из разработанных алгоритмов при решении задачи выбора эффективного режима работы станции подогрева нефти нефтепровода Узень-Самара в пункте Атырау.

Ключевые слова: *нечеткое принятие решений, многокритериальная оптимизация, информация, функция принадлежности, принципы оптимальности, лицо, принимающее решение (ЛПР).*

Information about the authors

Yerbol Amangazovich Ospanov* – PhD, Associate Professor of the Department of Automation, Information Technology and Urban Planning; Shakarim Semey University; e-mail: 78oea@mail.ru.

Nurbek Suykimbekuly Turarbai – is a 2nd–year undergraduate student at the National Academy of Sciences Shakarim University of Semey.

Mert Bayraktar – is a PhD doctoral student at Akdeniz University, Turkey.

Авторлар туралы мәліметтер

Ербол Амангазұлы Оспанов – PhD, автоматтандыру, ақпараттық технологиялар және қала құрылысы кафедрасының қауымдастырылған профессоры; Семей қаласының Шәкәрім атындағы университеті; e-mail: 78oea@mail.ru.

Нұрбек Сүйікімбекұлы Тұрарбай – Семей қаласының Шәкәрім атындағы университеті 2 курс магистранты.

Мирт Байрактар – Акдениз университетінің PhD докторанты, Түркия.

Сведения об авторах

Ербол Амангазович Оспанов – PhD, ассоциированный профессор кафедры Автоматизация, информационные технологии и градостроительство; Университет имени Шакарима города Семей; e-mail: 78oea@mail.ru.

Нурбек Сүйкімбекұлы Турарбай – магистрант 2 курса; Университет имени Шакарима города Семей.

Мирт Байрактар – PhD докторант университета Агдениз, Турция.

Received 04.03.2024
Accepted 04.04.2024

DOI: 10.53360/2788-7995-2024-2(14)-6

ҒТАХР: 32.61.11

Н.Е. Рахимбай*, К.Б. Тусупова
Әл-Фараби атындағы Қазақ ұлттық университеті,
050040, Қазақстан Республикасы, Алматы қ., әл-Фараби даңғылы, 71
*e-mail: nazerke.rkh@gmail.com

ВЕБ-БЕТТЕРДЕГІ ЗИЯНДЫ ЖАРНАМАЛЫҚ БАҒДАРЛАМАЛАРДЫ АНЫҚТАУДА МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІН ПАЙДАЛАНУ

Аңдатпа: *Мақалада интернет қолданушыларының құпиялылығы мен қауіпсіздігіне айтарлықтай қауіп төндіретін веб-беттер арқылы зиянды жарнамалық бағдарламаларды тарату мәселесі қарастырылады. Веб-беттерге енгізілген зиянды жарнамалық бағдарламаларды анықтау және бейтараптандыру үшін машиналық оқыту алгоритмдерін қолдану. Деректерді өңдеу, белгілерді алу және жіктеу әдістеріне назар аудара отырып, машиналық оқыту зиянды бағдарламарларды анықтау процестерін қалай жақсарта алатынын егжей-тегжейлі талдайды. Зиянды және заңды жарнамалық мазмұнды ажыратудағы тиімділігін анықтау үшін машиналық оқытудың әртүрлі алгоритмдері, соның ішінде логистикалық регрессия, шешім ағаштары, кездейсоқ орман, аңғал Байес және ансамбльдік әдістер зерттеледі.*

Зиянды және қауіпсіз жарнама модульдері туралы деректерді қамтитын оқыту және тест үлгілерін құру әдістемесі сипатталған. Зиянды мінез-құлықтың жасырын үлгілерін анықтау үшін машиналық оқытудың әртүрлі тәсілдері, соның ішінде мұғаліммен, мұғалімсіз оқыту және терең оқыту әдістері талданады. Зерттеу нәтижелері машиналық оқыту алгоритмдерін қолдану зиянды жарнамалық бағдарламаларды жоғары дәлдікпен анықтауға мүмкіндік беретінін көрсетеді, бұл тиімдірек киберқауіпсіздік құралдарын әзірлеуге негіз бола алады. Сондай-ақ, қолданыстағы

әдістердің ықтимал мәселелері мен шектеулері талқыланады және машиналық оқыту арқылы зиянды жарнамалық бағдарламаларды анықтау бойынша қосымша зерттеулер жүргізу үшін бағыттар ұсынылады.

