit card, safety

For citation: R.Ye. Baitiles, B.S. Omarov. Detecting credit card fraud using machine learning // INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES. 2022. Vol. 3. Is. 4. Number 12. Pp. 57–69 (In Russ.). DOI: *10.54309/IJICT.2022.12.4.005*.

МАШИНАЛЫҚ ОҚЫТУ АРҚЫЛЫ НЕСИЕ КАРТАСЫНЫҢ АЛАЯҚТЫҒЫН АНЫҚТАУ

Р.Е. Байтілес, Б.С. Омаров*

Омаров Батырхан Сұлтанұлы — PhD докторы, Халықаралық ақпараттық технологиялар университетінің «Математикалық компьютерлік модельдеу» кафедрасының ассистенті
ORCID: 0000-0002-8341-7113. E-mail: rabinurye@gmail.com;
Байтілес Рабинұр Еркінқызы — Халықаралық ақпараттық технологиялар университеті «Математикалық компьютерлік модельдеу» кафедрасының студенті/магистрант.

Аннотация. Банктік алаяқтық — «Тұлғаның құпия ақпаратын сатып алу немесе пайдаланушының шотынан ақша алу үшін рұқсатсыз пайдалану». Электрондық коммерция қарқынды дамып келеді, ал әлем цифрландыруға, қолма-қол ақшасыз транзакцияларға, несиелік карталарды пайдалануға бет бұруда, пайдаланушылар саны және онымен байланысты алаяқтықтардың саны тез өсуде. Технологияның дамуына және онлайн транзакциялар санының артуына байланысты алаяқтық та өсіп, үлкен қаржылық шығындарға әкеледі. Сондықтан шығынды азайтудың тиімді әдістері қажет. Сонымен қатар, алаяқтар жалған SMS және қоңыраулар жіберу, сондай-ақ маскарадтық шабуылдар, фишингтік шабуылдар және т.б. арқылы пайдаланушының несие картасының ақпаратын ұрлау жолдарын табады. Бұл мақала алаяқтықтың пайда болуын болжау үшін Қолдау векторлық машинасы (SVM), Шешім ағашы, Байездік сенім желілері, логистикалық регрессия, k-ең жақын көрші (Knn) және жасанды нейрондық желі (ANN) сияқты бірнеше машиналық оқыту алгоритмдерін қолдануға бағытталған. Бұған қоса, біз жалған және алаяқтық емес транзакцияларды ажырату үшін енгізілген бақыланатын машиналық оқыту мен терең оқыту әдістерін ажыратамыз.

Түйін сөздер: SVM, k-En Nearest Neighbor, ANN, алаяқтықты анықтау, несие картасы, қауіпсіздік

Дәйексөз үшін: Р.Е. Байтілес, Б.С. Омаров. Машиналық оқыту арқылы несие картасының алаяқтығын анықтау // ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ-КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ. 2022. Том. 3. Is. 4. Нөмірі 12. 57–69 бет (орыс тілінде). DOI: 10.54309/IJICT.2022.12.4.005.

ВЫЯВЛЕНИЕ МОШЕННИЧЕСТВА С КРЕДИТНЫМИ КАРТАМИ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ

Р.Е. Байтілес, Б.С. Омаров*

Омаров Батырхан Султанович — к.т.н., доцент кафедры «Математическое компьютерное моделирование» Международного университета информационных технологий
ORCID: 0000-0002-8341-7113. E-mail: rabinurye@gmail.com;

Байтилес Рабинур Еркинкызы — студент/магистрант кафедры «Математическое компьютерное моделирование» Международного университета информационных технологий.

