Регион и мир, 2019, № 6

Մեքենայական ուսուցման մեթոդների կիրառումը քարտային զեղծարարությունների բացահայտման խնդրում

Սարգսյան Ա. Հ.
Հայաստանի Պետական Տնտեսագիտական Համալսարան (Հայաստան, Երևան)
and.sargsyan@yahoo.com

Վճռորոշ բառեր՝ մեքենայական ուսուցում, զեղծարարությունների բացահայտում, չբալանսավորված դասակարգում, լոգիստիկ ռեգրեսիա, նեյրոնային ցանց, գրադիենտ բուստինգ

Применение методов машинного обучения для обнаружения карточного мошенничества

Саргсян А. О.
Армянский государственный экономический университет (Армения, Ереван)
and.sargsyan@yahoo.com

Резюме: Объемы мошенничества с кредитными картами увеличиваются по мере роста числа пользователей банковских карт. Необходимо принять контрмеры для предотвращения подобных преступлений. Компании-эмитенты карт постоянно совершенствуют механизмы безопасности для защиты владельцев карт от различных видов мошенничества. Одним из таких механизмов является обнаружение мошеннических транзакций с помощью методов машинного обучения. В статье представлены результаты и сравнительный анализ различных моделей машинного обучения, созданных для обнаружения онлайн-мошенничества.
Ключевые слова: машинное обучение, выявление мошенничества, несбалансированная классификация, нейронная сеть, логистическая регрессия, градиент бустинг

Application of machine learning methods for credit card fraud detection

Sargsyan A. H.
Armenian State University of Economics (Armenia, Yerevan)
and.sargsyan@yahoo.com

Abstract: Credit card fraud volumes increase as the number of bank card users grows. It is necessary to take countermeasures to prevent such crimes. Card issuing companies continuously improve security mechanisms to protect cardholders from various types of fraud. One of these mechanisms is the discovery of fraudulent transactions through machine learning techniques. The article presents the results and comparative analysis of various machine learning models built for online fraud detection.
Keywords: Machine Learning, Fraud Detection, Unbalanced Classification, Neural Network, Logistic Regression, Gradient Boosting

Բանկային քարտեր օգտագործողների թվաքանակի աճին զուգընթաց աճում են նաև քարտային զեղծարարության ծավալները: Դա առաջ է բերում նման հանցագործությունների կանխման միջոցառումներ ձեռնարկելու անհրաժեշտություն: Քարտեր թողարկող ընկերությունները շարունակաբար կատարելագործում են քարտապաններին տարբեր տեսակի զեղծարարություններից պաշտպանելու անվտանգության մեխանիզմները: Այդ մեխանիզմներից մեկը մեքենայական ուսուցման մեթոդների օգտագործման միջոցով կեղծ գործարքների բացահայտումն է: Հոդվածում ներկայացվում են օնլայն քարտային զեղծարարությունների բացահայտման խնդրի համար կառուցված մեքենայական ուսուցման տարբեր մոդելների գրանցած արդյունքները և համեմատական վերլուծությունը:

Զեղծարարությունների գծով հավաստագրված փորձագետների ասոցիացիան (ACFE) զեղծարարությունը սահմանում է որպես կանխամտածված խաբեություն, որը իրականացվում է անձնական շահի կամ ուրիշ անձին վնասելու համար[1]:

Քարտային զեղծարարությունները լինում են 2 տիպի՝ օֆլայն և օնլայն: Օֆլայն զեղծարարությունը իրականացվում է գողացված ֆիզիկական քարտով խանութներում գնումներ կատարելու միջոցով: Շատ դեպքերում քարտը թողարկող բանկը կարող է արգելա-

[1] Տե՛ս ACFE European Fraud Conference, 2012

փակել քարտը մինչ դրանով զեղծարարություն իրականացվելը: Օնլայն զեղծարարությունները իրականացվում են ինտերնետային կամ հեռախոսային գնումների միջոցով, որի դեպքում միայն քարտի տվյալները բավարար են գործարք իրականացնելու համար[2]:

Օնլայն զեղծարարությունների կանխատեսման մոդելները կարելի է բաժանել 2 խմբի՝ մոդելներ, որոնք հիմնված են առանց հսկողության մեքենայական ուսուցման վրա (անոմալիայի բացահայտում), և մոդելներ, որոնք իրականացնում են կեղծ և ոչ կեղծ գործարքների դասակարգում (դասակարգման մոդելներ): Առաջինի առավելությունը կայանում է նրանում, որ ուսուցման համար հարկավոր չեն պիտակավորված տվյալներ, և մոդելը ունակ է բացահայտելու նաև այնպիսի կեղծիքներ, որոնք մինչև տվյալ գործարքի պահը դեռ չէին պատահել:

Դասակարգման համար գոյություն ունեն մեքենայական ուսուցման տարբեր ալգորիթմներ, որոնցից են՝ լոգիստիկ ռեգրեսիան, նեյրոնային ցանցերը, պատահական անտառները, որոշումների ծառերով գրադիենտ բուստինգը և այլն: Զեղծարարությունների բացահայտման խնդրի առանձնահատկություններից է հանդիսանում տվյալների ոչ բալանսավորված լինելը, քանի որ գրանցված գործարքների մեծ մասը ոչ կեղծ են, և չնչին մասն է, որ կեղծ է: Եթե հաշվի չառնենք այդ կարևոր փաստը, ապա ալգորիթմների մեծ մասը պարզապես բոլոր գործարքները համարելով ոչ կեղծ, շատ մեծ ճշգրտություն կապահովեն, սակայն օգտակար չեն լինի: Պարզելու համար, թե որ ալգորիթմն է ավելի ճշգրիտ արդյունքներ տալիս նշված խնդրի համար, կառուցվել են տարբեր մոդելներ և իրականացվել համեմատական վերլուծություն:

Մոդելների կառուցման և գնահատման համար անհրաժեշտ էր ունենալ գործարքների պատմության տվյալների բազա և յուրաքանչյուր գործարքի պիտակը՝ կեղծ կամ ոչ կեղծ: Ուսումնասիրության համար օգտագործվել է եվրոպական Worldline վճարահաշվարկային կազմակերպության կողմից հավաքագրված և Kaggle հարթակում հրապարակված գործարքների բազան: Այն բաղկացած է 2013 թվականի սեպտեմբեր ամսին եվրոպացի քարտապանների կողմից 2 օրվա ընթացքում իրականացված գործարքների պատմությունից: Բազայում ընդհանուր գրանցված է 284,807 գործարք, որից 492-ը կեղծ են: Տվյալները չափազանց ոչ բալանսավորված են. կեղծ գործարքները կազմում են ընդհանուր գործարքների ընդամենը 0.1727%-ը:

Քանի որ գործարքների պատմությունը կոնֆիդենցիալ տեղեկատվություն է, հրապարակված բազայում դրանք ներկայացված չեն այնպես, ինչպես գեներացվել են: Մեզ տրված է միայն PCA (Principal Component Analysis) ձևափոխության արդյունքում ստացված 28 կոմպոնենտ: Չեն ներկայացվում նաև այն հատկանիշները, որոնցից ստացվել են կոմպոնենտները: Միակ հատկանիշները, որոնք ձևափոխված չեն, առաջին գործարքի իրականացումից հետո անցած վայրկյանների քանակն է և գործարքի արժեքը: Ստացվում է, որ ընդհանուր տրված է 30 հատկանիշ և գործարքի պիտակը. եթե գործարքը կեղծ է՝ 1, հակառակ դեպքում՝ 0:

Հաջորդ քայլը մոդելների գնահատման համար ճիշտ չափորոշիչների ընտրությունն էր: Ճշգրտությունը, այսինքն՝ ճիշտ դասակարգված գործարքների քանակի հարաբերությունը թեստավորման ենթակա բոլոր գործարքների քանակին, լավ չափորոշիչ չէ այս խնդրի համար, քանի որ մոդելը կարող էր ապահովել բավականին բարձր ճշգրտություն բոլոր գործարքները պարզապես ոչ կեղծ համարելով, ինչը բխում է տվյալների չափազանց ոչ բալանսավորված լինելուց (մեր դեպքում ճշգրտությունը կկազմեր 99.83%): Գնահատման համար կօգագործենք հետևյալ գնահատականները.

