ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ
АКАДЕМИЯСЫ» РҚБ

ҚАЙЫРЫМДЫЛЫҚ ҚОРЫ
HALYK
CHARITY FOUNDATION

«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ
«ХАЛЫҚ» ЖҚ

ХАБАРЛАРЫ

ИЗВЕСТИЯ

РОО «НАЦИОНАЛЬНОЙ
АКАДЕМИИ НАУК РЕСПУБЛИКИ
КАЗАХСТАН»
ЧФ «Халық»

N E W S

OF THE ACADEMY OF SCIENCES
OF THE REPUBLIC OF
KAZAKHSTAN
«Halyk» Private Foundation

SERIES
PHYSICS AND INFORMATION TECHNOLOGY

3 (347)

JULY – SEPTEMBER 2023

PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

ЧФ «ХАЛЫҚ»

В 2016 году для развития и улучшения качества жизни казахстанцев был создан частный Благотворительный фонд «Халык». За годы своей деятельности на реализацию благотворительных проектов в областях образования и науки, социальной защиты, культуры, здравоохранения и спорта, Фонд выделил более 45 миллиардов тенге.

Особое внимание Благотворительный фонд «Халык» уделяет образовательным программам, считая это направление одним из ключевых в своей деятельности. Оказывая поддержку отечественному образованию, Фонд вносит свой посильный вклад в развитие качественного образования в Казахстане. Тем самым способствуя росту числа людей, способных менять жизнь в стране к лучшему – профессионалов в различных сферах, потенциальных лидеров и «великих умов». Одной из значимых инициатив фонда «Халык» в образовательной сфере стал проект Ozgeris powered by Halyk Fund – первый в стране бизнес-инкубатор для учащихся 9-11 классов, который помогает развивать необходимые в современном мире предпринимательские навыки. Так, на содействие малому бизнесу школьников было выделено более 200 грантов. Для поддержки талантливых и мотивированных детей Фонд неоднократно выделял гранты на обучение в Международной школе «Мирас» и в Astana IT University, а также помог казахстанским школьникам принять участие в престижном конкурсе «USTEM Robotics» в США. Авторские работы в рамках проекта «Тәлімгер», которому Фонд оказал поддержку, легли в основу учебной программы, учебников и учебно-методических книг по предмету «Основы предпринимательства и бизнеса», преподаваемого в 10-11 классах казахстанских школ и колледжей.

Помимо помощи школьникам, учащимся колледжей и студентам Фонд считает важным внести свой вклад в повышение квалификации педагогов, совершенствование их знаний и навыков, поскольку именно они являются проводниками знаний будущих поколений казахстанцев. При поддержке Фонда «Халык» в южной столице был организован ежегодный городской конкурс педагогов «Almaty Digital Ustaz.

Важной инициативой стал реализуемый проект по обучению основам финансовой грамотности преподавателей из восьми областей Казахстана, что должно оказать существенное влияние на воспитание финансовой грамотности и предпринимательского мышления у нового поколения граждан страны.

Необходимую помощь Фонд «Халык» оказывает и тем, кто особенно остро в ней нуждается. В рамках социальной защиты населения активно проводится работа по поддержке детей, оставшихся без родителей, детей и взрослых из социально уязвимых слоев населения, людей с ограниченными возможностями, а также обеспечению нуждающихся социальным жильем, строительству социально важных объектов, таких как детские сады, детские площадки и физкультурно-оздоровительные комплексы.

В копилку добрых дел Фонда «Халык» можно добавить оказание помощи детскому спорту, куда относится поддержка в развитии детского футбола и карате в нашей стране. Жизненно важную помощь Благотворительный фонд «Халык» оказал нашим соотечественникам во время недавней пандемии COVID-19. Тогда, в разгар тяжелой борьбы с коронавирусной инфекцией Фонд выделил свыше 11 миллиардов тенге на приобретение необходимого медицинского оборудования и дорогостоящих медицинских препаратов, автомобилей скорой медицинской помощи и средств защиты, адресную материальную помощь социально уязвимым слоям населения и денежные выплаты медицинским работникам.

В 2023 году наряду с другими проектами, нацеленными на повышение благосостояния казахстанских граждан Фонд решил уделить особое внимание науке, поскольку она является частью общественной культуры, а уровень ее развития определяет уровень развития государства.

Поддержка Фондом выпуска журналов Национальной Академии наук Республики Казахстан, которые входят в международные фонды Scopus и Wos и в которых публикуются статьи отечественных ученых, докторантов и магистрантов, а также научных сотрудников высших учебных заведений и научно-исследовательских институтов нашей страны является не менее значимым вкладом Фонда в развитие казахстанского общества.

С уважением,
Благотворительный Фонд «Халык»!

«ҚР ҰҒА Хабарлары. Физика және информатика сериясы».
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.). Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 14.02.2018 ж. берілген **No 16906-Ж** мерзімдік басылым тіркеуіне қойылу туралы куәлік.
Тақырыптық бағыты: *физика және ақпараттық коммуникациялық технологиялар сериясы.*
Қазіргі уақытта: *«ақпараттық технологиялар»бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*
Мерзімділігі: *жылына 4 рет.*
Тиражы: *300 дана.*
Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

«Известия НАН РК. Серия физика и информатики».
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Собственник: *Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).*
Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **No 16906-Ж** выданное 14.02.2018 г.
Тематическая направленность: *серия физика и информационные коммуникационные технологии.*
В настоящее время: *вошел в список журналов, рекомендованных ККСОН МОН РК по направлению «информационные коммуникационные технологии».*
Периодичность: *4 раз в год.*
Тираж: *300 экземпляров.*
Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

News of the National Academy of Sciences of the Republic of Kazakhstan.
Series of physics and informatics.
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty). The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. 16906-Ж,** issued 14.02.2018
Thematic scope: *series physics and information technology.*
Currently: *included in the list of journals recommended by the CCSES MES RK in the direction of «information and communication technologies».*
Periodicity: *4 times a year.*
Circulation: *300 copies.*
Editorial address: *28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19*
http://www.physico-mathematical.kz/index.php/en/

NEWS OF THE NATIONAL ACADEMY OF SCIENCESOF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X
Volume 3. Namber 347 (2023). 147–160
https://doi.org/10.32014/2023.2518-1726.210

UDC 004.056
МРНТИ 81.93.29

© B. Rzayev*, Zh. Beldeubayeva, I. Uvaliyeva, 2023
S. Seifullin Kazakh Agro-Technical Research University, Astana, Kazakhstan.
E-mail: pathinchaos@gmail.com

IDENTIFICATION OF MALICIOUS DATA IN THE INFORMATION NETWORK BY USING THE STACKING METHOD

Rzayev Babyr Temirbekuly — doctoral student, S. Seifullin Kazakh Agro-Technical Research University. 010000. Astana, Kazakhstan
E-mail: pathinchaos@gmail.com. ORCID ID: https://orcid.org/0000-0002-9671-650X;
Beldeubayeva Zhanar Toleubayevna — PhD, Senior lecturer, S. Seifullin Kazakh Agro-Technical Research University. 010000. Astana, Kazakhstan
E-mail: zh.beldeubayeva@mail.ru. ORCID ID: https://orcid.org/0000-0003-4056-6220;
Uvaliyeva Indira Makhmutovna — PhD, associate professor, D. Serikbayev East Kazakhstan Technical University. 070000. Ust-Kamenogorsk, Kazakhstan
E-mail: iuvalieva@mail.ru. ORCID ID: https://orcid.org/0000-0002-2117-5390.

