КОМПЬЮТЕРЛІК ТЕХНИКА ЖӘНЕ СӘУЛЕТ
КОМПЬЮТЕРНАЯ ТЕХНИКА И АРХИТЕКТУРА
COMPUTER TECHNOLOGY AND ARCHITECTURE

DOI 10.51885/1561-4212_2023_3_116
МРНТИ 28.23.37

О.А. Усатова[1,2], А.Т. Жұмабекова[1], В.И. Карюкин[1], Ж.Б. Медетбек[1], Г.Е. Алпысбай[1], Қ.Т. Жұмабекова[3], Е.Е. Бегимбаева[4]
[1]әл-Фараби атындағы Қазақ Ұлттық Университеті, Алматы қ., Қазақстан
[2]Ғ. Дәукеев атындағы Алматы энергетика және байланыс университеті, Алматы қ., Қазақстан
[3]Қазақстан Республикасы ІІМ М. Есболатов атындағы Алматы академиясы, Алматы қ., Қазақстан
[4]Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы қ., Қазақстан
*E-mail: zhumabekova2702@gmail.com**
E-mail: vladislav.karyukin@gmail.com
E-mail: medetbek.zhanar@gmail.com
E-mail: gulbanu.alpysbay@gmail.com
E-mail: uoa_olga@mail.ru
E-mail: kz_kuanysh93@mail.ru
E-mail: enlik_89@mail.ru

ЖАСАНДЫ ИНТЕЛЛЕКТТІ ҚОЛДАНУ АРҚЫЛЫ SQL ИНЪЕКЦИЯЛАРЫН АНЫҚТАУ

ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ОБНАРУЖЕНИЯ SQL ИНЪЕКЦИИ

APPLICATION OF ARTIFICIAL INTELLIGENCE TO DETECT SQL INJECTIONS

Аңдатпа. *Ақпараттық технологиялардың дамуы адамдардың өміріне айтарлықтай әсер етті. Сонымен бірге ақпараттық қауіпсіздіктің маңыздылығы барған сайын маңызды бола түсуде. Деректер қорында пайдаланушылардың жеке мәліметтерінің тұтастығын, қолжетімділігін және құпиялылығын қорғау ақпараттық қауіпсіздіктің негізгі мақсаты болып табылады. Веб-қосымшалардың қауіпсіздігіне ең үлкен қауіптердің бірі – SQL инъекциясы. SQL инъекциясы қосымшаның ақпараттық қауіпсіздігіне нұқсан келтіретін мәліметтер базасын тікелей басқару үшін веб-қосымшаларға кіруді бақылауды айналып өтуге мүмкіндік беретін шабуылдың осы түріне жатады. Бұл мақалада біз жасанды интеллект әдістерін қолдана отырып, SQL инъекциясын анықтауды қарастырамыз. Бұл тапсырма үшін келесі алгоритмдерді қолдана отырып, машиналық оқыту және нейрондық желілердің модельдері жасалды: Наив Байес классификаторы, қолдау векторлық машинасы, логистикалық регрессия, шешім ағашы, кездейсоқ орман, XGBoost, AdaBoost және терең нейрондық желілер. Әзірленген модельдер арқылы деректерді жіктеудің тиімділігін бағалау үшін 0,95-0,99 диапазонындағы мәндерге жеткен дұрыстығы, дәлдігі, толықтығы және F-өлшемі көрсеткіштері пайдаланылды.*

Түйін сөздер: *жасанды интеллект, машиналық оқыту, нейрондық желілер, SQL инъекция.*

Аннотация. *Развитие информационных технологий существенно повлияло на жизнедеятельность людей. При этом важность информационной безопасности становится все более и более существенной. Защита целостности, доступности и конфиденциальности персональных данных пользователей в базе данных является основной целью информационной безопасности. Одной из самых больших угроз безопасности веб-приложений являются SQL инъекции. SQL инъекция относится к такому типу атак, которая позволяет обходить контроль доступа к веб-приложениям для непосредственного управления базами данных, что ставит под угрозу информационную безопасность приложения. В данной статье мы рассматриваем выявление SQL-инъекций с использованием методов искусственного интеллекта. Для этой задачи были разработаны модели машинного обучения и нейронных сетей, использующий следующие алгоритмы: Наивный*

байесовский классификатор, Машина опорных векторов, Логистическая регрессия, Дерево решений, Случайный лес, XGBoost, AdaBoost и Глубокие нейронные сети. Для оценки эффективности классификации данных с помощью разработанных моделей использовались показатели правильности, точности, полноты и F-критерия, достигнувшие значений в диапазоне 0,95-0,99.

Ключевые слова: *искусственный интеллект, машинное обучение, нейронные сети, SQL-инъекция.*

Abstract. *The development of information technologies has significantly affected people's lives. At the same time, the importance of information security is becoming more and more significant. Protecting the integrity, accessibility, and confidentiality of users' personal data in the database is the main goal of information security. One of the biggest threats to the security of web applications is SQL injection. SQL injection refers to a type of attack that allows bypassing access control to web applications for direct database management, jeopardizing the application's information security. In this article, we consider the identification of SQL injections using artificial intelligence methods. For this task, machine learning and neural network models were developed using the following algorithms: Naive Bayesian Classifier, Support Vector Machine, Logistic Regression, Decision Tree, Random Forest, XGBoost, AdaBoost, and Deep Neural Networks. To assess the effectiveness of data classification using the developed models, the indicators of accuracy, precision, recall, and F1-score were used, which reached values in the range of 0.95-0.99.*

Keywords: *artificial intelligence, machine learning, neural networks, SQL injection.*