Аннотация. Банковское мошенничество — это «Несанкционированное использование конфиденциальной информации физического лица для совершения покупок или снятия средств со счета пользователя». Электронная коммерция быстро растет, и мир движется в сторону оцифровки, безналичных расчетов, использования кредитных карт, стремительно увеличивается количество пользователей, а вместе с ним и количество связанных с ним мошенничеств. В связи с развитием технологий и увеличением количества онлайн-транзакций увеличивается и мошенничество, приводящее к огромным финансовым потерям. Поэтому необходимы эффективные методы снижения потерь. Кроме того, мошенники находят способы украсть информацию о кредитной карте пользователя путем отправки поддельных SMS и звонков, а также с помощью маскарадных атак, фишинговых атак и так далее. Эта статья направлена на использование нескольких алгоритмов машинного обучения, таких как машина опорных векторов (SVM), дерево решений, байесовские сети доверия, логистическая регрессия, k-ближайший сосед (Knn) и искусственная нейронная сеть (ANN), для прогнозирования возникновения мошенничества. Кроме того, мы различаем реализованные контролируемые методы машинного обучения и методы глубокого обучения, чтобы различать мошеннические и немошеннические транзакции.

Ключевые слова: SVM, k-Nearest Neighbor, ANN, обнаружение мошенничества, кредитная карта, безопасность

Для цитирования: Р.Е. Байтілес, Б.С. Омаров. Выявление мошенничества с кредитными картами с использованием машинного обучения // МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ. 2022. Том. 3. Is. 4. Номер 12. Стр. 57–69 (на русском языке). DOI: 10.54309/IJICT.2022.12.4.005.

Introduction

As one of the most used financial products, the credit card is for making purchases such as gasoline, groceries, TVs, travel, shopping bills, etc. due to lack of funds at the moment. Credit cards are the most valuable because they provide different benefits when used for different types of transactions. Usually, large hotels, as well as various car rental companies, require the buyer to have a credit card.

A credit card usually refers to a card that is assigned to a customer (cardholder) and usually allows him to purchase goods and services up to a credit limit or withdraw cash in advance. A credit card gives the cardholder a time advantage, i.e., it gives its customers time to pay later at a set time by carrying it over to the next billing cycle.

Fraud is considered to be methods of obtaining money or services and goods in illegal or unethical ways.

The relevance of the topic - along with a high list of advantages, bank cards also have certain disadvantages, the most significant of which is their vulnerability to unauthorized influence by third parties in order to organize illegal access to the holder's account and subsequent theft of funds. The problem of ensuring the security of financial transactions using bank cards and, first of all, reducing the risk of fraud, is rightfully considered global, since all participants in the global payment instruments market are involved in the process of solving it.

The absence of a mechanism to prevent fraud using bank cards can potentially lead to the risks of the issuing bank associated with direct financial losses, deterioration of business reputation and distrust of the products provided by customers. Taking into account the rapid pace of development of the banking services market, solving the problem of ensuring the complexity and effectiveness of the measures taken to manage the risk of fraud in card transactions and operations across systems is a key aspect of the formation of a security policy, both at the level of an individual credit institution and throughout the banking system. The formation of a management system for these processes seems to be relevant both for commercial banks and for the Banks of Kazakhstan.

The scientific novelty of the study lies in the development of a classification of operational risks according to internal and external sources of their occurrence, indicating losses, as well as the development of guidelines to reduce the risk of fraud.

The purpose of the study is to scientifically substantiate and identify the most effective tools used by banks to reduce the risk of unauthorized (fraudulent) transactions using bank cards and systems, as well as to develop recommendations for managing risks when performing transactions using bank cards and systems.

To achieve our goals, consider the main tasks:

1) Expand the concept of operational risks, including cyber risks and develop a classification of operational risks, as well as study the history of the development of payment systems with a focus on reducing the risks of fraudulent transactions;

2) Identify mechanisms for organizing unauthorized access to the client's system and to card data by third parties for subsequent transactions without the client's consent;

3) Examine the current tools of the Bank of Kazakhstan and commercial banks to reduce the risk of fraud on bank cards and systems, as well as analyze data from the Bank of Kazakhstan on the dynamics of transactions without the consent of customers in recent years.

Credit card fraud is an easy target. Without any risks, a significant amount can be withdrawn without the knowledge of the owner, in a short time. Fraudsters are always trying to make every fraudulent transaction legitimate, which makes fraud detection a very difficult task. According to a 2020 report from the U.S. Payments Forum, criminals have shifted their focus to activities involving CNP transactions as chip card security has been enhanced.

Even then, thieves have a chance to misuse credit cards. There are many machine learning methods to solve this problem.