Abstract. Information security is now more relevant than ever, and information is now as valuable to criminals as our physical property. The attacker's motives may include stealing information, obtaining financial benefits, spying or sabotage. Organizations should allocate funds to ensure security and be ready to detect, respond and proactively prevent attacks such as phishing, malicious software, viruses, malicious insiders and ransomware. Since the number of cyber threats is growing rapidly, organizations cannot prepare for all of them. Often, the information security systems used are not enough to identify new types of attacks and vulnerabilities — it is necessary to complement existing security systems with intelligent solutions. This paper proposes an approach to solving the problem of detecting malicious traffic in data transmission networks based on processing the received tuples of information sequences of network packets by the ensemble classification method — stacking machine learning algorithms. The approach does not require special data preparation, the resulting classification errors of individual algorithms are smoothed out by the solution of the metaclassifier. The proposed solution, in order to increase the accuracy and completeness of detecting destructive effects, makes it possible to use its classification algorithms optimized for different types of anomalies, which are trained on their own subsets of data

presented as a tuple of values of information sequences of network packets. The experiment is described using the machine learning classifiers Naïve Bayes, Hoeffding Three, Random Tree, REP Tree and J48. The evaluation was carried out using classifiers separately and using stacking, which was based on the same classifiers. Experimental results were obtained on the NSL-KDD public dataset. The software implementation of the approach, as a full-fledged intellectual solution, will make it possible to more effectively identify destructive effects. The approach can be applied as an addition to the existing monitoring systems of organizations related to network traffic processing. The essential advantages of the approach are its versatility for various technologies and data processing systems, the purpose of which is the accurate classification of data, and scalability, through the use of additional algorithms beyond those used in the approach. The purpose of this study is to increase the accuracy of detecting malicious network traffic by using stacking machine learning algorithms.

Keywords: information security, machine learning, stacking of algorithms, network traffic, NSL-KDD

© Б.Т. Рзаев*, Ж.Т. Бельдеубаева, И.М. Увалиева, 2023
С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті,
Астана, Қазақстан.
E-mail: pathinchaos@gmail.com

СТЕКИНГ ӘДІСІН ҚОЛДАНУ АРҚЫЛЫ АҚПАРАТТЫҚ ЖЕЛІДЕГІ ЗИЯНДЫ ДЕРЕКТЕРДІ АНЫҚТАУ

Рзаев Бабыр Темірбекұлы — докторант. С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті. 010000. Астана, Қазақстан
E-mail: pathinchaos@gmail.com. ORCID ID: https://orcid.org/0000-0002-9671-650X;
Бельдеубаева Жанар Толеубаевна — PhD, С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті. 010000. Астана, Қазақстан
E-mail: zh.beldeubayeva@mail.ru. ORCID ID: https://orcid.org/0000-0003-4056-6220;
Увалиева Индира Махмутовна — PhD, қауымдастырылған профессор, Д. Серікбаев атындағы Шығыс Қазақстан техникалық университеті. 010000. Өскемен, Казақстан
E-mail: iuvalieva@mail.ru. ORCID ID: https://orcid.org/0000-0002-2117-5390.

Аннотация. Қазіргі таңда ақпараттық қауіпсіздік бұрынғыдан да өзекті ал ақпарат біздің физикалық мүлкіміз сияқты қылмыскерлер үшін құнды болып келеді. Қылмыскердің мотивтеріне ақпаратты ұрлау, қаржылық пайда табу, тыңшылық немесе диверсия кіруі мүмкін. Ұйымдар қауіпсіздікті қамтамасыз етуге қаражат бөліп, фишинг, зиянды бағдарламалық жасақтама, вирустар, зиянды инсайдерлер және төлем бағдарламалары сияқты шабуылдарды анықтауға, әрекет етуге және алдын-алуға дайын болуы керек. Киберқауіптердің саны тез өсіп келе жатқандықтан, ұйымдар олардың барлығына дайын бола алмайды. Көбінесе ақпараттық қауіпсіздікті қамтамасыз ету жүйелері шабуылдардың жаңа түрлерін және осалдықтарды анықтау үшін жеткіліксіз,

сондықтан қолданыстағы қауіпсіздік жүйелерін зияткерлік шешімдермен толықтыру қажет. Бұл жұмыста ансамбльдік жіктеу әдісі — машиналық оқыту алгоритмдерінің стекингімен желілік пакеттердің ақпараттық тізбектер кортеждерін өңдеуге негізделген деректерді тарату желілеріндегі зиянды трафикті анықтау мәселесін шешу тәсілі ұсынылады. Тәсіл деректерді арнайы дайындауды қажет етпейді, алынған жеке алгоритмдердің жіктеу қателері метаклассификатор шешімімен тегістеледі. Диструктивті әсерлерді анықтаудың дәлдігі мен толықтығы көрсеткіштерін арттыру мақсатында ұсынылған шешім желілік пакеттердің ақпараттық тізбектер мәндерінің кортежі түрінде ұсынылған деректердің өзіндік ішкі жиындарында оқытылған түрлі аномалиялар үшін оңтайландырылған олардың жіктеу алгоритмдерін пайдалануға мүмкіндік береді. Naïve Bayes, Hoeffding Three, Random Tree, REP Tree және J48 машиналық оқыту классификаторларын қолданылған эксперименттердің сипаттамасы берілген. Бағалау классификаторларды жеке қолдану және сол классификаторларды жіктеу негізіндегі стекингті қолдану арқылы жүргізілді. Эксперименттік нәтижелер NSL-KDD жалпыға ортақ деректер жинағын өңдеу арқылы алынған. Диструктивті әсерлерді тиімді анықтауға мүмкіндік беретін толыққанды интеллектуалды шешім ретінде тәсілді бағдарламалық қамтамасыз ету қажет. Тәсіл желілік трафикті өңдеуге қатысты ұйымдардың қолданыстағы бақылау жүйелеріне қосымша ретінде қолданылуы мүмкін. Тәсілдің маңызды артықшылықтары оның әртүрлі технологиялар мен деректерді өңдеу жүйелері үшін әмбебаптығы және тәсілде қолданылғаннан тыс қосымша алгоритмдерді қолдана отырып масштабталуы болып табылады. Зерттеудің мақсаты машиналық оқыту алгоритмдер стекингі арқылы зиянды желілік трафикті анықтау дәлдігін арттыру болып табылады.

Түйін сөздер: ақпараттық қауіпсіздік, машиналық оқыту, алгоритмдер стекингі, желілік трафик, NSL-KDD

© Б.Т. Рзаев*, Ж.Т. Бельдеубаева, И.М. Увалиева, 2023
Казахский агротехнический исследовательский университет
имени С. Сейфуллина, Астана, Казахстан.
E-mail: pathinchaos@gmail.com

ИДЕНТИФИКАЦИЯ ВРЕДОНОСНЫХ ДАННЫХ В ИНФОРМАЦИОННОЙ СЕТИ С ИСПОЛЬЗОВАНИЕМ МЕТОДА СТЕКИНГА

Рзаев Бабыр Темирбекулы — докторант, Казахский агротехнический исследовательский университет имени С.Сейфуллина. 010000. Астана, Казахстан
E-mail: pathinchaos@gmail.com. ORCID ID: https://orcid.org/0000-0002-9671-650X;
Бельдеубаева Жанар Толеубаевна — PhD, Казахский агротехнический исследовательский университет имени С.Сейфуллина. 010000. Астана, Казахстан
E-mail: zh.beldeubayeva@mail.ru. ORCID ID: https://orcid.org/0000-0003-4056-6220;

Увалиева Индира Махмутовна — PhD, ассоциированный профессор, Восточно-Казахстанский технический университет им. Д. Серикбаева. 010000. Усть-Каменогорск, Казахстан
E-mail: iuvalieva@mail.ru. ORCID ID: https://orcid.org/0000-0002-2117-5390.