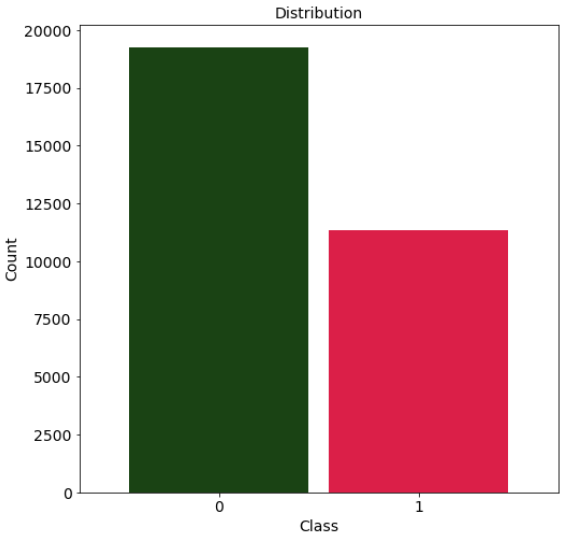
Кіріспе. Компьютерлік технологияның дамуымен адамдар интернет дәуіріне толық қадам басты, ал интернеттің дамуымен көптеген веб-қосымшалар пайда болды. Веб-қосымша деп веб-браузер арқылы кез келген операциялық жүйеде қолданушыларға интерфейсті қамтамасыз ететін және түрлі күнделікті негізгі қызметтерді өңдейтін программалық жүйені айтамыз. Веб-қосымшалар пайдаланушылардың жеке деректері жиналады, сондықтан да веб-қосымшалар мен веб сайттардың ақпарат қауіпсіздігінің маңыздылығы барған сайын артуда. Деректер қорында сақталған құнды мәліметтер зиянкестердің кибершабуылдары нысанасына айналды [1], әсіресе қауіпсіздігі әлсіз жобаланған веб-қосымшаларда кибершабуылдар жиі кездеседі. Сондықтан түрлі шабуылдар мен қауіптер көбейіп келеді. Өте танымал шабуыл әдісінің бірі – SQL инъекция. Бұл веб-қосымшалардың дерек қорынан құпия мәліметтерді алу, өзгерту немесе өшіру мақсатында сұраулардағы осалдықтарды пайдаланатын әдіс болып табылады [2]. Оның табыстылығы да басқалармен салыстырмалы түрде қарағанда жоғары болып келеді. SQL инъекциясы кіру пәрмендері сияқты сайт рұқсат еткен сұрауларды қолданып, зиянды SQL пәрмендерін дерекқор серверлеріне жіберу арқылы орындалады. Сұрау логикасын өзгертетін ерікті SQL кодын енгізуге негізделген [3]. Ұсынылған мақаланың негізгі мақсаты нейрондық желілері мен машиналық оқыту алгоритмдерін қолдану арқылы SQL инъекциясын анықтау. Нәтижеге қол жеткізу үшін SQL инъекциясының қауіптерін анықтау алгоритмін құру және SQL инфекциясына негізделген қауіптерді анықтау үшін жасанды интеллект әдістерін қолдана отырып модель құру тапсырмалары қойылды.

Әдеби шолу. SQL инъекциясын анықтау үшін [4] мақаласында авторлар LSTM моделін пайдаланғанын және LSTM маңызы жоғары екенін көрсеткенін анықтадық. SQL инъекциясын анықтау мақсатында LSTM моделін оқу деректер жинағы ретінде URL мекенжайынан векторлық енгізуді қолданған. SQL инъекциямен қатар, XSS шабуылдар мен фишинг сайттарын анықтауға LSTM моделін пайдалану мүмкіндігін сипаттап, талдау жасаған. Ал [5] мақаласында зерттеушілер SQL инъекциялық шабуылын анықтау үшін CNN-BiLSTM тәсілін, сондай-ақ әртүрлі машиналық оқыту алгоритмдерін суреттеген. CNN-BiLSTM тәсілі сипатталған басқа машиналық оқыту алгоритмдерімен салыстырғанда шамамен 98 % дәлдікті көрсетті. Келесі [6] мақаласының авторлары SQL инъекциясын анықтау әдістеріне талдау жасады. Деректер жиынында түрлі машиналық оқыту әдістерді және терең нейрондық желілерді пайдалану арқылы көптеген эксперименттер жүргізіп, SQL инъекция мәлімдемелерін анықтау тиімділігін салыстырды. Нәтижесінде зияткерлік көлік

жүйесіндегі шабуылдарды анықтауға арналған SQL инъекция әдісі жоғары нәтиже көрсетеді, дәлдікті дәлелдейді және жалған негативтер мен жалған позитивтердің жиілігін төмендетеді. SQL инъекциясын анықтау үшін Elastic-Pooling CNN пайдалану әдісін [7] жұмысында ұсынған және оны дәстүрлі анықтау әдістерімен салыстыру жүргізген. Бұл әдіс тіркелген 2D матрицасын деректерді қысқартпай шығара алады, сондай-ақ веб-программаның SQL инъекцияларын тиімді анықтады. Зерттелген келесі [8] мақаласының зерттеушілері SQL инъекция шабуылдары мәселесін шешу мақсатында нейрондық желілерге негізделген тиімді үлгісін ұсынған. Ұсынылған модель үш негізгі элементті қамтиды: URL генераторы, URL классификаторы және нейрондық желілер моделі. URL генераторы және URL классификаторы нейрондық желілер үлгісін тестілеу, тексеру және оқытудың үш қадамы үшін қажетті зиянды және қауіпсіз URL мекенжайларын қамтамасыз ету үшін пайдаланылады. Түсірілген нәтижелерді ескере отырып, SQL инъекция шабуылын анықтауға арналған нейрондық желілер негізіндегі модель дәлдік, шынайы оң жылдамдық және жалған оң жылдамдық тұрғысынан жақсы жұмыс істейді. Соңғы зерттелген [9] мақаласында зерттеушілер журнал файлдарын талдау арқылы SQL инъекциясын анықтауға арналған ATTAR деп аталатын модель үлгісін ұсынды. Эксперимент нәтижесі бойынша ATTAR грамматикалық және мінез-құлық мүмкіндіктерін пайдалана отырып, төмен жалған теріс және жалған оң көрсеткіштермен SQL инъекциясын тиімді анықтай алатынын көрсетеді. Машиналық оқытудың бірнеше алгоритмі пайдаланылған: Naive Bayesian, SVM, ID3, Random Forest, and K-means. Осы алгоритмдердің ішінде ID3 және Random Forest алгоритмдері арқылы жоғары нәтижеге қол жеткізген.