This article uses several machine learning algorithms such as support vector machine (SVM), decision tree, Bayesian belief networks, logistic regression, k-nearest neighbor

(Knn), and artificial neural network (ANN) to predict the occurrence of fraud. In addition, we distinguish between implemented supervised machine learning methods and deep learning methods to distinguish between fraudulent and non-fraudulent transactions.

Machine learning for fraud detection.

Machine learning is defined as a set of computer algorithms that make systems autonomously learn and produce results and improve them based on various analyzes and results. The data will be fed into these algorithms, which will automatically train them to perform a certain task, get a certain result, and therefore we can apply this to our real business scenarios. Machine learning algorithms can be used to solve business problems like regression, classification, prediction, clustering, associations, etc.

Based on the style and method used, machine learning algorithms are divided into four main types: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. In the following sections, let's take a closer look at each of the algorithms.

Supervised learning is a technique that involves learning using labeled past data and the algorithm must predict the label for unseen or future data. The supervised machine learning algorithm is actually told what to look for, and so on until it finds basic patterns that give the expected result with a satisfactory degree of accuracy. In other words, using these previously known results, the machine learning algorithm learns from the past data and then generates an equation for the label or value. This stage is called the learning stage.

The supervised learning algorithm has the following set of tasks — data collection, data preparation, modeling, model evaluation, deployment, and monitoring.

Gathering or data collection is the collection of relevant data required for a supervised learning algorithm. This data can be obtained through regular activities such as transactions, demographics, surveys, etc.

Data preparation is where we modify and transform the data using the necessary steps. It is critical to remove unnecessary data points and fill in inconsistencies in the data. This step ensures accuracy.

The simulation or training phase in which a link is established between the label and other variables.

During the evaluation phase, we check for errors and try to improve the model.

Deployment and monitoring occurs with invisible data, during the implementation of the model and the creation of forecast results.

Typical applications for supervised learning:

Image segmentation - various image classification activities are performed using the image data and the predefined labels we are looking for.

Medical diagnosis - using medical images and past tagged data that contain labels for disease states, we can identify disease for new patients.

Fraud detection − classification algorithms can be used to detect fraudulent transactions, fraudulent customers, etc. using historical data to identify patterns that could lead to possible fraud.

Spam detection - again, classification algorithms are used to classify email as safe or spam.

Speech recognition − the algorithm is trained using voice data and various identifications can be performed with it, such as voice passwords, voice commands, etc.

We have seen how supervised learning algorithms can help us predict value and event, various forms of supervised learning algorithms are designed to solve these business problems. Consider commonly used supervised learning algorithms such as:

1. Decision tree

The decision tree algorithm classifies objects by answering "questions" about their attributes located at key points. Depending on the answer, one of the branches is selected, and so on until the "leaf" is reached — the final answer.

Decision tree applications include knowledge management platforms for customer service, predictive pricing, and product planning. Figure 1 - The figure below shows the algorithm of the decision tree [1].

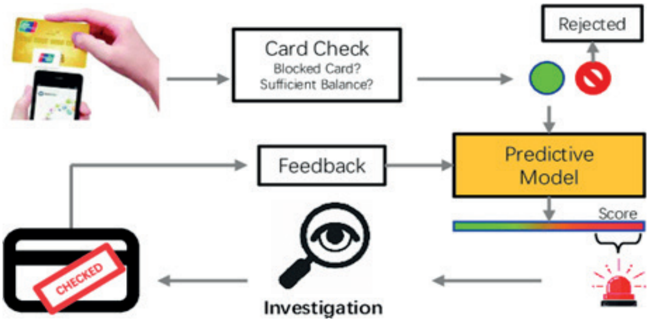

Figure 1- Algorithm of the decision tree

2. Bayesian Belief Networks

This method uses Bayes' theorem to calculate the probability of a hypothesis and determine whether it is true or false. The classifier is used to compute conditional probabilities for all possible classes and insert it into the class that has the highest conditional probability for a particular value of X. Also represents the case of a directed acyclic graph graphical model, with directed edges encoding probabilistic dependency relationships between variables.