Аннотация. Информационная безопасность сейчас как никогда актуальна, а информация теперь так же ценна для преступников, как и наше физическое имущество. Мотивы злоумышленника могут включать кражу информации, получение финансовой выгоды, шпионаж или саботаж. Организации должны выделять средства на обеспечение безопасности и быть готовыми к обнаружению, реагированию и упреждающему предотвращению таких атак, как фишинг, вредоносное программное обеспечение, вирусы, вредоносные инсайдеры и программы — вымогатели. Поскольку количество киберугроз быстро растет, организации не могут подготовиться ко всем из них. Зачастую, имеющихся систем обеспечения информационной безопасности недостаточно для выявления новых видов атак и уязвимостей. Необходимо доукомплектовывать существующие системы безопасности новыми интеллектуальными решениями. В данной работе предлагается подход к решению проблемы выявления вредоносного трафика в сетях передачи данных, основанный на обработке полученных кортежей информационных последовательностей сетевых пакетов ансамблевым методом классификации – стекингом алгоритмов машинного обучения. Подход не требует специальной подготовки данных и полученные ошибки классификации отдельных алгоритмов сглаживаются решением метаклассификатора. Предложенное решение с целью повышения показателей точности и полноты выявления диструктивных воздействий дает возможность использовать оптимизированные для разных типов аномалий свои алгоритмы классификации, которые обучены на собственных подмножествах данных, представленных в виде кортежа значений информационных последовательностей сетевых пакетов. Приведено описание эксперимента с использованием классификаторов машинного обучения Naïve Bayes, Hoeffding Tree, Random Tree, REP Tree и J48. Оценка производилась с использованием классификаторов в отдельности и с применением стекинга, в основе которого были использованы те же классификаторы. Экспериментальные результаты получены на публичном наборе данных NSL-KDD. Программная реализация подхода как полноценного интеллектуального решения позволит более эффективно выявлять диструктивные воздействия. Подход может быть применим как дополнение к существующим системам мониторинга организаций, связанных с обработкой сетевого трафика. Существенными преимуществами подхода является его универсальность для различных технологий и систем обработки данных, целью которых является точная классификация данных и масштабируемость путем применения дополнительных алгоритмов сверх используемых в подходе.

Целью данного исследования является повышение точности выявления

вредоносного сетевого трафика, путем применения стекинга алгоритмов машинного обучения.

Ключевые слова: информационная безопасность, машинное обучение, стекинг алгоритмов, сетевой трафик, NSL-KDD

Introduction

The functioning of corporate telecommunications networks (CTS) requires constant monitoring of the occurrence of all types of failures, collisions associated with the processing of network traffic. The development of the concept of the industrial Internet, the Internet of Things makes it necessary to assess the operability, functional security of individual network devices and the network segments formed by them. The analysis of the CTS operability is carried out using various monitors that process internal and external information containing statistical data of network packets and indicators of their processing. As a result, multidimensional time series are formed, which can contain many time-varying parameters reflecting the functioning of the system.

The variety of elements of the Internet of Things, a large number of objects, protocols of interaction in network traffic, data processing technologies, heterogeneity of formats, constantly changing architecture, configuration changes and improvement of attacks cause problems with prompt detection and response to subsequent information security (IS) incidents. This shows that existing approaches and methods may not always be effective in conditions of constant changes in CTS.

The task of identifying malicious data refers to the tasks of detecting anomalies (Cyral, 2022). There are specialized anomaly detection systems (Knapp, 2011), which allow you to detect unusual behavior or events in network traffic. They can help network administrators detect and respond to security threats, network errors, and performance issues. In recent years, there has been a growing interest of the scientific community in investigating the problem of anomaly detection using machine learning and deep learning methods (Yuan Gao et al., 2023).

Related works

Here is an overview of scientific studies on key aspects of detecting anomalies in network traffic.

Types of anomalies. Study (Hayes et al., 2015) offers the division of anomalies in network traffic into three types: point anomalies, contextual anomalies and collective anomalies. Point anomalies refer to individual data points that differ significantly from the rest of the data. Contextual anomalies occur when the behavior of the system deviates from the expected context or pattern. Collective anomalies occur when a group of data points deviates from an expected pattern.

Problems with detecting anomalies. Detection of anomalies in network traffic faces a number of problems, including high dimensionality (Zheng et al., 2022) and the complexity of network data, the dynamic nature of network traffic (Stephen Ranshous et al., 2015) and the presence of noisy and incomplete data. In addition, it is necessary to carefully balance the ratio between false positive and false negative results in order not to miss real anomalies and minimize false positives.

Areas of application of anomaly detection. Anomaly detection systems in network traffic have a wide range of applications, including intrusion detection, network monitoring and performance analysis. For example, anomaly detection can be used to detect malicious activity in network traffic, such as DDoS attacks (Purwanto et al., 2014; Haiping et al., 2022; Purwanto et al., 2015; Chovanec et al, 2023; Lopez et al., 2019) and botnet activity (Zhao et al., 2013; Alaa Obeidat et al., 2022). It can also be used to detect network performance issues (Wawrowski et al., 2023; Igor Fosić et al., 2023), such as packet loss and latency, and to identify optimization opportunities.

Approaches to detecting anomalies in network traffic include statistical methods, machine learning methods, and rule-based methods.

One of the common approaches to detecting anomalies in network traffic are statistical methods (Iglesias Vázquez et al., 2014; Liu et al., 2023). This includes analyzing the statistical properties of network traffic data to detect unusual patterns or behaviors. For example, an anomaly can be detected if the volume of traffic or the frequency of certain types of traffic deviates significantly from the expected levels. An example of a statistical method for detecting anomalies is the use of moving averages (Zhang et al., 2020; Zhou Zeng-Guang et al., 2016) or exponential smoothing (Tang et al., 2022) to identify trends and anomalies in network traffic data.

Machine learning methods (Eduardo Weber Wächter et al., 2022; Nassif Ali et al., 2021; Thudumu et al., 2020). They can also be used to detect anomalies in network traffic. These methods include training the model on a large set of network traffic data and using the model to identify unusual patterns or behaviors in new data. For example, clustering methods can determine the current state of IoT devices (Sukhoparov et al., 2020). Also, an example of a machine learning method for detecting anomalies is the use of neural networks (Benjamin Staar et al., 2019; AlDahoul et al., 2021), who are able to study complex patterns and relationships in data.

Rule-based methods (Elfaki Abdelrahman, 2014; Duffield et al., 2009), They can also be used to detect anomalies in network traffic. These methods include defining a set of rules or thresholds that trigger an alert when certain conditions are met. For example, a rule can be defined to trigger an alert if the number of failed login attempts exceeds a certain threshold within a specified time period. An example of a rule-based anomaly detection method is the use of Snort (Szmit Maciej et al., 2007). An open source intrusion detection system that uses a set of predefined rules to detect various types of network threats.

In conclusion, it should be noted that each conducted study can make a significant contribution to the implementation of anomaly detection systems, which are an important tool for network administrators and security specialists. They help detect unusual behavior and events in network traffic, identify security threats and network performance issues. With the advent of machine learning and deep learning methods, the accuracy and efficiency of anomaly detection systems continue to increase.

Proposed solution

This work involves the use of a stacking algorithm to identify abnormal, potentially malicious data in network traffic.

The operation of the algorithm can be represented as follows (Figure 1).