Материалдар және зерттеу әдістері. Бұл мақалада SQL инъекциялық шабуылдарын анықтау және алдын алу үшін нейрондық жүйелер мен машиналық оқыту моделдерінің құрылысы талқыланады. Модельді әзірлеу кезеңдері келесі қадамдарды қамтиды: деректер жиынын дайындау, алдын ала өңдеу, векторлау және нейрондық жүйелер мен машиналық оқыту алгоритмдері бойынша жіктеу.

Деректер жиынтығы. SQL инъекциясы арқылы шабуылдар туралы мәліметтер мен әр түрлі веб-сайттардан қауіпсіз деректер ағыны жиналды [10]. Деректер жиынтығында мәтіндік белгі мәні (Sentence) және тегтер (Label) бар 11341 қауіпті SQL командалары мен 19268 қауіпсіз ұсыныстар бар. Класс бойынша бөлу 1-суретте көрсетілген.



1-сурет. Кластар бойынша бөлу

Белгілерді алу. Деректер жиынында Sentence өрісінің мәтіндік мәнінен белгілерді алу керек. Ол үшін кең таралған term frequency – inverse document (tf-idf) әдісі қолданылады [11]. Бұл әдіс векторлаудың ең тиімді және жиі қолданылатын әдістерінің бірі болып табылады. Метрика tf (термиялық жиілік – сөз жиілігі) және idf (инверсиялық құжат жиілігі – құжаттың кері жиілігі) екі құрамдас бөлігін қамтиды.

tf-белгілі бір сөздің пайда болу санының құжат сөздерінің жалпы санына қатынасы. Осылайша, жеке құжаттағы сөздің маңыздылығы бағаланады (1):

$$tf(t,d) = \frac{n_i}{\sum_{i=1}^{k} n_i}, \quad (1)$$

мұндағы n_i 5 – құжаттағы сөздің кездесу саны, ал бөлгіш – құжаттағы сөздердің жалпы саны.

idf – жинақ құжаттарында белгілі бір сөздің кездесетін жиілігіне кері шамасы. idf пайдалану жиі қолданылатын сөздердің салмағын азайтады. Берілген құжаттар жинағындағы әрбір бірегей сөз үшін бір ғана idf мәні бар (2):

$$idf(t,D) = \log \frac{|D|}{|(d_i \supset t_i)|}, \quad (2)$$

мұндағы $|D|$ – корпустағы құжаттардың саны; $|(d_i \supset t_i)|$ – t_i-тен тұратын құжаттар саны. tf және idf мәндерін есептегеннен кейін екі бөлік те көбейтіледі (3) [8]:

$$tf\text{-}idf = tf \times idf \quad (3)$$

Машиналық оқыту алгоритмдері және нейрондық желілер. Функцияларды шығару кезеңінен кейін алынған деректер бірнеше машиналық оқыту алгоритмдері бойынша жіктеледі: Наив Байес классификаторы (Naïve Bayes) [12], қолдау векторлық машинасы (Support vector machine) [13], логистикалық регрессия (Logistic regression) [14], шешім ағашы (DT) [15], кездейсоқ орман (RF) [16], XGBoost [17] және AdaBoost [18] және терең нейрондық желі (Deep neural network) [19].

NB мәтінді жіктеу үшін ең қарапайым және жиі қолданылатын машиналық оқыту алгоритмдерінің бірі болып табылады, деректердің тәуелсіздігі туралы күшті болжамдары бар Байес теоремасына негізделген ықтималдық тәсілді пайдаланады (4):

$$P(c \mid z) = \frac{P(c) \times P(z \mid c)}{P(z)}, \quad (4)$$

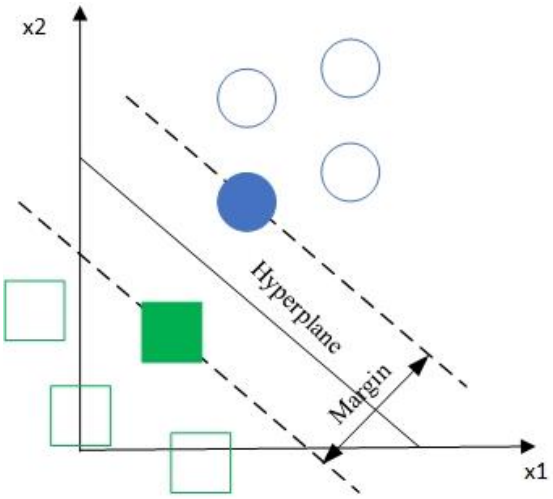
мұндағы $z = \{x_1, x_2, ..., x_n\}$, x_i – салмағы i^{th} мәтін белгісіндегі сөздер z, және c – құжат класы.

SVM – танымал Машиналық оқыту алгоритмі [20]. Бұл алгоритм гиперпланмен бөлінген белгілер кеңістігімен жұмыс істейді. Бұл жағдайда жақсы бөлуге гиперплан арқылы қол жеткізіледі, ол екі сыныптың ең жақын оқу деректер нүктелеріне (функционалдық шекара деп аталады) ең үлкен қашықтыққа ие, өйткені шекара неғұрлым үлкен болса, жіктеуіш қатесі соғұрлым төмен болады.

Гиперплан теңдеуі келесі түрде жазылады (5):

$$y_i(\vec{w} \times \vec{x} + b) \geq 0, \quad (5)$$

мұндағы $\vec{x} = (x_1, x_2, ..., x_n)$ – ерекшелік векторы; $\overrightarrow{w} = (w_1, w_2, ..., w_n)$ – масштаб векторы; y_i – шығу мәні; b – ауысым. Егер мән нөлден үлкен немесе оған тең болса, ол оң класқа жатады. Әйтпесе, ол теріс класқа жатады (2-сурет).