3. Logistic Regression

Logistic regression is a type of multiple regression whose general purpose is to analyze the relationship between multiple independent variables (also called regressors or predictors) and a dependent variable. This is an appropriate method that can be used in predictive analysis when the dependent variable is dyadic or binary. Since the classification of transactions as fraud is a double-edged variable, this method can be used. This probability-based statistical classification model detects fraud using a logistic curve. Since the value of this logistic curve ranges from 0 to 1, it can be used to interpret the probabilities of belonging to a class. The figure-2 shows the algorithm of the Logistic Regression [2].

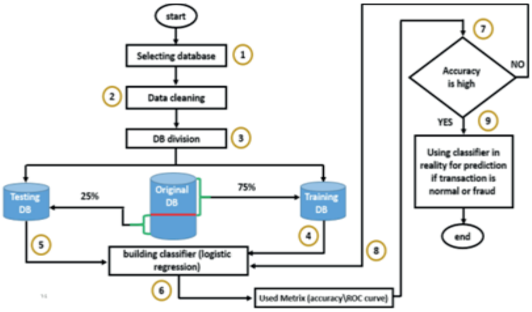
start
Selecting database
1
2
Data cleaning
DB division
3
Testing DB
25%
Original DB
75%
Training DB
4
5
building classifier (logistic regression)
6
Used Metrix (accuracy\ROC curve)
8
7
Accuracy is high
NO
YES
9
Using classifier in reality for prediction if transaction is normal or fraud
end

Figure 2 - algorithm of the Logistic Regression

4. SVM

The support vector machine (SVM) is one of the most popular learning methods used to solve classification and regression problems. The main idea of the method is to construct a hyperplane that separates the sample objects in an optimal way. The algorithm works on the assumption that the greater the distance (gap) between the separating hyperplane and the objects of separable classes, the smaller the average classifier error. In the Figure 3 illustrates support vector machine algorithm [3].

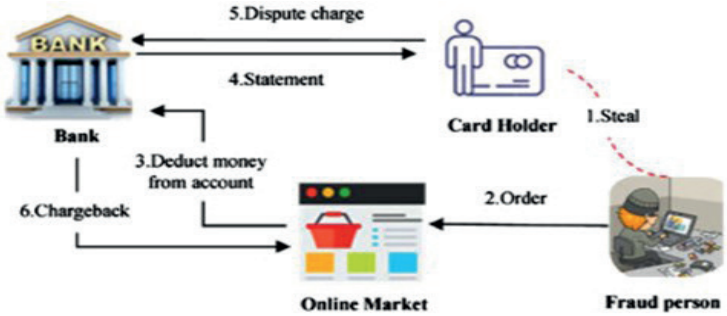

Figure 3 – SVM

5. K-Nearest Neighbors

The k-nearest neighbours method is a simple supervised machine learning algorithm that can be used to solve classification and regression problems. It is simple to implement and understand but has a significant drawback - a significant slowdown when the amount of data grows.

The algorithm finds the distances between the query and all examples in the data by choosing a certain number of examples (k) closest to the query, then votes for the most frequently occurring label (in the case of a classification problem) or averages the labels (in the case of a regression problem) [4].

KNN has a 97.69 % accuracy in detecting fraudulent card transactions. It showed suitable performance. It has been proven that KNN is effective for all existing indicators that are used in the classification and not a single false result has been recorded. Other evidence was executed using KNN, where punctuality of 72 % was achieved for credit card fraud [5].