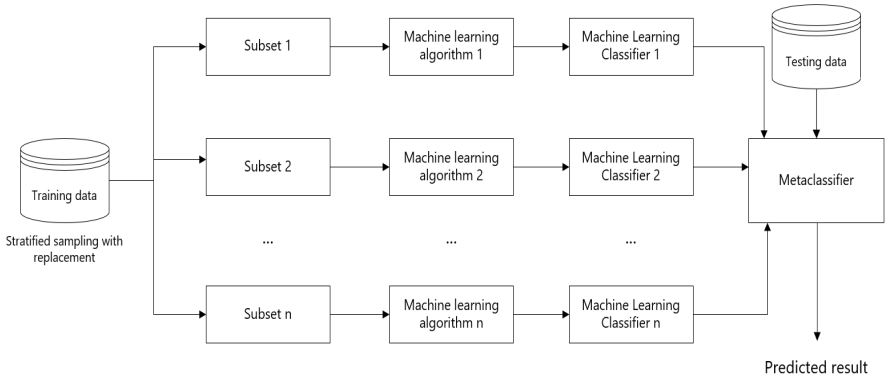

Fig. 1. Stacking operation scheme

There are -distinct subsets of the training dataset, which are created using a stratified sample with substitution, where the relative proportion of different classes is preserved in all subsets (Sikora Riyaz et al., 2014). Each subset of the training set is used to determine the performance of classifiers in the training set. A metaclassifier in the form of a relative weight for each classifier is created by assigning a weight proportional to its performance to the classifier.

When evaluating an instance from a test set, each algorithm outputs a class distribution vector for this instance, which gives the probability that this particular instance belongs to this class. We can represent the vector of class distribution over *c* classes for the *j*-th classifier by the *1 x c* vector as follows:

$$\Delta_j = [\delta_{1j}\ \delta_{2j}\ \dots\ \delta_{cj}]\ 1 \le j \le n \quad (1)$$

where,

$$0 \le \delta_{ij} \le 1\ \forall\ 1 \le i \le c$$
$$\sum_i \delta_{ij} = 1$$

The class distribution vectors for n classifiers can then be represented by an *n x c* matrix as follows:

$$\boldsymbol{\Delta} = [\Delta_1\ \Delta_2\ \dots\ \Delta_n]^T \quad (2)$$

The metaclassifier creates a weight distribution vector that gives relative weight to various advertisements. The vector of weight distribution over *n* classifiers is represented as follows:

$$\Theta = [\theta_1\ \theta_2 \dots\ \theta_n] \quad (3)$$

where,

$$0 \le \delta_j \le 1$$
$$\sum_j \theta_j = 1$$

Given a class distribution matrix and a weight distribution vector, the metaclassicator evaluates each instance of the test set using the following class distribution *1 x c* vector:

$$\Delta' = \Theta.\boldsymbol{\Delta} = [\delta'_1\ \delta'_2 \dots\ \delta'_c] \quad (4)$$

где,

$$\delta'_i = \sum_j \theta_j \delta_{ij}$$

At the same time, as mentioned above, the stacking algorithm assumes that the weight distribution vector of the metaclassifier Θ it is created by assigning a weight to the classifier proportional to its performance.

Evaluation of the proposed solution and results

The experimental evaluation of the proposed solution was carried out using the publicly distributed NSL-KDD dataset (Bhupendra Ingre et al., 2015), which contains 125973 entries for training and 22544 entries for testing.

The description of the features of the NSL-KDD set is presented in Table 1.

Table 1. Description of the features of the NSL-KDD dataset

№	Feature	Description
1	Duration	Duration of connection time
2	Protocol_type	Protocol used for connection
3	Service	Destination network service used
4	Flag	Connection Status - Normal or Error
5	Src_bytes	The number of bytes of data transferred from the source to the destination in a single connection
6	Dst_bytes	The number of bytes of data transferred from the destination to the source in a single connection
7	Land	if the source and destination IP addresses and port numbers are equal, then this variable takes the value 1, another 0
8	Wrong_fragment	The total number of incorrect fragments in this connection
9	Urgent	The number of urgent packets in this connection. Urgent packets are packets with the urgency bit activated.
10	Hot	The number of "burning" indicators in the content, such as: entering the system directory, creating programs and executing programs
11	Num_failed_logins	Number of failed login attempts
12	Logged_in	Login status: 1 on successful login; 0 otherwise
13	Num_compromised	Number of "compromise" conditions
14	Root_shell	1 if the root shell is obtained; 0 otherwise
15	Su_attempted	1 if the "su root" command was tried or used; 0 otherwise

16	Num_root	The number of "root" accesses or the number of operations performed as root in the connection
17	Num_file_creations	The number of operations to create a file when connecting
18	Num_shells	Number of shell hints
19	Num_access_files	Number of operations with access control files
20	Num_outbound_cmds	Numbering of outgoing commands in an ftp session
21	Is_hot_login	1 if the login belongs to the "hot" list, i.e. root or administrator; otherwise 0
22	Is_guest_login	1 if login is "guest"; 0 otherwise
23	Count	The number of connections to the same destination node as the current connection in the last two seconds.
24	Srv_count	The number of connections to the same service (port number) as the current connection in the last two seconds.
25	Serror_rate	Percentage of connections that activated the flag (4) s0, s1, s2 or s3 among the connections combined in count (23)
26	Srv_serror_rate	Percentage of connections that activated the flag (4) s0, s1, s2 or s3 among the connections combined in srv_count (24)
27	Rerror_rate	Percentage of connections that activated the REJ flag (4) among the connections combined in count (23)
28	Srv_rerror_rate	Percentage of connections that activated the REJ flag (4) among connections combined in srv_count (24)
29	Same_srv_rate	Percentage of connections to the same service, among the connections combined in the account (23)
30	Diff_srv_rate	Percentage of connections to various services, among the connections combined in the column (23)
31	Srv_diff_host_rate	Percentage of connections that were to different destination machines among the connections combined in srv_count (24)
32	Dst_host_count	The number of connections having the same destination host IP address
33	Dst_host_srv_count	The number of connections having the same port number
34	Dst_host_same_srv_rate	Percentage of connections to the same service, among connections combined in dst_host_count (32)
35	Dst_host_diff_srv_rate	Percentage of connections to various services, among connections combined in dst_host_count (32)
36	Dst_host_same_src_port_rate	Percentage of connections that were to the same source port among the connections combined in dst_host_srv_count (33)
37	Dst_host_srv_diff_host_rate	Percentage of connections that were to different destination machines among the connections combined in dst_host_srv_ count (33)
38	Dst_host_serror_rate	Percentage of connections that activated the flag (4) s0, s1, s2 or s3 among the connections combined in dst_host_count (32)
39	Dst_host_srv_serror_rate	Percentage of connections that activated the flag (4) s0, s1, s2 or s3 among the connections combined in dst_host_srv_count (33)
40	Dst_host_rerror_rate	Percentage of connections that activated the (4) REJ flag among the connections combined in dst_host_count (32)
41	Dst_host_srv_rerror_rate	Percentage of connections that activated the REJ flag (4) among the connections combined in dst_host_srv_count (33)
42	Class	Class of data

Training is performed based on KDDTrain data, which contains 22 types of attacks, and testing is performed on KDDTest data, which contains an additional 17 types of attacks. These attacks are divided into four classes of attacks:

Denial of Service (DoS) - malicious attempt to block system or network resources and services.

Probe – this attack collects information about potential vulnerabilities of the target system, which can later be used to launch attacks on these systems.

Remote to Local (R2L) – Unauthorized ability to send data packets to a remote system over the network and gain access either as a user or as root to perform their unauthorized actions.

User to Root (U2R) – In this case, attackers gain access to the system as a normal user and hack vulnerabilities to gain administrative privileges.

In addition to attacks, for classification purposes, the set also contains normal data that does not contain malicious components. The number of NSL-KDD set entries distributed by classes is shown in the Table 2.