- ճշգրիտ դրական կանխատեսումների քանակի և բոլոր դրական կանխատեսումների քանակի հարաբերությունը (precision)

$$precision = \frac{true\ positive}{true\ positive + false\ positive}$$

- ճշգրիտ դրական կանխատեսումների քանակի և բոլոր դրական (կեղծ) գործարքների հարաբերությունը (recall)

$$recall = \frac{true\ positive}{true\ positive + false\ negative}$$

- F1 գնահատական (F1 score), որը հաշվարկվում է որպես recall-ի և precision-ի միջին հարմոնիկ (լավագույն արժեքը՝ 1,

[2]Տե՛ս Yufeng Kou, Chang-Tien Lu, S. Sirwongwattana, Yo-Ping Huang (2004). Survey of fraud detection techniques

վատագույն արժեքը՝ 0)

$$F1 = \left(\frac{recall^{-1} + precision^{-1}}{2}\right)^{-1}$$

Մոդելի թեստավորումը չպետք է կատարել այն տվյալների վրա, որոնց վրա մոդելը սովորել է: Թեստավորման նպատակով օրիգինալ տվյալների բազայից առանձնացվել է բոլոր տվյալների 20%-ը պատահականության սկզբունքով, իսկ մնացած 80%-ը օգտագործվել է ուսուցման համար:

Համեմատության համար կառուցվել են լոգիստիկ ռեգրեսիայի, նեյրոնային ցանցի և գրադիենտ բուստինգի վրա հիմնված մոդելներ:

Լոգիստիկ ռեգրեսիա

Քարտային գեղծարարությունների կանխատեսման համար կօգտագործենք լոգիստիկ ռեգրեսիայի մոդելը, որը որպես մուտք ընդունում է x փոփոխականների վեկտոր և որպես ելք տալիս է [0, 1] միջակայքից իրական թիվ, որը մեր դեպքում ներկայացնում է հավանականությունը այն բանի, որ գործարքը կեղծ է: Եթե ելքային արժեքը մեծ է 0.5-ից, կասենք որ գործարքը կեղծ:

Մոդելը ներկայացվում է հետևյալ բանաձևով[3]

$$p(C_1 \mid x) = \sigma(\omega^T \cdot x)$$

որտեղ՝

$p(C_1 \mid x)$ - գործարքի կեղծ լինելու հավանականությունն է,

ω - մոդելի պարամետրերի վեկտորն է, որը պետք է գտնել ուսուցման միջոցով

σ - լոգիստիկ սիգմոիդ ֆունկցիան է (logistic sigmoid function)

$$\sigma(x) = \frac{1}{1 + exp(-x)}$$

Նեյրոնային ցանց

Նեյրոնային ցանցը կարելի է ներկայացնել որպես մուտքային տվյալների գծային և ոչ գծային ձևափոխությունների հաջորդականություն: Այն բաղկացած է մուտքի շերտից, ելքի շերտից և թաքնված շերտերից: Իսկ շերտերը բաղկացած են նեյրոններից, որոնք ֆունկցիաներ են և կատարում են իրենց մուտքային տվյալների նախ գծային, այնուհետև ոչ գծային ձևափոխություն, և որի ելքը կարող է մուտք հանդիսանալ մեկ ուրիշ նեյրոնի համար[4]: Ուսումնասիրության ենթակա ցանցը կազմված է 2 թաքնված շերտերից, որոնցից առաջինը ունի 32 նեյրոն, իսկ երկրորդը՝ 16: Ելքային շերտը բաղկացած է 1 նեյրոնից, որը ներկայացնում է հավանականությունը այն բանի, որ գործարքը կեղծ է: Թաքնված շերտերում որպես ակտիվացման (ոչ գծային) ֆունկցիա կիրառվել է ReLU-ն, իսկ ելքային շերտում՝ լոգիստիկ սիգմոիդը:

Որոշման ծառերով գրադիենտ բուստինգ

Գրադիենտ բուստինգը ռեգրեսիայի և դասակարգման համար նախատեսված մեքենայական ուսուցման ալգորիթմ է, որը հիմնված է թույլ կանխատեսման մոդելների համախմբման վրա[5]: Որպես թույլ մոդելներ սովորաբար օգտագործվում են որոշման ծառերը: Համեմատության համար կառուցված մոդելը բաղկացած է 1000 ծառերից, որոնցից յուրաքանչյուրի առավելագույն խորությունը 2 է: Սա հնարավորություն է տալիս խուսափել overfitting-ից:

Քանի որ ոչ կեղծ գործարքների տեսակարար կշիռը ավելի մեծ էր, ուսուցման ժամանակ կեղծ գործարքներին տրվել է ավելի մեծ կշիռ (կեղծ գործարքների համար՝ 400, ոչ կեղծ գործարքների համար՝ 1):

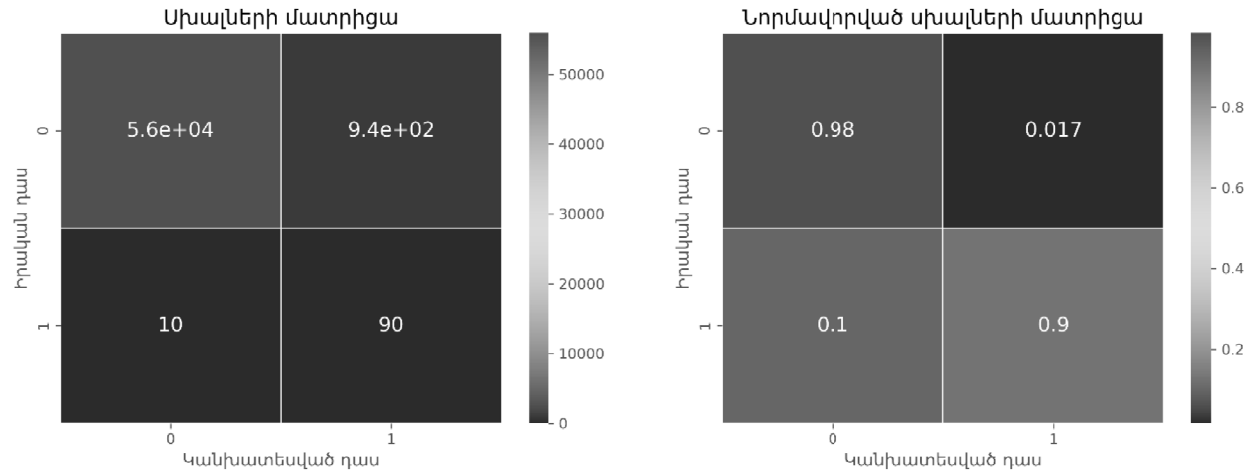
Գծապատկեր 1-ում պատկերված է կառուցված լոգիստիկ ռեգրեսիայի մոդելի սխալների մատրիցան և նորմավորված սխալների մատրիցան: Սխալների մատրիցան ցույց է տալիս, թե քանի ոչ կեղծ գործարքներ են սխալմամբ դասակարգվել որպես կեղծ (false positive, գծապատկերի վերին աջ անկյունում) և թե քանի գործարքներ են եղել կեղծ, սակայն մոդելի կողմից դասակարգվել որպես նորմալ (false negative, գծապատկերի ստորին ձախ անկյունում): Մատրիցան ցույց է տալիս նաև ճիշտ կանխատեսված կեղծ և ոչ կեղծ գործարքների քանակը (true positive և true negative): Նորմավորված սխալների մատրիցան ցույց է տալիս նույնը, ինչ սխալների մատրիցան, սակայն թեստային տվյալներում

[3] Տե՛ս Christopher M. Bishop (2006). Pattern Recognition and Machine Learning (Information Science and Statistics)