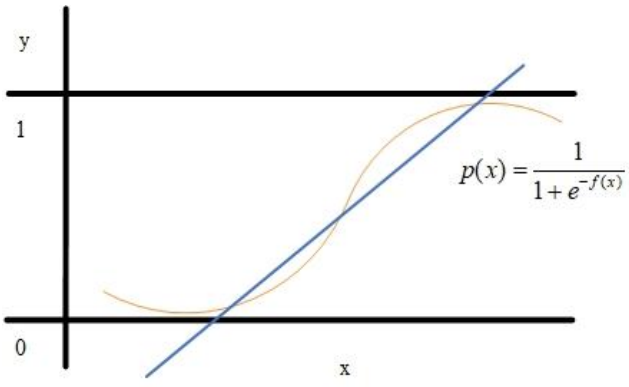


2-сурет. Гипержазықтықты бөлу

Логистикалық регрессия логистикалық функцияны (6) пайдалана отырып, тәуелсіз айнымалының [0,…,1] интервалында болу ықтималдығын болжау арқылы модельді жіктейді:

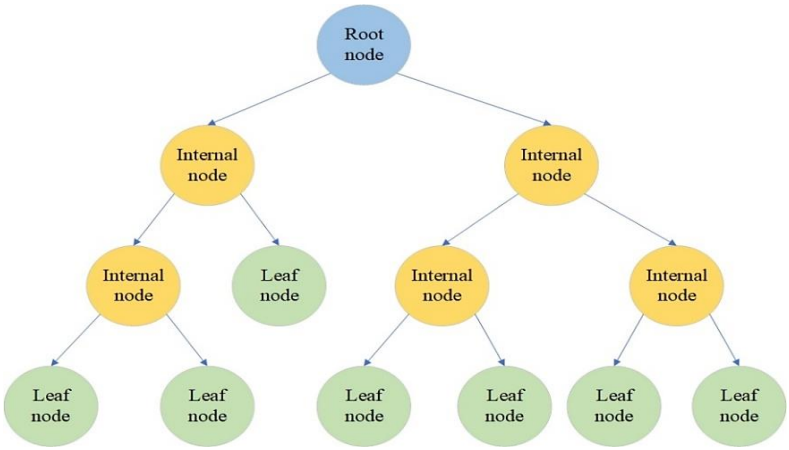
$$p(x) = \frac{1}{1 + e^{-f(x)}}, \tag{6}$$

мұндағы $f(x) = w_0 + w_1 x_1 + ... + w_r x_r$ – сызықтық жіктеуіш функциясы, $\vec{x} = (x_1, x_2, ..., x_n)$ – мүмкіндік векторы; $\overrightarrow{w} = (w_1, w_2, ..., w_n)$ – масштаб векторы. Логистикалық функция $p(x)$ ықтималдық мәндері 0-ден 1-ге дейінгі сигма тәрізді (3-сурет) түрінде болады. $p(x)$ мәні нөлге жақын болса, z құжаты бірінші сыныпқа жатады. Әйтпесе, екінші сыныпқа қойылады.



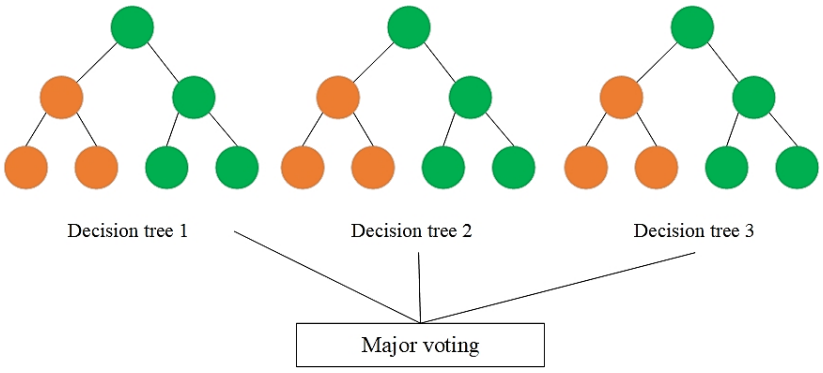
3-сурет. Логистикалық функция

Шешім ағашы – шешім қабылдау үшін ережелер жиынтығын пайдаланатын машиналық оқыту алгоритмі. Бұл әдіс барлық деректер нүктелері белгілі бір классқа жатқанға дейін белгілі бір сұрақтарға жауап бере отырып, мәліметтер жиынтығын белгілер бойынша бөлуге негізделген. Осылайша, әр сұраққа түйін қосып, ағаш тәрізді құрылым пайда болады (4-сурет).



4-сурет. Шешім ағашы

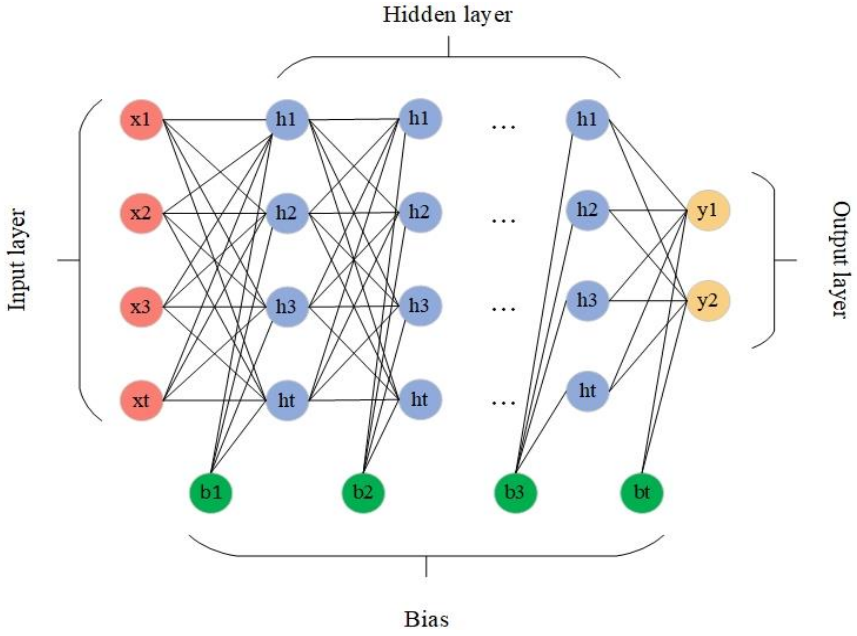
Кездейсоқ орман – ансамбльді оқыту тұжырымдамасына негізделген тағы бір танымал машиналық оқыту алгоритмі. Бұл тұжырымдама үлгі өнімділігін жақсарту үшін бірнеше классификаторларды біріктіруді қамтиды. Кездейсоқ орман бір шешім ағашын емес, тұтас қатарды қамтиды (5-сурет). Жіктеу есептерінде әрбір құжат бір-бірінен тәуелсіз барлық ағаштар бойынша жіктеледі. Шығаруда құжаттың классы барлық ағаштар арасындағы ең көп дауыс санымен анықталады.