Table 2. Number of NSL-KDD set entries distributed by category

Testing dataset		Training dataset	
Class	Number of entries	Class	Number of entries
Normal	67343	Normal	9711
DOS	45927	DOS	7458
Probe	11656	Probe	2421
R2L	995	R2L	2754
U2R	52	U2R	200
Total	125973	Total	22544

The freely distributed Weka application for data processing and machine learning was used for the experiment. This application is written in Java at the University of Waikato (New Zealand), distributed under the GNU GPL license (Wikipedia, 2021).

The first part of the experiment involved the use of classifiers separately. Naïve Bayes (NB), Hoeffding Tree (HT), Random Tree (RT), REP Tree (REP) and J48 were selected from the classifiers.

The classifiers were evaluated using the ROC error curve, which displays the ratio between correctly classified records and incorrectly classified ones, Figure 2.

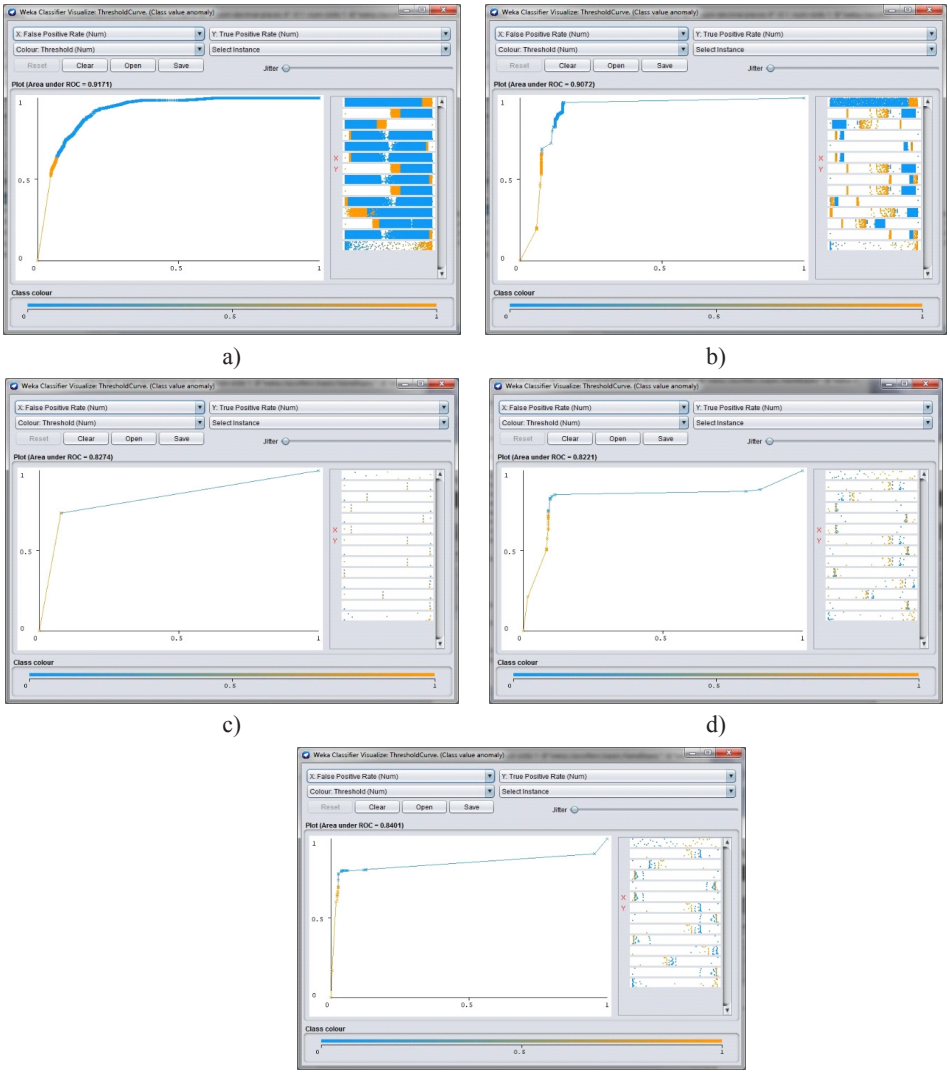

Fig. 2. ROC curves: a) NB, b) HT, c) RT, d) REP, e) J48

The obtained results of data classification using ML algorithms are presented in the Table 3.

Table 3. Classification results

Parameter\Classifier	NB	HT	RT	REP	J48
Accuracy, %	76,1	77,1	81,3	81,5	81,5
Precision, %	80,9	81,2	83,7	83,5	85,8
Recall, %	76,1	77,2	81,4	81,5	81,5
F-measures, %	75,9	77,1	81,4	81,6	81,5
ROC	91,7	90,7	82,7	82,2	84,0

The best results on the ROC curve were shown by the NB classifier (81,5 %).

In the second part of the experiment, the stacking of ML algorithms (ST) was implemented using the same classifiers - Naïve Bayes (NB), Hoeffding Tree (HT), Random Tree (RT), REP Tree (REP) and J48. The metaclassifier for stacking was chosen – Logistic Regression.

The assessment of the use of stacking was carried out according to the ROC curve, Figure 3.

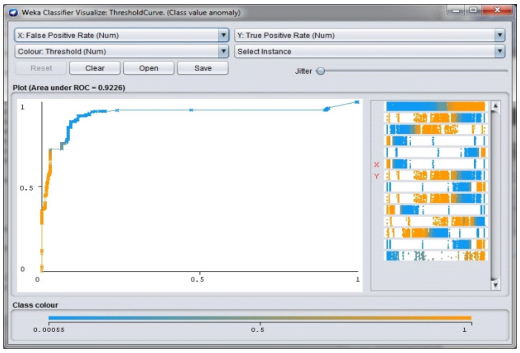

Fig. 3. Indicators of the ROC curve when using stacking

The results of classifiers individually and with the use of stacking are shown in the Table 4.

Table 4. Comparative classification results

Parameter\Classifier	NB	HT	RT	REP	J48	ST
Accuracy, %	76,1	77,1	81,3	81,5	81,5	82,2
Precision, %	80,9	81,2	83,7	83,5	85,8	86,2
Recall, %	76,1	77,2	81,4	81,5	81,5	82,3
F-measures, %	75,9	77,1	81,4	81,6	81,5	82,2
ROC	91,7	90,7	82,7	82,2	84,0	92,2

A comparative histogram using classifiers separately and using stacking is shown in Figure 4.

Fig. 4. Comparative histogram with using stacking

The stacking results are superior to the result of the best single classifier. Therefore, it can be argued about the effectiveness of the use of stacking algorithms more than the use of algorithms separately.

Conclusion

This article proposes an approach to the identification of malicious traffic based on the use of an ensemble method of machine learning – stacking. The approach is based on the use of individual ML classifiers as the basic ones, and on their basis the training of the metaclassifier, in our case, logistic regression is used for subsequent data classification.

The application of the proposed solution based on stacking makes it possible to improve classification accuracy indicators rather than using separate classifiers.

The advantage of using stacking is versatility for various data processing systems and scalability, through the use of additional algorithms beyond those used in the approach.

The disadvantage of the proposed approach is sensitivity to data quality, as one of the main requirements for data classification by MO methods, as well as the need for additional computing resources when the model becomes more complex.

REFERENCES

Cyral, 2022. Anomaly detection. URL: https://cyral.com/glossary/anomaly-detection/

Knapp E., 2011. Exception, Anomaly, and Threat Detection. Industrial Network Security. Pp. 189–214. DOI:10.1016/b978-1-59749-645-2.0000.