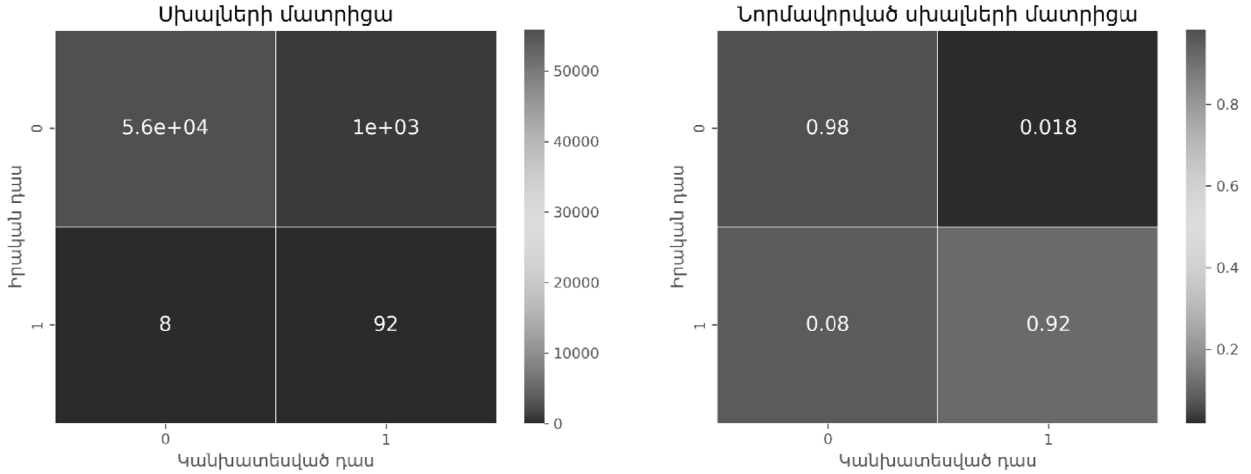
[4] Տե՛ս Hastie, Trevor, Robert, Tibshirani and J. H. Friedman (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer

[5] Տե՛ս J. Friedman (2001). Greedy function approximation: a gradient boosting machine. Annals of Statistics, 29(5):1189–1232

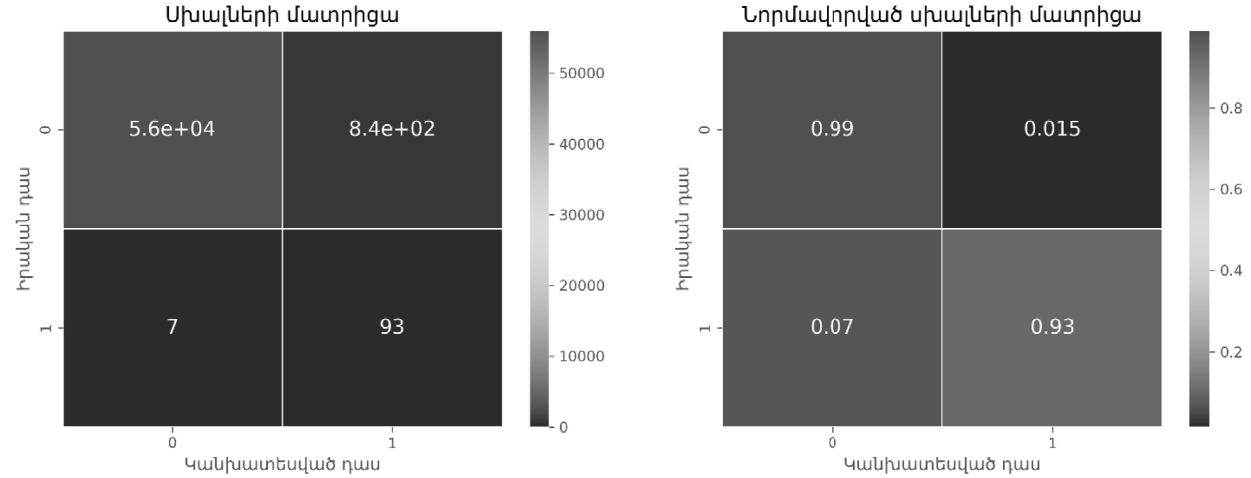
Գծապատկեր 1: *Լոգիստիկ ռեգրեսիայի մոդելի սխալների մատրիցան*