5-сурет. Шешім ағашы

XGboost – күшейту принципін қолданатын машиналық оқытудың жетілдірілген алгоритмдерінің бірі. Ол жақсы өнімділікке ие және регрессия мен жіктеу мәселелерінің көпшілігін шешеді. Күшейту – жаңа үлгіде бұрынғы қателер жойылатын ансамбльдік әдіс. Оқытылған ансамбль болжамдарының ауытқулары әрбір итерацияда жаттығу жиынында есептеледі. Осылайша, оңтайландыру модельдің орташа ауытқуын азайта отырып, ансамбльге жаңа ағаш болжамдарын қосу арқылы орындалады. Бұл процедура қажетті қате деңгейіне немесе «ерте тоқтату» критерийіне жеткенше жалғасады.

Терең нейрондық желілер екі немесе одан да көп жасырын қабаттары бар нейрондық желілердің үлгісі болып табылады. Нейрондық желі кіріс деректерін қамтитын кіріс деңгейінен, нейрондар деп аталатын түйіндерді қамтитын жасырын қабаттардан және бір немесе бірнеше нейрондарды қамтитын шығыс қабаттан тұрады (6-сурет).



6-сурет. Терең нейрондық желі

Бұл жағдайда $x = x_1, x_2, ..., x_f$ кіріс векторы, $w_1, w_2, ..., w_i$ - әр деңгейдің қосылу салмағы, ал $b_1, b_2, ..., b_i$ – орын ауыстыру векторы. l_2 -тен l_{n-1} -ке дейінгі деңгейлер жасырын қабаттарды құрайды, ал l_n шығыс деңгейі тиісті $y_1, y_2, ..., y_m$ шығыс векторымен ұсынылған. Жасырын және шығу қабаттарының элементтері нейрондар деп аталады. Олар кіріс пен жауап айнымалысы арасындағы сызықтық емес функционалды картаға жауап беретін белсендіру функцияларымен ұсынылған.

Нәтижелері. Векторланған деректер Python – Scikit-learn кітапханасын қолдана отырып, Машиналық оқыту алгоритмдері арқылы жіктеледі. Нәтижелер Matplotlib кітапханасы мен Seaborn көмегімен де көрсетілді. Деректерді жіктеудің тиімділігін бағалау үшін келесі көрсеткіштер қолданылды: дұрыстығы (accuracy), дәлдігі (precision), толықтығы (recall) және F-өлшемі (F-score). Олар келесі формулалармен өрнектеледі (8-12) [21-23]:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \tag{8}$$

$$precision = \frac{TP}{TP + FP}, \tag{9}$$

$$recall = \frac{TP}{TP + FN}, \tag{10}$$

$$F1\_score = 2\frac{precision \times recall}{precision + recall}, \tag{11}$$

мұндағы шынайы оң (True positive – TP) оң көңіл – күй классы дұрыс жіктеген тест данасын көрсетеді; шынайы теріс (True negative-TN) теріс көңіл-күй классы дұрыс жіктеген тест данасын көрсетеді; жалған оң нәтиже (False positive – FP) оң көңіл-күй классы қате жіктеген тест данасын көрсетеді; жалған теріс нәтиже (False negative-FN) теріс көңіл-күй классы қате жіктеген сынақ данасын көрсетеді.

Алгоритмдерді тиімді бағалау үшін пайдалы графикалық шара да бар. Ол Area under the curve – Receiver operating characteristics (AUC–ROC) деп аталады. AUC-ROC жіктеу нәтижелерін визуализациялау үшін өте ыңғайлы. Ол нөлден бірлікке дейінгі ось жазықтығындағы қисық астындағы ауданды білдіреді. Жазықтық осьтері келесі формулалар бойынша есептелетін True Positive Rate және False Positive Rate мәндерін көрсетеді (12, 13):

$$TruePositiveRate = \frac{TP}{TP + FN} \tag{12}$$

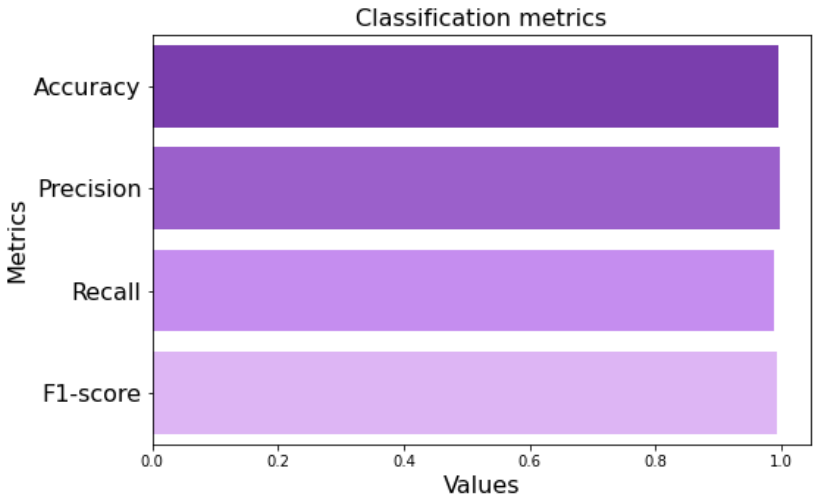
$$FalsePositiveRate = \frac{FP}{FP + TN} \tag{13}$$

Аудан мәні неғұрлым үлкен болса, классификация үлгілерінің тиімділігі соғұрлым жоғары болады. Деректерді жіктеу нәтижелері 1-кестеде берілген.