Yuan Gao, Xianhui Yin, Zhen He, Xueqing Wang, 2023. A deep learning process anomaly detection approach with representative latent features for low discriminative and insufficient abnormal data. Computers & Industrial Engineering. Volume 176, 108936. ISSN 0360-8352. DOI: https://doi.org/10.1016/j.cie.2022.108936.

Hayes M.A., Capretz M.A., 2015. Contextual anomaly detection framework for big sensor data. Journal of Big Data 2. DOI: https://doi.org/10.1186/s40537-014-0011-y.

Zheng J., Li J., Liu C. et al., 2022. Anomaly detection for high-dimensional space using deep hypersphere fused with probability approach. Complex Intell. Syst. 8. 4205–4220. DOI: https://doi.org/10.1007/s40747-022-00695-9.

Stephen Ranshous, Shitian Shen, Danai Koutra, Steve Harenberg, Christos Faloutsos, Nagiza F. Samatova, 2015. Anomaly detection in dynamic networks: a survey WIREs Comput Stat 2015. 7. Pp. 223–247. DOI: 10.1002/wics.1347.

Y. Purwanto, Kuspriyanto Hendrawan, B. Rahardjo, 2014. Traffic anomaly detection in DDos flooding attack. 8th International Conference on Telecommunication Systems Services and Applications (TSSA), Kuta. Bali. Indonesia. Pp. 1–6. DOI: 10.1109/TSSA.2014.7065953.

Haiping Lin, Chengwen Wu, Mohammad Masdari, 2022. A comprehensive survey of network traffic anomalies and DDoS attacks detection schemes using fuzzy techniques. Computers and Electrical Engineering. Volume 104. Part B. 108466. ISSN 0045-7906. DOI: https://doi.org/10.1016/j.compeleceng.2022.108466.

Purwanto Yudha, Kuspriyanto, Hendrawan Temmy, Rahardjo Budi, 2015. Traffic anomaly detection in DDos flooding attack. Proceedings of 2014 8th International Conference on Telecommunication Systems Services and Applications. TSSA 2014. DOI: 10.1109/TSSA.2014.7065953.

Chovanec M., Hasin M., Havrilla M., Chovancová E., 2023. Detection of HTTP DDoS Attacks Using NFStream and TensorFlow. Applied Sciences. 13(11):6671. DOI: https://doi.org/10.3390/app13116671.

Lopez Alma D., Mohan Asha P., Nair Sukumaran, 2019. Network Traffic Behavioral Analytics for Detection of DDoS Attacks. SMU Data Science Review. Vol. 2. No. 1. Article 14.

Zhao David, Traore Issa, Sayed, Bassam Lu, Wei Saad, Sherif Ghorbani, Ali Garant Dan, 2013. Botnet detection based on traffic behavior analysis and flow intervals. Computers & Security. 39. 2–16. DOI: 10.1016/j.cose.2013.04.007

Alaa Obeidat, Rola Yaqbeh, 2022. Smart Approach for Botnet Detection Based on Network Traffic Analysis. Journal of Electrical and Computer Engineering. vol. 2022. Article ID 3073932. DOI: https://doi.org/10.1155/2022/3073932

Wawrowski Ł., Białas A., Kajzer A., Kozłowski A., Kurianowicz R., Sikora M., Szymańska-Kwiecień A., Uchroński M., Białczak M., Olejnik M., Michalak M., 2023. Anomaly Detection Module for Network Traffic Monitoring in Public Institutions. Sensors. 23(6):2974. DOI: https://doi.org/10.3390/s23062974.

Igor Fosić, Drago Žagar, Krešimir Grgić, Višnja Križanović, 2023. Anomaly detection in NetFlow network traffic using supervised machine learning algorithms. Journal of Industrial Information Integration. Volume 33. 100466. ISSN 2452–414X. DOI: https://doi.org/10.1016/j.jii.2023.100466.

Iglesias Vázquez, Félix, Zseby Tanja, 2014. Analysis of network traffic features for anomaly detection. Machine Learning. 101. DOI: 10.1007/s10994-014-5473-9.

Liu H., Wang H., 2023. Real-Time Anomaly Detection of Network Traffic Based on CNN. Symmetry. 15(6):1205. DOI: https://doi.org/10.3390/sym15061205.

Zhang M., Guo J., Li X., Jin R., 2020. Data-Driven Anomaly Detection Approach for Time-Series Streaming Data. Sensors (Basel). 20(19):5646. DOI: 10.3390/s20195646.

Zhou, Zeng-Guang, Tang Ping, 2016. Improving time series anomaly detection based on exponentially weighted moving average (EWMA) of season-trend model residuals. DOI: 10.1109/IGARSS.2016.7729882.

Tang H., Wang Q., Jiang G., 2022. Time Series Anomaly Detection Model Based on Multi-Features. Comput Intell Neurosci. 2022:2371549. DOI: 10.1155/2022/2371549.

Eduardo Weber Wächter, Server Kasap, Şefki Kolozali, Xiaojun Zhai, Shoaib Ehsan, Klaus D. McDonald-Maier, 2022. Using machine learning for anomaly detection on a system-on-chip under gamma radiation. Nuclear Engineering and Technology. Volume 54. Issue 11. Pp. 3985–3995. ISSN 1738-5733. DOI: https://doi.org/10.1016/j.net.2022.06.028.

Nassif Ali, Abu Talib Manar, Nasir Qassim, Dakalbab Fatima, 2021. Machine Learning for Anomaly Detection: A Systematic Review. IEEE Access. Pp. 1–1. DOI: 10.1109/ACCESS.2021.3083060.

Thudumu S., Branch P., Jin J. et al., 2020. A comprehensive survey of anomaly detection techniques for high dimensional big data. J Big Data 7. 42. DOI: https://doi.org/10.1186/s40537-020-00320-x.

Sukhoparov M. E., Lebedev I. S., 2020. Identification of the state of information security of Internet of Things devices in information and telecommunication systems. Control systems, communications and security. № 3. Pp. 252–268. DOI: 10.24411/2410-9916-2020-10310.

Benjamin Staar, Michael Lütjen, Michael Freitag, 2019. Anomaly detection with convolutional neural networks for industrial surface inspection. Procedia CIRP. Volume 79. Pp. 484–489. ISSN 2212–8271. https://doi.org/10.1016/j.procir.2019.02.123.

AlDahoul N., Abdul Karim H., Ba Wazir A.S., 2021. Model fusion of deep neural networks for anomaly detection. J Big Data 8. 106 (2021). DOI: https://doi.org/10.1186/s40537-021-00496-w.

Elfaki Abdelrahman, 2014. Using a Rule-based Method for Detecting Anomalies in Software Product Line. Research Journal of Applied Sciences. Engineering and Technology. 7.

N. Duffield, P. Haffner, B. Krishnamurthy, H. Ringberg, 2009. Rule-Based Anomaly Detection on IP Flows. IEEE INFOCOM 2009. Rio de Janeiro. Brazil. Pp. 424–432. DOI: 10.1109/INFCOM.2009.5061947.

Szmit Maciej, Wężyk Radosław, Skowroński Maciej, Szmit Anna, 2007. Traffic anomaly detection with Snort.

Sikora Riyaz, Al-laymoun O'la Hmoud, 2014. A Modified Stacking Ensemble Machine Learning Algorithm Using Genetic Algorithms. Journal of International Technology and Information Management. Vol. 23. Iss. 1. Article 1. DOI: https://doi.org/10.58729/1941-6679.1061. Available at: https://scholarworks.lib.csusb.edu/jitim/vol23/iss1/1.

Bhupendra Ingre, Anamika Yadav, 2015. Performance Analysis of NSL-KDD dataset using ANN. 2015 International Conference on Signal Processing And Communication Engineering Systems (SPACES). Pp. 92–96. DOI: 10.1109/SPACES.2015.7058223.