Түйін сөздер: *Машиналық оқыту, зиянды бағдарламалар, киберқауіпсіздік, жасанды интеллект, веб-беттер.*

Кіріспе

Веб-технологиялар күнделікті өмірде басты рөл атқаратын қазіргі цифрлық әлемде деректердің қауіпсіздігі маңызды болып табылады. Зиянды жарнамалық бағдарламалар интернет пайдаланушыларының құпиялылығы мен қауіпсіздігіне қауіп төндіретін қауіптердің бірі болып табылады. Бұл бағдарламалар пайдаланушы тәжірибесіне кедергі келтіріп қана қоймайды, сонымен қатар жеке ақпарат пен қаржылық деректерді ұрлауды қоса алғанда, ауыр зардаптарға әкелуі мүмкін. Осыны ескере отырып, зиянды жарнамаларды табу және алдын алу киберқауіпсіздік саласындағы зерттеушілер мен әзірлеушілер үшін бірінші кезектегі міндет болып табылады [1].

Машиналық оқыту, үлкен көлемдегі деректерді өңдеу және күрделі заңдылықтарды анықтау қабілетімен, зиянды бағдарламалық жасақтамамен күресудің қуатты құралы болып табылады. Зиянды жарнамалық бағдарламаларды веб-беттерде анықтау үшін машиналық оқыту алгоритмдерін қолдану тиімдірек және автоматтандырылған қауіпсіздік жүйелерін құрудың жаңа перспективаларын ашады [2].

Материалдар мен тәсілдер.

Машиналық оқыту кибершабуылдарды анықтау мен алдын алудың тиімді құралын қамтамасыз ете отырып, веб-беттердегі зиянды бағдарламалармен күресу үшін қолданылады. Бұл сала нақты уақыттағы қауіптерді тану және бейтараптандыруда машиналық оқыту технологияларын киберқауіпсіздікпен біріктіреді. Машиналық оқыту зиянды бағдарламалардың болуын көрсететін ауытқуларды анықтау үшін веб-трафик деректерінің үлкен көлемін талдауға көмектеседі [3]. Алгоритмдер қауіпсіз веб-беттерді зиянды беттерден ажыратуды үйрену арқылы қалыпты мазмұнының мысалдарын қамтитын мәліметтерден үйренеді [4].

1 кесте – Модельді құруда пайдаланатын машиналық оқыту алгоритмдеріне талдау

Алгоритм	Артықшылығы	Кемшілігі	Мүмкіндігі	Қауіпі
Логистикалық регрессия	Түсінікті және орындауға оңай	Сызықтық емес деректермен жұмыс істеуде нашар	Екілік классификациялау мәселелері үшін жақсы	Модельдің шектеулі күрделілігі
Шешім ағашы (Decision Tree)	Интерпретациялауға және визуализациялауға оңай	Артық сәйкестендіруге бейім	Күрделі жүйелерде пайдалануға болады	Деректердегі өзгерістерге сезімтал.
Наивті Байес (Naive Bayes)	Үлкен деректер жиынтығымен жақсы жұмыс істейді	Сипаттардың тәуелсіздігі туралы болжам жасайды	Мәтіндік классификацияда тиімді	Байланысты сипаттармен тиімсіз болуы мүмкін
Кездейсоқ ормандар	Жоғары дәлдік, артық сәйкестендіруге төзімді	Есептеу құны жоғары болуы мүмкін	Көптеген мәселелерге қолдануға болады	Көп ресурс қажет етуі мүмкін
Градиенттік бустинг	Өте жоғары дәлдік	Орнату қиын және артық сәйкестендіруге бейім	Деректер анализіндегі бәсекелестікте тиімді	Үлкен деректермен жұмыс істеу кезінде баяу болуы мүмкін

Машиналық оқыту моделі – белгілі бір оқу мәселесінің қалай шешілетіндігінің математикалық көрінісі. Бұл машиналық оқыту алгоритмі деректерді талдайтын және олардан заңдылықтарды немесе тәуелділіктерді шығаратын оқу процесінің нәтижесі. Машиналық оқыту алгоритмдері ағымдағы қауіптерді анықтап қана қоймай, сонымен қатар тенденциялар мен мінез-құлық үлгілерін талдау арқылы ықтимал зиянды шабуылдарды болжай алады. Бұл шабуылдарды зиян келтірмес бұрын алдын алуға мүмкіндік береді. Көбінесе желілік мониторлар, кірудің алдын алу жүйелері және антивирустық бағдарламалар сияқты басқа киберқауіпсіздік жүйелерімен біріктіріліп, веб-беттердегі зиянды шабуылдарға қарсы кешенді қорғаныс жасайды [5].

Нәтижелер мен талқылаулар

Әрбір модель бірегей сипаттамаларға ие және әртүрлі деректер түрлерімен және тапсырмалармен жақсырақ жұмыс істей алады. Модельді оңтайлы таңдау тапсырманың нақты мақсаттары мен талаптарына байланысты.

1 суретте көрсетілген корреляциялық жылу картасында біз объектінің (мысалы, веб-сайттың) зиянды екенін білдіретін әртүрлі сипаттамалар мен мақсатты айнымалы арасындағы сызықтық байланыс дәрежесін көреміз.