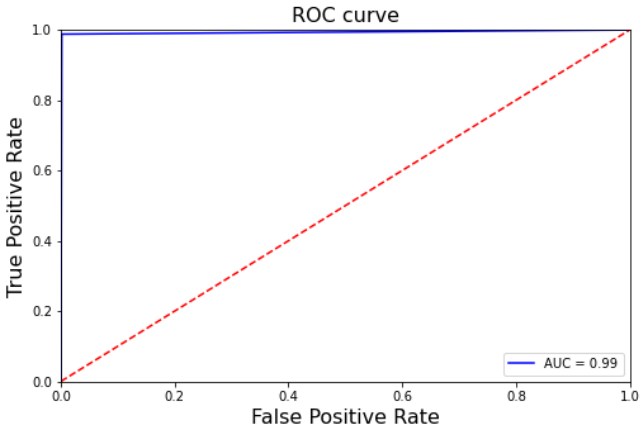
1-кесте. Мәтіндердің екілік жіктелуі

Классификатор	NB	SVM	LR	DT	RF	XGBoost	AdaBoost	Нейрондық желілер	Орташа мәні
Accuracy	0,96	0,99	0,99	0,99	0,99	0,98	0,99	0,99	0,99
Precision	0,97	0,99	0,99	0,99	0,99	0,99	0,99	0,99	0,99
Recall	0,93	0,98	0,98	0,99	0,98	0,97	0,98	0,98	0,97
F-score	0,95	0,99	0,99	0,99	0,99	0,98	0,99	0,99	0,98
Орташа мәні	0,95	0,99	0,99	0,99	0,99	0,98	0,99	0,99	

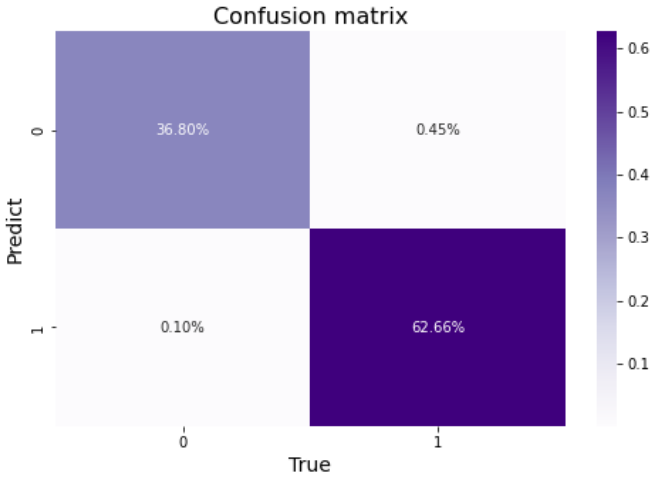
Жіктеу модельдерін бағалау графиктері гистограмма, AUC-ROC қисық және қате матрицасы 7-9 суреттерде көрсетілген.



7-сурет. Классификациялық гистограммалар



8-сурет. AUC-ROC қисығы



9-сурет. Қате матрицасы

Қорытынды. Қазіргі таңда интернет пен веб-қосымшалар мүмкіндіктерін кеңінен пайдалану ақпараттық қауіпсіздік мәселесін шиеленістіріп отыр. Ұсынылған мақалада кең таралған SQL инъекция шабуылы қарастырылған. Бұл шабуыл түрі әдетте деректер қорымен әрекеттесетін жүйелердегі қателерді SQL пәрмендері арқылы орындалатын кибершабуыл. Эксперименттік деректердің нәтижелері бойынша машиналық оқыту алгоритмдері мен нейрондық желі SQL инъекциялық деректер жинағын жіктеуде тамаша нәтижелер көрсетіп, дұрыстығы, дәлдігі, толықтығы және өлшемі бойынша 0,95 - 0,99 мәндеріне жеткенін көруге болады.

Әдебиеттер тізімі

1. Жұмабекова А.Т., Усатова О.А. Нейрондық желілер арқылы интернет қауіптерді анықтау // Қазіргі әлемдегі ғылым және білім: XXI ғасырдағы қауіптер" IX Халықаралық ғылыми-практикалық конференциялар, желтоқсан 2021 ж.
2. К.В. Смирнова, А.О. Смирнов, В.М.Плотников Веб-жүйелерге шабуылдарды жіктеу тапсырмалары үшін машиналық оқытудың қолданылуы. 3 бөлім. // Технологиялық және бизнес – процестерді автоматтандыру, Одесса ұлттық тамақ технологиялары академиясы, Одесса. 10 Том, 1/2018 шығарылым.