Գծապատկեր 2: *Որոշման ծառերով գրադիենտ բուստինգի մոդելի սխալների մատրիցան*



Գծապատկեր 3: *Նեյրոնային ցանցի մոդելի սխալների մատրիցան*



եղած կեղծ և ոչ կեղծ գործարքների նկատմամբ տոկոսային հարաբերակցությամբ:

Կառուցված նեյրոնային ցանցի սխալների մատրիցան ունի գրեթե նույն recall մակարդակը, ինչ գրադիենտ բուստինգի մոդելը, սակայն ունի շատ ավելի քիչ կեղծ ահազանգեր (գծապատկեր 3):

Աղյուսակ 1-ում ներկայացված են կառուցված մոդելների գնահատականները: Նեյրոնային ցանցի մոդելը ունի ամենաբարձր գնահա-

տականները թե՛ կեղծ գործարքների բացահայտմամբ, թե՛ կանխատեսումների ճշգրտությամբ։

Աղյուսակ 1: Լոգիստիկ ռեգրեսիայի, որոշման ծառերով գրադիենտ բուստինգի և նեյրոնային ցանցի արդյունքների համեմատությունը

Չափորոշիչ	Լոգիստիկ ռեգրեսիա	Որոշման ծառերով գրադիենտ բուստինգ	Նեյրոնային ցանց
precision	0.09	0.08	0.10
recall	0.90	0.92	0.93
F1-score	0.16	0.15	0.18

Բոլոր մոդելների դեպքում էլ կեղծ ահազանգերը ավելի շատ են, քան իրականները։ Դա կարելի է փոքրացնել՝ բարձրացնելով հավանականության այն շեմը, որն անցնելուց հետո գործարքը համարվում է կեղծ։ Սակայն այդ դեպքում կիջնի recall-ի ցուցանիշը, այսինքն՝ ավելի քիչ կեղծ գործարքներ կբացահայտվեն։

Այսպիսով՝ վճարահաշվարկային համակարգերում անվտանգություն ապահովելու համար կենտրոնական բանկերը և վճարահաշվարկային կազմակերպությունները կարող են մշակել ու ներդնել կեղծ և կասկածելի գործարքների բացահայտման համակարգ՝ հիմնված իրականացվող գործարքների վերլուծության վրա։ Նման համակարգ մշակելու դժվարություններից են հանդիսանում տվյալների ոչ բալանսավորված լինելը և պիտակավորված տվյալների ձեռք բերման հետ կապված դժվարությունները։ Համակարգի մշակման ժամանակ պետք է ուշադրություն դարձվի կեղծ գործարքների նկատմամբ մոդելի զգայունության աստիճանի որոշմանը։ Զեղծարարությունների բացահայտման համար ավելի նպատակահարմար է ունենալ բարձր զգայունություն և սխալ ահազանգերի առկայություն, քան կեղծ գործարքների չբացահայտում։ Տրված խնդրի համար ավելի նպատակահարմար էր օգտագործել նեյրոնային ցանցի մոդելը, որը բոլոր ցուցանիշներով գերազանցեց լոգիստիկ ռեգրեսիայի և գրադիենտ բուստինգի մոդելներին։ Հետագա ուսումնասիրություններում կարելի է համեմատել առանց հսկողության մեքենայական ուսուցման մեթոդներով կառուցված մոդելները դասակարգման վրա հիմնված մոդելների հետ։

Օգտագործված գրականության ցանկ

1. ACFE European Fraud Conference (2012)
2. Yufeng Kou, Chang-Tien Lu, S. Sirwongwattana, Yo-Ping Huang (2004). Survey of fraud detection techniques
3. Christopher M. Bishop (2006). Pattern Recognition and Machine Learning (Information Science and Statistics)
4. Hastie, Trevor, Robert, Tibshirani and J. H. Friedman (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer
5. J. Friedman (2001) . Greedy function approximation: a gradient boosting machine. Annals of Statistics, 29(5):1189–1232

* *Հոդվածում ներկայացված մոդելների իրականացումը կարող եք գտնել այստեղ՝ https://github.com/AndranikSargsyan/Fraud-detection*