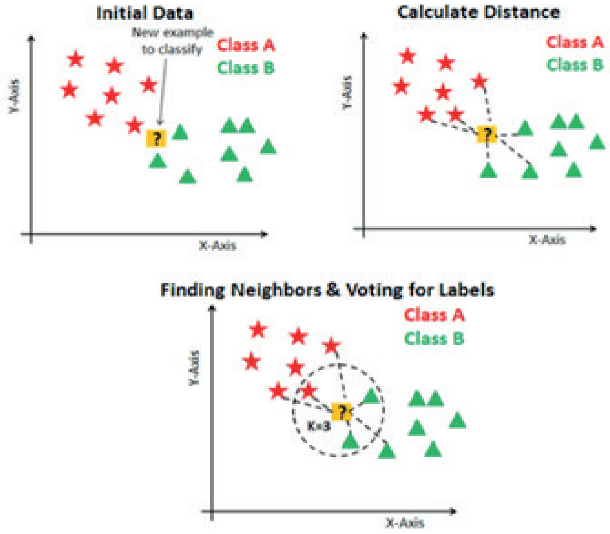

Figure 4 - K-Nearest Neighbours method

6. Artificial Neural Network (ANN) Method

ANN is a machine learning algorithm that functions like the human brain. In most cases, ANN is based on two kinds of methods: the supervised method and the unsupervised method. An unguided neural network is widely used to uncover fraud cases, since its accuracy is 95 % [6]. Basically, the unguided neural network tries to detect similar patterns among present credit card holders and those found in more early transactions. Let's assume that details seen in current transactions correlate with previous transactions. Then, most likely, the fraud incident will be revealed [6]. ANN methods are characterized by high fault tolerance. For example, the generation of output information is held up even if one or some cells are damaged. Due to its sublime speed and efficient processing capabilities, it is possible for ANN to compute a successful output for the sake of credit card fraud.

ANN appears to be a successful algorithm that is allowed to be used in credit card fraud. This can be seen from the literature that it has shown good performance in long-term use in overload with various functions and algorithms. These features have their own personal disadvantages. Also, the application of ANN in credit card fraud has become promising due to its ability to contain a high amount of granted and a texture of computed memory.

Classification of credit card frauds

The most popular types of fraud are:

- transfers to the accounts of fraudsters under their pressure;
- telephone fraud;
- Internet fraud;
- forgery of maps and websites;
- withdrawal of funds from a stolen or lost card.

Naturally, banks are aware of this and, for their part, are fighting scammers.

Constantly improving client identification methods, all kinds of passwords for Internet banks, PIN codes and CVV / CVC codes for cards, including dynamically changing OTP passwords for online transactions, and other degrees of protection.

Judicial practice makes it clear that most often in the dock are citizens who have chosen such a type of enrichment as theft of funds from other people's bank cards. Most often, scammers use the tricks of experienced fishermen, offering bank card holders an interesting “bait”. This technique is called "phishing". With its use, due to the inattention of citizens, there are tens of thousands of cases of fraud per year.

1. Phishing is a type of Internet fraud, the purpose of which is to obtain user identification data. This includes the theft of passwords, credit card numbers, bank accounts, and other sensitive information.
2. Skimming — this model of an attempt on the funds of individuals using the services of ATMs to withdraw cash, is the installation of a small device in an ATM that "collects cream" — copies all the data from the magnetic line of a bank card. Together with it, a device can be installed to remove information about the pin code when it is entered. Interestingly, such devices can also be installed in stores; they will copy information when buying goods with payment by bank card.
3. Telephone fraud - a type of fraud in the field of information technology, in particular, unauthorized actions and misuse of resources and services, theft of someone else's property or the acquisition of the right to someone else's property by entering, deleting, modifying information or otherwise interfering with the operation of data processing or transmission tools information and telecommunication networks.
4. Fraudulent call centers - fraudsters organize fraudulent call centers in which they call people to carry out mass illegal actions.
5. Fake website scam: A scammer injects malicious code that does its job on a website.
6. Lost/Stolen Card: This type of fraud involves the loss of the card by the cardholder or theft of the card from the cardholder.
7. Fake card scam: A type of scam where the scammer copies all the data from the magnetic strip and the real card looks like the original card and only works like the original card. This card is being used for fraud.