Wikipedia, 2022. Weka. URL: https://ru.wikipedia.org/

МАЗМҰНЫ

Г. Әбдіқалық, Ә. Мұқанова, А. Назырова
CRF ЖӘНЕ RANDOM FOREST МОДЕЛДЕРІНІҢ КӨМЕГІМЕН ҚАЗАҚ ТІЛІНДЕ АТАЛҒАН ОБЪЕКТІЛЕРДІ ТАНУ: САЛЫСТЫРМАЛЫ ЗЕРТТЕУ.....7
Г.Б. Абдикеримова, М.Б. Есенова, Т.Т. Оспанова, У.Ж. Айтимова, М. Айтимов
ҒАРЫШТЫҚ КЕСКІНДЕРДІ ӨҢДЕУДЕ АҚПАРАТТЫҚ ТЕКСТУРАЛЫҚ LABC МАСКАЛАР ӘДІСТЕРІН ҚОЛДАНУ.....18
Б.У. Асанова, Б.Б. Оразбаев, Ж.Ж. Молдашева, Г.Ж. Шуйтенов, Э.М. Дюсембина
ТҮРЛІ СИПАТТАҒЫ ҚОЛ ЖЕТІМДІ АҚПАРАТТАР НЕГІЗІНДЕ БАЯУ КОКСТЕУ ҚОНДЫРҒЫСЫНЫҢ ӨЗАРА БАЙЛАНЫСҚАН ТЕХНОЛОГИЯЛЫҚ АГРЕГАТТАРЫ МОДЕЛЬДЕРІН ҚҰРУ ӘДІСТЕМЕСІ.....28
Г.Б. Бахадирова, Н. Тасболатұлы, А.С. Муканова, Ш. Тураев
MATLAB SIMULINK-ТЕ СЫЗЫҚТЫҚ ЕМЕС ЖҮЙЕ ҮШІН КЕРІ БАЙЛАНЫСТЫ СЫЗЫҚТЫҚ БАСҚАРУДЫ ЖОБАЛАУ.....44
Е.С. Голенко, А.А. Исмаилова
ПРЕДСКАЗАНИЕ ФУНКЦИЙ БЕЛКА С ИСПОЛЬЗОВАНИЕМ КОМБИНАЦИИ BILSTM И АЛГОРИТМА САМОВНИМАНИЯ.....62
Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева
CNN НЕГІЗІНДЕ ҚАЗАҚ ЫМ ТІЛІН ТАНУ.....76
К.К. Кадиркулов, А.А. Исмаилова, Ә.Б. Бейсеғұл
ЛАБОРАТОРИЯЛЫҚ ЗЕРТТЕУ НӘТИЖЕЛЕРІН ТАЛДАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУДЫҢ МОДЕЛІН ТАҢДАУ.....88
А. Мукашова, А. Муханова, Т. Оспанова, А. Бакиева, В. Махатова
ҚҰЗЫРЕТТІК ТӘСІЛДЕР НЕГІЗІНДЕГІ БІЛІМ БЕРУ БАҒДАРЛАМАЛАРЫН ӘЗІРЛЕУДІҢ МАҢЫЗДЫ АСПЕКТІЛЕРІ.....99
Ш.Ж. Мусиралиева, М.А. Болатбек, М. Сағынай, Ж.Ы. Елтай, К.Б. Багитова
ЭКСТРЕМИСТІК МӘЛІМЕТТЕР ТҮСІНІГІ ЖӘНЕ ЭКСТРЕМИЗМГЕ ҚАРСЫ КҮРЕС ЖОБАЛАРЫНА ЖҮЙЕЛІК ШОЛУ.....112
Д. Оралбекова, О. Мамырбаев, А. Жунусова, Б. Жумажанов
КҮРДЕЛІ МОРФОЛОГИЯЛЫҚ ҚҰРЫЛЫМЫ БАР ТІЛГЕ АРНАЛҒАН ЗАМАНАУИ ТІЛДІК МОДЕЛЬДЕУ ӘДІСТЕРІН ЗЕРТТЕУ.....131
Б.Т. Рзаев, Ж.Т. Бельдеубаева, И.М. Увалиева
СТЕКИНГ ӘДІСІН ҚОЛДАНУ АРҚЫЛЫ АҚПАРАТТЫҚ ЖЕЛІДЕГІ ЗИЯНДЫ ДЕРЕКТЕРДІ АНЫҚТАУ.....147
Н.С. Баймулдина, Г.Н. Скабаева, А.Д. Жақсыбаева
БИОТЕХНОЛОГИЯ САЛАСЫНДАҒЫ ЖОБАЛАРДЫ БАСҚАРУДЫҢ БАҒДАРЛАМАЛЫҚ ҚАМТАМАСЫЗ ЕТУІ.....161
А.Ә. Таурбекова, Ө.Ж. Мамырбаев, Б. Т. Қарымсакова, Б. Ж. Жұмажанов
МАГМАНЫҢ ШЫҒУ ПРОЦЕСІН ЗЕРТТЕУ.....176
Г.С. Шаймерденова, Р.А. Саркулакова, М.М. Тұрғанбекова, Б.Ө. Тастанбекова, М.Т. Байжанова,
МОБИЛЬДІ ЖӘНЕ ОНЛАЙН-БАНКИНГТЕГІ ЖЕТІСТІКТЕР: ТЕХНОЛОГИЯЛАР МЕН ИННОВАЦИЯЛАРДЫ КЕШЕНДІ ТАЛДАУ.....193
Я. Кучин, Н. Юничева, Р.И. Мухамедиев, Е. Мухамедиева
МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІМЕН ҚАБАТТЫҢ ТОТЫҒУ АЙМАҚТАРЫН ОҚШАУЛАУ МҮМКІНДІГІН БАҒАЛАУ.....210