3. Ю.Д. Шабалин, В.Л. Елисеев. XSS осалдықтары мен SQL инъекцияларын автоматтандырылған іздеу кезінде желілік хост мінез-құлқындағы ауытқуларды анықтауға арналған нейрондық желі алгоритмін зерттеу // Ақпараттық технологиялар, механика және оптика ғылыми-техникалық хабаршысы, 2016, 16 том, № 2.
4. Джоши Падма, Н. Равишанкар, М. Б. Раджу, Н. Ч. Рави. және т. б. LSTM көмегімен SQL инъекциялық шабуылдарымен хирургиялық соққылар // Үндістанның информатика және инженерия журналы (IJCSE). 13 Том, № 1, 2022 жылғы қаңтар-ақпан.
5. Ганди, Н. SQL инъекциялық шабуылдарын анықтауға арналған CNN-BiLSTM әдісі. // 2021 жылғы есептеу интеллектісі және білім экономикасы жөніндегі халықаралық конференция материалдарында (ICCIKE), Дубай, Біріккен Араб Әмірліктері, 17-18 наурыз 2021 ж; 378-383 б.
6. Ли, К.; Ван, Ф.; Ван, Дж.; Ли, У. Ақылды көлік жүйесіне арналған LSTM негізіндегі SQL инъекциясын анықтау әдісі // IEEE Trans. Veh. Технол. 2019, 68, 4182-4191.
7. Син Се, Чунхуэй Рен, Ишэн Фу, Джи Сю, Джинхун Го, CNN серпімді пулына негізделген веб-қосымшаларға арналған SQL инъекциясын анықтау // IEEE Access, 2019.
8. Н.М. Шейхканлу. SQL инъекциялық шабуылын анықтау үшін нейрондық желілерді пайдалану. // Ақпарат және желі қауіпсіздігі бойынша 7-ші халықаралық конференция материалдары, Просидинг. Қыркүйек 2014, 318-323 беттер.
9. Хонцан Гао, Цзинвэнь Чжу, Лэй Лю Цзин, Сю Янфэн Ву, Ао Лю. Грамматикалық үлгіні тану және қол жеткізу әрекетін талдау арқылы SQL инъекциялық шабуылдарын анықтау. // 2019 жылғы IEEE халықаралық энергетикалық интернет конференциясы (ICEI).
10. Сайед Саклейн Хусейн Шах. SQL инъекцияларына арналған мәліметтер жиынтығы // URL: https://www.kaggle.com/datasets/syedsaqlainhussain/sql-injection-dataset
11. Т.Н. Манжунат, Дипа Йогиш, С. Махалакшми, Х.К. Йогиш. Векторлау тәсілін және статистикалық баллдық әдісті қолдана отырып, сұрақтарға жауап берудің интеллектуалды жүйесі // Бүгінгі материалдар: Просидинг, 2021.
12. Мишра, Шайлендра, Айман Альбаракати, Сунил Кумар Шарма. Машиналық оқытуды қолдана отырып, интернет заттарына арналған киберқауіптерді талдау // *Energies* 10, Нөмір 12: 2673. 2022 жыл.
13. Стрижек, Шимон, Марек Натканец. LSTM, IF және SVM көмегімен желілік трафикті талдау негізінде интеллектуалды желілердегі интернет қауіптерін анықтау // *Energies* 16, нөмір 1: 329. 2023 жыл.
14. Ф. Аниш МОН Соломон, г. Уинстер Сатианесан, Р. Рамеш, Логистикалық регрессияға деген сенім – регрессиялық талдауды қолдана отырып, интернет заттарына деген сенім моделі // компьютерлік жүйелер туралы ғылым және инженерия, 44 том, №2, 1125-1142 ББ, 2023.
15. Цзекун Ниу, Цзинфэн Сюэ, Дачэн Куй, Али Ван, Джун Чжэн, Хунфэй Чжу. Iiot-те зиянды бағдарламаларды анықтауға арналған шифрланған трафикті бейімделген онлайн талдауға негізделген жаңа тәсіл // Ақпараттық ғылымдар, 601 том, 2022, 162-174 ББ.
16. Карюкин В.И., Жұмабекова А.Т., Есенжанова С.Б. Әлеуметтік медианы талдауға арналған машиналық оқыту және нейрондық желі әдістемелері. ACM халықаралық конференциясының жалғасы сериясы, инженерия және MIS, // ICEMIS 2020 жөніндегі халықаралық конференция; есептеу техникасы қауымдастығы (Канада), IITU Манас көшесі 34/1 Алматы; Қазақстан, 1-7 б, 2020 ж.
17. Шахин М., Чен Ф. Ф., Хоссейнзаде А. және т. б. заттар интернетінің өнеркәсіптік құрылғыларында кибершабуылдарды анықтауға арналған терең гибридті оқыту моделі. // Халықаралық озық өндіріс технологиялары журналы 123, 1973-1983 (2022).
18. П.Л. Индрасири, М.Н. Халгамуге, А. Мұхаммед, Фишингтік URL мекен-жайларын сүзуге арналған машиналық оқытудың сенімді ансамбльдік моделі: Кеңейтілген кездейсоқ градиентті дауыс беру классификаторы (ERG-SVC), // IEEE Access, 9 том, 150142-150161 ББ, 2021.
19. Бхандари, Гуру, Андреас Лит, Андрей Шалагинов, Тор-Мортен Гренли. 2023 жыл. Smart IoT экожүйесіндегі кибершабуылдарды анықтауға арналған терең нейрондық желілерге негізделген таратылған аралық бағдарламалық құрал: жаңа құрылым және өнімділікті бағалау тәсілі // Электроника 12, № 2: 298.
20. Усатова О.А., Жұмабекова А.Т., Мэтсон Э., Карюкин В.И., Ілесова Б.Е. «Ақпараттық ресурстарға төнетін қауіп түрлері және оларды машиналық оқытуды әдістерін қолдану арқылы анықтау» // Известие НАН РК № 6, Алматы, 2021 г. С. 48-58.
21. О. Усатова, А. Жұмабекова, Е. Бегимбаева, Э. Мэтсон, Н. Усатов, Машиналық оқыту алгоритмдерін қолдана отырып, ddos шабуылдарының кешенді жіктелуі, // Компьютерлер, материалдар және континуум, 73-том, №1, 577-594, 2022 ББ.
22. Мутанов Г., Карюкин В., Мамыкова Ж. Машиналық оқыту алгоритмдерін қолдана отырып, осы әлеуметтік желілердің көңіл-күйін көп класты талдау. // СМС-компьютерлер, материалдар және жалғасы. 2021; 69(1): 913-930.