Rarely seen but bringing huge profits to scammers - a skimmer. A skimmer is a miniature portable reader that can be attached to an ATM. With the help of such devices, fraudsters steal bank card data: its details, PIN code, etc., in other words, all the information recorded on the magnetic strip. A skimmer can be a plastic pad attached to a card reader, a miniature video camera in a brochure holder next to an ATM. There are also special keyboard overlays that read the order of typing the PIN code. Skimmers are attached to ATMs using ordinary double-sided tape or Velcro fasteners. For example, if the keyboard was concave, then a special overlay will make the panel flatter. Also, the skimming device can change the keys themselves: they will either be recessed into the keyboard panel, or, conversely, protrude too much. In recent years, ATM manufacturers have begun to install special devices on ATMs that allow them to recognize skimmers. Finding a skimmer at an ATM is not easy, so bank employees recommend using only machines located in bank branches, large shopping centers, in a protected area. The skimmer can only steal information from the magnetic strip, not from the chip. For this reason, (and not only) chip cards are considered more secure.

There are also portable skimmers that allow you to make a copy of the card when it is in the hands of an attacker (for example, if he is also a part-time waiter in a restaurant where customers often pay with plastic cards).

While reading e-mail or browsing the Internet, you should be aware of scammers who seek to steal your personal data or money, and, as a rule, both. Such fraudulent activities or schemes are called “phishing” (from the English word “fish”, which means “fish” or “to fish”), since their goal is to “extract” your personal data from the bank card holder.

I would like to clarify another type of fraud called the Lebanese loop. "Lebanese loop" — for its application, a small piece of photographic film is used, which is folded in half, and the edges are bent at an angle of 90 degrees. This device is inserted into the ATM. The "highlight" is a small petal cut out on the underside of the film at a certain distance from the edge, bent up along the card. The film is located in the card reader so as not to interfere with the transaction. The bent petal does not allow the ATM to issue a plastic card back. That is, having completed the operation, the cardholder cannot get it back from the ATM. At this time, an “advisor” comes up, who recommends urgently going and calling the service department, for example. The owner of the card leaves, and in the meantime, the "adviser", who saw how he dialed the PIN code, pulls out the card and withdraws the money.

Also, no one is immune from the banal robbery of a bank card holder. This is the most uncomplicated method of the existing ones: the client withdrew cash — the swindler robbed.

The level of illegal transactions with bank cards in our country is lower than in developed countries. Among the types of fraud, skimming (illegal copying of card data using a special device) is still in the lead.

Ways of taking possession of other people's money, associated with deceit and abuse of trust or with the use of modern technical means, are diverse. You can protect yourself from them only with absolute care. Do not fall for the tricks of scammers.

Applying machine learning for fraud detection.

Table 1 - Application of machine learning techniques for fraud detection

Work	Technique used	Dataset used	Performance metrics	Result
[7]	Bayesian Neural Networks	PagSeguro (Brazil Online Payment Service)	HM between precision and Recall, and Economic Efficiency.	2 times improvement in performance
[8]	SVM, KNN, Logistic regression, Naive bayes	Real time data	Data and predictive analytics which is performed by ML models and an API module to detect the transaction is fraud or not.	Accuracy for SVM 91 %, KNN 72 %, LR 74 %, NB 83 %
[9]	Random forest, Adaboost algorithm	Kaggle	Accuracy, the confusion matrix is used to plot the ROC curve	Random forest has highest than adaboost algorithm

[10]	SVM, Artificial neural network	Provided by Serge Waterschoot at Europay International	TP, FP	Bayesian Belief, better than ANN. 8% more frauds detected. But ANN detects faster
[11]	Long-Short Term Memory	Dataset Recorded from March to May2015	Resistance to imbalance classes, attention to specific business interests	LSTM is more accurate than RF, improves personal transactions, and LSTM is prone to overfitting (layers have fewer nodes)
[12]	Big Data Analysis	German dataset	TP, FP	RF the decision tree performs best in terms of accuracy and precision among LR, DT and DTRF
[13]	Deep Learning	German Credit Data	Accuracy, Variance	High precision data processing
[14]	CNN	Credit card fraud data	TP-FP, FN-TN, Precision and Recall with their HM	With SMOTE, out-performs NN.
[15]	Deep Neural network	Real credit card data in the bank of US	AUC comparison	Better performance
[16]	Logistic regression, Naive Bayes, Random forest	Kaggle	Accuracy	Accuracy for LR 97.46 %, NB 99.23 %, Random forest 99.96 %
[17]	K-reverse nearest neighbor (KRNN)	dataset of European CC holders	Accuracy, problems of class imbalance that exists in the dataset	The findings showed that the RF method had a 91.24 % accuracy rate for fraud detection. The accuracy of the LR technique, in contrast, was 95.16 %.
[18]	KNN, random forest, LR, Svm	Uci library	Prediction of defaults	Assess the dataset in this study, then do feature selection and apply various machine learning methods.