СОДЕРЖАНИЕ

Г. Абдикалык, А. Муканова, А. Назырова
РАСПОЗНАВАНИЕ ИМЕНОВАННЫХ ИМЕНОВАННЫХ ОБЪЕКТОВ В КАЗАХСКОМ ЯЗЫКЕ С ПОМОЩЬЮ МОДЕЛЕЙ CRF И RANDOM FOREST: СРАВНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ.....7
Г.Б. Абдикеримова, М.Б. Есенова, Т.Т. Оспанова, У.Ж. Айтимова, М. Айтимов
ИСПОЛЬЗОВАНИЕ МЕТОДОВ ИНФОРМАТИВНОЙ ТЕКСТУРНОЙ МАСОК ЛАВСА ПРИ ОБРАБОТКЕ КОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ.....18
Б.У. Асанова, Б.Б. Оразбаев, Ж.Ж. Молдашева, Г.Ж. Шуйтенов, Э.М. Дюсембина
МЕТОДИКА РАЗРАБОТКИ МОДЕЛЕЙ ВЗАИМОСВЯЗАННЫХ ТЕХНОЛОГИЧЕСКИХ АГРЕГАТОВ УСТАНОВКИ ЗАМЕДЛЕННОГО КОКСОВАНИЯ НА ОСНОВЕ ДОСТУПНОЙ ИНФОРМАЦИИ РАЗЛИЧНОГО ХАРАКТЕРА.....28
Г.Б. Бахадирова, Н. Тасболатұлы, А.С. Муканова, Ш.Тураев
ПРОЕКТИРОВАНИЕ ЛИНЕЙНОГО УПРАВЛЕНИЯ С ОБРАТНОЙ СВЯЗЬЮ ДЛЯ НЕЛИНЕЙНОЙ СИСТЕМЫ В MATLAB SIMULINK.....44
Е.С. Голенко, А.А. Исмаилова
ПРЕДСКАЗАНИЕ ФУНКЦИЙ БЕЛКА С ИСПОЛЬЗОВАНИЕМ КОМБИНАЦИИ BILSTM И АЛГОРИТМА САМОВНИМАНИЯ.....62
Л.З. Жолшиева, Т.К. Жукабаева, Ш. Тураев, М.А. Бердиева
РАСПОЗНАВАНИЕ КАЗАХСКОГО ЖЕСТОВОГО ЯЗЫКА НА ОСНОВЕ CNN.....76
К.К. Кадиркулов, А.А. Исмаилова, Ә.Б. Бейсеғұл
ВЫБОР МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ ПО ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ ЛАБОРАТОРНЫХ ИССЛЕДОВАНИЙ.....88
А. Мукашова, А. Муханова, Т. Оспанова, А. Бакиева, В. Махатова
ВАЖНЫЕ АСПЕКТЫ РАЗРАБОТКИ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ, ОСНОВАННЫХ НА КОМПЕТЕНТНОСТНОМ ПОДХОДЕ.....99
Ш.Ж. Мусиралиева, М.А. Болатбек, М. Сағынай, Ж.Ы. Елтай, К.Б. Багитова
ПОНЯТИЕ ЭКСТРЕМИСТСКИХ ДАННЫХ И СИСТЕМНЫЙ ОБЗОР ПРОЕКТОВ ПО БОРЬБЕ С ЭКСТРЕМИЗМОМ.....112
Д. Оралбекова, О. Мамырбаев, А. Жунусова, Б. Жумажанов
ИССЛЕДОВАНИЕ СОВРЕМЕННЫХ МЕТОДОВ ЯЗЫКОВОГО МОДЕЛИРОВАНИЯ ДЛЯ ЯЗЫКА СО СЛОЖНОЙ МОРФОЛОГИЧЕСКОЙ СТРУКТУРОЙ.....131
Б.Т. Рзаев, Ж.Т. Бельдеубаева, И.М. Увалиева
ИДЕНТИФИКАЦИЯ ВРЕДОНОСНЫХ ДАННЫХ В ИНФОРМАЦИОННОЙ СЕТИ С ИСПОЛЬЗОВАНИЕМ МЕТОДА СТЕКИНГА.....147
Н.С. Баймулдина, Г.Н.Скабаева, А.Д. Жақсыбаева
ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ УПРАВЛЕНИЯ ПРОЕКТАМИ В ОБЛАСТИ БИОТЕХНОЛОГИИ.....161
А.А. Таурбекова, О.Ж. Мамырбаев, Б.Т. Карымсакова, Б.Ж. Жумажанов
ИССЛЕДОВАНИЯ ПРОЦЕССА ИСТЕЧЕНИЯ МАГМЫ.....176
Г.С. Шаймерденова, Р.А. Саркулакова, М.М. Турганбекова, Б.О. Тастанбекова, М.Т. Байжанова
ДОСТИЖЕНИЯ В МОБИЛЬНОМ И ОНЛАЙН-БАНКИНГЕ: КОМПЛЕКСНЫЙ АНАЛИЗ ТЕХНОЛОГИЙ И ИННОВАЦИЙ.....193
Я. Кучин, Н. Юничева, Р.И. Мухамедиев, Е. Мухамедиева
ОЦЕНКА ВОЗМОЖНОСТИ ВЫДЕЛЕНИЯ ЗОН ПЛАСТОВОГО ОКИСЛЕНИЯ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ.....210

CONTENTS

G. Abdikalyk, A. Mukanova, A. Nazyrova
NAMED ENTITY RECOGNITION FOR KAZAKH LANGUAGE USING CRF AND RANDOM FOREST MODELS: A COMPARATIVE STUDY....7
G.B. Abdikerimova, M.B. Yessenova, T.T. Ospanova, U.Zh Aitimova, M. Murat
USE OF INFORMATION TEXTURE LAWS MASK METHODS IN SPACE IMAGE PROCESSING....18
B. Assanova, B. Orazbayev, Zh. Moldasheva, G. Shuitenov, E. Dyussembina
METHODOLOGY FOR DEVELOPING MODELS OF INTERRELATED TECHNOLOGICAL UNITS OF A DELAYED COKING UNIT ON THE BASIS OF AVAILABLE INFORMATION OF A DIFFERENT NATURE....28
G.B. Bahadirova, H. Tasbolatuly, A.S. Mukanova, Sh. Turaev
DESIGNING LINEAR FEEDBACK CONTROL FOR A NONLINEAR SYSTEM IN MATLAB SIMULINK....44
Y.S. Golenko, A.A. Ismailova
PROTEIN FUNCTION PREDICTION USING THE COMBINATION OF BILSTM AND SELF-ATTENTION ALGORITHM....62
L. Zholshiyeva, T. Zhukabayeva, Sh. Turaev, M. Berdieva
KAZAKH SIGN LANGUAGE RECOGNITION BASED ON CNN....76
K. Kadirkulov, A. Ismailova, A. Beissegul
SELECTION OF A MACHINE LEARNING MODEL FOR INTERPRETING LABORATORY RESULTS....88
A. Mukashova, A. Mukanova, T. Ospanova, A. Bakiyeva, V. Makhatova
IMPORTANT ASPECTS OF DEVELOPING EDUCATIONAL PROGRAMS BASED ON THE COMPETENCY-BASED APPROACH....99
Sh. Mussiraliyeva, M. Bolatbek, M. Sagynay, Zh. Yeltay, K. Bagitova
THE CONCEPT OF EXTREMIST DATA AND A SYSTEMATIC REVIEW OF ANTI-EXTREMISM PROJECTS....112
D. Oralbekova, O. Mamyrbayev, A. Zhunussova, B. Zhumazhanov
STUDY OF MODERN METHODS OF LANGUAGE MODELING FOR A LANGUAGE WITH A COMPLEX MORPHOLOGICAL STRUCTURE....131
B. Rzayev, Zh. Beldeubayeva, I. Uvaliyeva
IDENTIFICATION OF MALICIOUS DATA IN THE INFORMATION NETWORK BY USING THE STACKING METHOD....147
N.S. Baimuldina, G.N. Skabayeva, A. Zhaksybayeva
PROJECT MANAGEMENT SOFTWARE IN THE FIELD OF BIOTECHNOLOGY....161
A.A. Taurbekova, O.Zh. Mamyrbaev, B.T. Karymsakova, B.Zh. Zhumazhanov
INVESTIGATIONS OF MAGMA OUTPUT PROCESS....176
G.S. Shaimerdenova, R.A. Sarkulakova, M.M. Turganbekova, B.O. Tastanbekova, M.T. Baizhanova
ADVANCEMENTS IN MOBILE AND ONLINE BANKING: A COMPREHENSIVE ANALYSIS OF TECHNOLOGIES AND INNOVATIONS....193
Y. Kuchin, N. Yunicheva, R.I. Mukhamediev, E. Mukhamedieva
ESTIMATION OF THE POSSIBILITY TO SELECT RESERVOIR OXIDATION ZONES BY MACHINE LEARNING METHODS....210

Publication Ethics and Publication Malpracticein the journals of the National Academy of Sciences of the Republic of Kazakhstan

For information on Ethics in publishing and Ethicalguidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any otherlanguage, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:
www:nauka-nanrk.kz
http://physics-mathematics.kz/index.php/en/archive
ISSN2518-1726 (Online),
ISSN 1991-346X (Print)

Подписано в печать 28.09.2023.
Формат 60x881/8. Бумага офсетная. Печать –ризограф.
18,0 п.л. Тираж 300. Заказ 3.

РОО «Национальная академия наук РК» 050010,
Алматы, ул. Шевченко, 28, т. 272-13-19