23. Карюкин В., Мутанов Г., Мамыкова З. және т.б. пайдаланушылардың пікірін және оның қоғам үшін рөлін бақылау үшін ақпараттық жүйені әзірлеу туралы // Үлкен деректер журналы 9, 110 (2022).

References

1. Zhumabekova A.T., Ussatova O.A. Detection of neural networks based on internet security. IX International Scientific and practical conference // Science and education in the modern world: threats in the XXI century, December 2021.
2. K.V. Smirnova, A.O. Smirnov, V.M.Plotnikov Applicability of machine learning for classification tasks of attacks on web systems. Part 3. // Automation technological and business – processes, Odessa National Academy of Food Technologies, Odessa. Volume 10, Issue 1 /2018.
3. Y.D. Shabalin, V.L. Eliseev. Investigation of a neural network algorithm for detecting anomalies in the behavior of a network host during automated search for XSS vulnerabilities and SQL injections. // Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics, 2016, volume 16, № 2.
4. Joshi Padma, Dr. N. Ravishankar, Dr. M. B. Raju, N.Ch. Ravi. Joshi Padma N et al. Surgical striking SQL injection attacks using LSTM // Indian Journal of Computer Science and Engineering (IJCSE). Vol. 13 No. 1 Jan-Feb 2022.
5. Gandhi, N. A CNN-BiLSTM based Approach for Detection of SQL Injection Attacks. // In Proceedings of the 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 17–18 March 2021; pp. 378–383.
6. Li, Q.; Wang, F.; Wang, J.; Li, W. LSTM-Based SQL Injection Detection Method for Intelligent Transportation System // IEEE Trans. Veh. Technol. 2019, 68, 4182–4191.
7. Xin Xie, Chunhui Ren, Yusheng Fu, Jie Xu and Jinhong Guo, SQL injection detection for web applications based on Elastic-Pooling CNN // IEEE Access, 2019. DOI 10.1109/ACCESS.2019.2947527,
8. N. M. Sheykhkanloo. Employing Neural Networks for the Detection of SQL Injection Attack. // Proceedings of the 7th International Conference on Security of Information and Networks. September 2014 Pages 318-323.
9. Hongcan Gao, Jingwen Zhu, Lei Liu Jing, Xu Yanfeng Wu, Ao Liu. Detecting SQL Injection Attacks Using Grammar Pattern Recognition and Access Behavior Mining. // 2019 IEEE International Conference on Energy Internet (ICEI).
10. Syed Saqlain Hussain Shah. SQL injection dataset. // URL: https://www.kaggle.com/datasets/syedsaqlainhussain/sql-injection-dataset
11. T.N. Manjunath, Deepa Yogish, S. Mahalakshmi, H.K. Yogish. Smart question answering system using vectorization approach and statistical scoring method. // Materials Today: Proceedings, 2021.
12. Mishra, Shailendra, Aiman Albarakati, and Sunil Kumar Sharma. 2022. Cyber Threat Intelligence for IoT Using Machine Learning // *Processes* 10, no. 12: 2673.
13. Stryczek, Szymon, and Marek Natkaniec. 2023. Internet Threat Detection in Smart Grids Based on Network Traffic Analysis Using LSTM, IF, and SVM // *Energies* 16, no. 1: 329.
14. F. Anish Mon Solomon, G. Winster Sathianesan and R. Ramesh, Logistic regression trust–a trust model for internet-of-things using regression analysis, // Computer Systems Science and Engineering, vol. 44, no.2, pp. 1125–1142, 2023.
15. Zequn Niu, Jingfeng Xue, Dacheng Qu, Yong Wang, Jun Zheng, Hongfei Zhu. A novel approach based on adaptive online analysis of encrypted traffic for identifying Malware in IIoT, // Information Sciences, vol. 601, 2022, pp. 162-174.
16. Karyukin V.I., Zhumabekova A.T., Esenzhanova S.B. Machine learning and neural network methodologies of analyzing social media. ACM International Conference Proceeding Series14 September 2020, International Conference on Engineering and MIS, ICEMIS 2020; Association for computing machinery (Canada), IITU Manas Street 34/1 Almaty; Kazakhstan, с. 1-7, 2020 г. DOI:10.1145/3410352.3410739.
17. Shahin, M., Chen, F.F., Hosseinzadeh, A. et al. A deep hybrid learning model for detection of cyber attacks in industrial IoT devices // Int J Adv Manuf Technol 123, 1973–1983 (2022).
18. P.L. Indrasiri, M.N. Halgamuge and A. Mohammad, Robust Ensemble Machine Learning Model for Filtering Phishing URLs: Expandable Random Gradient Stacked Voting Classifier (ERG-SVC), // *IEEE Access*, vol. 9, pp. 150142-150161, 2021.
19. Bhandari, Guru, Andreas Lyth, Andrii Shalaginov, and Tor-Morten Grønli. 2023. Distributed Deep Neural-Network-Based Middleware for Cyber-Attacks Detection in Smart IoT Ecosystem: A Novel

Framework and Performance Evaluation Approach // *Electronics* 12, no. 2: 298.
20. Ussatova O.A., Zhumabekova A.T., Matson E., Karyukin V.I., Ilessova B.E. Types of threats to information resources and their identification using machine learning methods // Journal NAS RK No. 6, Almaty, 2021, pp. 48-58.
21. O. Ussatova, A. Zhumabekova, Y. Begimbayeva, E.T. Matson, N. Ussatov, Comprehensive ddos attack classification using machine learning algorithms // Computers, Materials & Continua, vol. 73, no.1, pp. 577–594, 2022.
22. Mutanov G, Karyukin V, Mamykova Zh. Multi-class Sentiment Analysis of Social Media Data with Machine Learning Algorithms. // CMC–Computers, Materials & Continua. 2021; 69(1): 913–930.
23. Karyukin, V., Mutanov, G., Mamykova, Z. et al. On the development of an information system for monitoring user opinion and its role for the public. // Journal of Big Data 9, 110 (2022 y.).