Results

The use of information methods in the present period may give us more diverse data, while confidentiality remains an issue. And with our proposed method, we are given the opportunity to use sets of information in the present period to learn modification while maintaining confidentiality. A field of study with ANN could enhance the ability of the ML modification to detect fraudulent transactions. the recommended mixed move would be able to qualitatively correct the order of credit card fraud, using sets of true data, and uncover fresh interests in the field of banking and finance. The recommended method can support monetary institutions and banks to use the information sets in the system of the present period in a bilateral cooperation order, which will bring corporate benefits to develop a successful plastic card fraud system. Although the recommended method is effective in terms of credit card fraud when using bundles of information in a realistic period while maintaining confidentiality, it has limitations if the case is reported in this deployment. All banks and financial institutions have their own personal rules and regulations, and they are quite strict in the current regard. Adapting the recommended method will not be an easy task, because any bank and financial institution has its own limitations, and they rely on their own internal resources, and not on a general approach. although the originals are not transmitted centrally, and the learned pattern will learn patterns that can be decrypted by hackers. Therefore, while maintaining limitations, there is always work to be done to gain the confidence of banks and financial institutions for the introduction of this technology.

Conclusion

This article looks at the different methods used for the sake of credit card fraud. It can be analyzed that ML methods are a good method to increase the reliability of credit card fraud. However, we need huge sets of information to train the model, to eliminate the problem of data imbalance.

The purpose of this work was to study the classification of machine learning for the problem of detecting fraud with bank cards. As objects of study, in addition to ANN, several machine learning models were proposed for solving the problem, such as the model of the naive Bayes classifier on additions, random forest, etc.

In a further extension of this study, we will provide a conclusion after all this, how the prediction of fraudulent transactions classifies that we are given the opportunity to execute further. We also act with the information in the system of the present period, or we make a procedure for showing fraud in real time. We also apply a deep learning method for the best results. We are given the opportunity to create an application and a website that helps detect fraudulent transactions in real time.

From the literature, the authors mainly used an unbalanced set of information to test the correctness, truthfulness, and revocation of various machine learning algorithms in order to simulate a fraudulent transaction. But despite this, we will use selection methods for a balanced data set. For an unbalanced set of submissions, we cannot rely on accuracy, we should represent accuracy, completeness. From this evaluation, we can freely see which sample is most likely to function with an unbalanced dataset, also provide balanced datasets.

In accordance with the purpose, a review of the main ways of classifying fraud was carried out. Some existing systems and methods for detecting fraud are considered. The main complexity of solving this problem and the basic requirements for the fraud detection model are determined.

REFERENCES

Bharany S., Sharma S., Khalaf O.I., Abdulsahib G.M., Humaimeedy A.S., Aldhyani T.H.H., Maashi M., Alkahtani H. (2022). A Systematic Survey on Energy-Efficient Techniques in Sustainable Cloud Computing. Sustainability, 14, — 6256.

Cheng Dawei & Xiang Sheng & Shang Chencheng & Zhang Yiyi & Yang Fangzhou & Zhang Liqing.

(2020). Spatio-Temporal Attention-Based Neural Network for Credit Card Fraud Detection. Proceedings of the AAAI Conference on Artificial Intelligence. 34. — Pp. 362–369. 10.1609/aaai.v34i01.5371.

Cheng D., Xiang S., Shang C., Zhang Y., Yang F., & Zhang L. (2020). Spatio-temporal attention-based neural network for credit card fraud detection. — IEEE Access, 8, 135714–135724.

De Sa A., Pereira A. & Pappa G. (2018). A customized classification algorithm for credit card fraud detection. Retrieved from *https://arxiv.org/abs/1811.02810*

IRJET. (2019). Credit card fraud detection using machine learning algorithms. International Research Journal of Engineering and Technology (IRJET), 7(9). — Pp.167–171.

