БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет радиофизики и компьютерных технологий
Кафедра интеллектуальных систем

Аннотация к дипломной работе

Интеллектуальная система обнаружения мошенничества с кредитными картами

Пастушенко Антон Михайлович

Научный руководитель: кандидат физико-математических наук, доцент Е.И. Козлова

Минск, 2024

РЕФЕРАТ

Дипломная работа: 65 страниц, 21 рисунок, 22 источника.

BIG DATA ТЕХНОЛОГИИ, ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ, МАШИННОЕ ОБУЧЕНИЕ С УЧИТЕЛЕМ, БИНАРНАЯ КЛАССИФИКАЦИЯ.

Объект исследования – подходы и методы определения мошенничества с кредитными картами.

Цель работы – исследование в области организации систем сбора, хранения и обработки данных, с их последующей обработкой и анализом, в том числе систем машинного обучения, с анализом работы выбранных алгоритмов для решения задачи бинарной классификации банковских транзакций на факт мошенничества.

Методы исследования – компьютерное моделирование.

В работе рассматриваются основные подходы и методы, которые используются для обеспечения безопасных процессов пользования кредитными картами, их достоинства и недостатки. Рассматриваются существующие методы организации системы сбора, хранения и обработки данных. Ключевые компоненты данного рода систем. Разработана на основе рассмотренных компонентов собственная система сбора, хранения и обработки данных. Также исследованы основные алгоритмы машинного обучения для решения задачи бинарной классификации такие как линейная регрессия, деревья принятия решений, случайный лес и нейронная сеть прямого распространения (LSTM). Проведен интеллектуальный анализ данных и выявлены целевые признаки транзакций, влияющих на факт мошенничества. Проведен анализ результатов и выбор модели с наилучшим показателем точности в 79 %. Разработаны требования и описание к системе обнаружения мошенничества с кредитными картами. Реализованы основные ее компоненты.

Совокупность полученных результатов можно использовать как программный инструмент для внедрения в банковские системы работы с данными с целью определения потенциально мошеннических транзакций в реальном времени, анализа данных о транзакциях и дообучения алгоритма определения мошенничества на основе постоянно обновляющихся данных.

РЭФЕРАТ

Дыпломная праца: 65 старонак, 21 малюнак, 22 крыніцы.

BIG DATA ТЭХНАЛОГІІ, ІНТЭЛЕКТУАЛЬНЫ АНАЛІЗ ДАНЫХ, МАШЫННАЕ НАВУЧАННЕ З НАСТАЎНІКАМ, БІНАРНАЯ КЛАСІФІКАЦЫЯ.

Аб'ект даследавання - падыходы і метады вызначэння махлярства з крэдытнымі картамі.

Мэта працы – даследаванне ў галіне арганізацыі сістэм збору, захоўвання і апрацоўкі даных, з іх наступнай апрацоўкай і аналізам, у тым ліку сістэм машыннага навучання, з аналізам працы выбраных алгарытмаў для вырашэння задачы бінарнай класіфікацыі банкаўскіх транзакцый на факт махлярства.
Метады даследавання – камп'ютэрнае мадэляванне.

У працы разглядаюцца асноўныя падыходы і метады, якія выкарыстоўваюцца для забеспячэння бяспечных працэсаў карыстання крэдытнымі картамі, іх вартасці і недахопы. Разглядаюцца існуючыя метады арганізацыі сістэмы збору, захоўвання і апрацоўкі дадзеных. Ключавыя кампаненты дадзенага роду сістэм. На падставе разгледжаных кампанентаў распрацавана ўласная сістэма збору, захоўвання і апрацоўкі даных. Таксама даследаваны асноўныя алгарытмы машыннага навучання для вырашэння задачы бінарнай класіфікацыі такія як лінейная рэгрэсія, дрэвы прыняцця рашэнняў, выпадковы лес і нейронавая сетка прамога распаўсюджвання (LSTM). Праведзены інтэлектуальны аналіз дадзеных і выяўлены мэтавыя прыкметы транзакцый, якія ўплываюць на факт махлярства. Праведзены аналіз вынікаў і выбар мадэлі з найлепшым паказчыкам дакладнасці ў 79 %. Распрацаваны патрабаванні і апісанне да сістэмы выяўлення махлярства з крэдытнымі картамі. Рэалізаваны асноўныя яе кампаненты.

Сукупнасць атрыманых вынікаў можна выкарыстоўваць як праграмны інструмент для ўкаранення ў банкаўскія сістэмы працы з дадзенымі з мэтай вызначэння патэнцыйна ашуканскіх транзакцый у рэальным часе, аналізу дадзеных аб транзакцыях і данавучання алгарытму вызначэння махлярства на аснове пастаянна абнаўляючыхся дадзеных.

ABSTRACT

Diploma work: 65 pages, 21 drawings, 22 sources.

BIG DATA TECHNOLOGIES, DATA MINING, SUPERVIZED MACHINE LEARNING , BINARY CLASSIFICATION.

The object of the study is approaches and methods for determining credit card fraud.

The purpose of the work is to study the organization of data collection, storage and processing systems, with their subsequent processing and analysis, including machine learning systems, with an analysis of the operation of selected algorithms to solve the problem of binary classification of bank transactions for fraud.

The research methods are computer modeling.

The paper discusses the main approaches and methods that are used to ensure safe processes of using credit cards, their advantages and disadvantages. The existing methods of organizing the system of data collection, storage and processing are considered. The key components of this kind of systems. Based on the considered components, a proprietary data collection, storage and processing system has been developed. Basic machine learning algorithms for solving binary classification problems such as linear regression, decision trees, random forest and direct propagation neural network (LSTM) have also been studied. The data mining was carried out and the target signs of transactions affecting the fact of fraud were identified. The results were analyzed and the model with the best accuracy rate of 79% was selected. The requirements and description for the credit card fraud detection system have been developed. Its main components have been implemented.

The totality of the results obtained can be used as a software tool for implementing data management into banking systems in order to identify potentially fraudulent transactions in real time, analyze transaction data and train an algorithm for detecting fraud based on constantly updated data.