Itoo F. & Meenakshi S.S. (2020). Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. International Journal of Information Technology, — 13. — Pp.1503–151

Jurgovsky M., Granitzer M., Ziegler K., Calabretto S., Portier P.E., He L. & Caelen O. (2018). Sequence classification for credit card fraud detection. Procedia Computer Science, 126. — Pp. 201–210.

Kang F., Dawei C., Yi T. & Liqing Z. (2016). Credit card fraud detection using convolutional neural networks. In 2016 International Conference on Wavelet Analysis and Pattern Recognition. — Pp. 483–490. IEEE.

Mirtaheri M., Abu-El-Haija S., Morstatter F., Steeg G.V. & Galstyan A. (2021). Identifying and analyzing cryptocurrency manipulations in social media. IEEE Transactions on Computational Social Systems. — 8. — Pp. 607–617.

Ogwueleka F.N. (2011). Data mining application in credit card fraud detection system. Journal of Engineering and Applied Sciences, — 6. — Pp. 311–322.

Rajesh R. & Usha R. (2018). Fraud detection in credit cards using data analytics. Journal of advanced research in dynamical and control systems, 11(12). Pp. 65–74.

Roy A., Sun J., Mahoney R., Alonzi L., Adams S. & Beling P. (2018). Deep learning detecting fraud in credit card transactions. In Systems and Information Engineering Design Symposium (SIEDS). — Pp. 129–134. IEEE.

Thennakoon A., Bhagyani C., Premadasa S., Mihiranga S. & Kuruwitaarachchi N. (2019). Realtime credit card fraud detection using machine learning. In 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). — Pp. 1–6). IEEE.

Varmedja D., Karanovic M., Sladojevic S., Arsenovic M. & Anderla A. (2019). Credit card fraud detection-machine learning methods. In 2019 18th International Symposium INFOTEH-JAHORINA. — Pp. 1–6. IEEE.

Sailusha R., Gnaneswar V., Ramesh R.G. & Rao R. (2020). Credit card fraud detection using machine learning. In 2020 International Conference on Intelligent Computing and Control Systems (ICICCS). — Pp. 216–220. IEEE.

Suraj P., Varsha N. & Kumar S.P. (2018). Predictive modelling for credit card fraud detection using data analytics. Procedia Computer Science, — 132. — Pp. 385–395.

Saurabh A., Sushant B., Survesh S. & Vinay K.N. (2021). “Prediction of credit card defaults through data analysis and machine learning techniques”, scientific committee of the 1st International Conference on Computations in Materials and Applied Engineering.

Wang Y., Adams S., Beling P., Greenspan S., Rajagopalan S., Velez-Rojas M., ... & Boker S. (2018). Privacy preserving distributed deep learning and its application in credit card fraud detection. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. — Pp. 1070–1078).

ХАЛЫҚАРАЛЫҚ АҚПАРАТТЫҚ ЖӘНЕ КОММУНИКАЦИЯЛЫҚ ТЕХНОЛОГИЯЛАР ЖУРНАЛЫ

МЕЖДУНАРОДНЫЙ ЖУРНАЛ ИНФОРМАЦИОННЫХ И КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ

INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES

Правила оформления статьи для публикации в журнале на сайте:
https://journal.iitu.edu.kz
ISSN 2708–2032 (print)
ISSN 2708–2040 (online)
Собственник: АО «Международный университет информационных технологий» (Казахстан, Алматы)

ОТВЕТСТВЕННЫЙ РЕДАКТОР
Ералы Диана Русланқызы

КОМПЬЮТЕРНАЯ ВЕРСТКА
Жадыранова Гульнур Даутбековна

Подписано в печать 15.12.2022.
Формат 60х881/8. Бумага офсетная. Печать - ризограф.7,0 п.л. Тираж 100
050040 г. Алматы, ул. Манаса 34/1, каб. 709, тел: +7 (727) 244-51-09.

Издание Международного университета информационных технологий
Издательский центр КБТУ, Алматы, ул. Толе би, 59