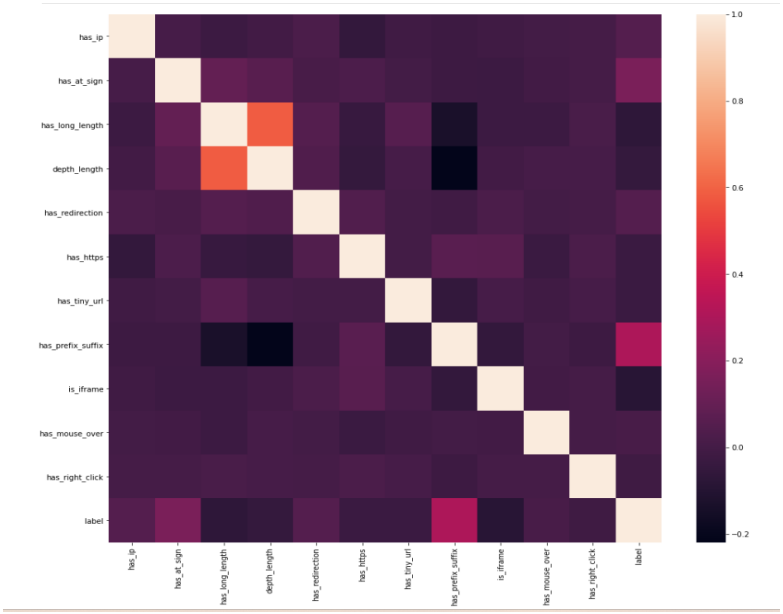
1 суреттегі картадан бірнеше бақылаулар жасауға болады:

1. «Label» – мен қиылысатын жол мен бағанда жеңілірек ұяшық бар «has_long_length» белгісі ерекше ерекшеленеді. Бұл ұзын URL мекенжайлары бар веб-сайттардың зиянды болуы ықтимал екенін көрсетеді, бұл қисынды, өйткені алаяқтар пайдаланушыларды алдау үшін күрделі және ұзын URL мекенжайларын жиі пайдаланады.

2. «Has_at_sign» белгісінде «label» қиылысында жеңіл ұяшық бар, бұл зиянды сайттармен оң корреляцияны көрсетуі мүмкін. URL мекен-жайында «@» белгісі болуы мүмкін, шабуылдаушылар заңды мекен-жайларға еліктеуге тырысатын фишингтік шабуылдар үшін қолданылады.

3. «Has_ip» және «depth_length» үшін «label» – мен корреляция да оң болып көрінеді, бірақ онша айқын емес. Бұл зиянды сайттар кейде IP мекенжайларын домендік атаулар ретінде пайдаланатынын және тереңірек URL құрылымын көрсетуі мүмкін.

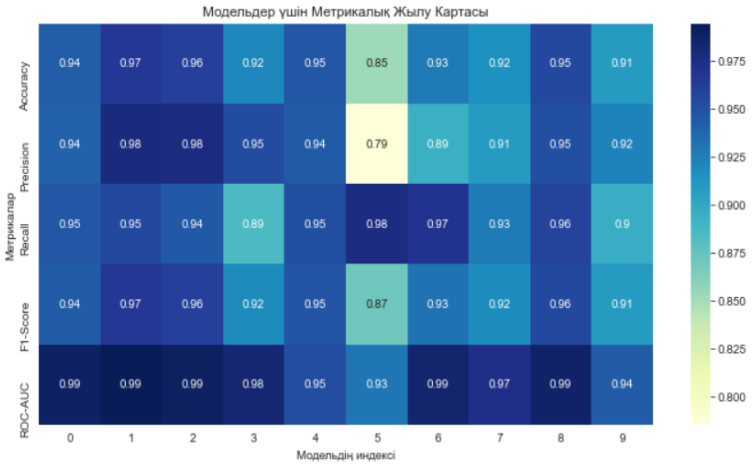
4. «Has_redirection», «has_https», «has_tiny_url», «has_prefix_suffix», «is_iframe», «has_mouse_over» және «has_right_click» сияқты басқа белгілер осы жылу картасындағы мақсатты айнымалымен аз айқын байланысты көрсетеді. Бұл олардың жіктеу тапсырмасы үшін маңызды емес екенін білдірмейді, өйткені тек Пирсонның корреляция коэффициенттері арқылы талдауда көрінбейтін сызықтық емес қатынастар немесе аралас әсерлер болуы мүмкін.



1 сурет – Корреляциялық жылу картасы

2 суретте көрсетілген жылу картасының әрбір ұяшығы белгілі бір модель үшін метрикалық мәнге сәйкес келеді. Метрикалық мәндер ұяшықтардың ішінде көрсетілген және ұяшықтың түсі картаның оң жағындағы түс шкаласына сәйкес осы мәннің шамасын көрсетеді. Шкаланың жоғарғы жағы (көк түстің күңгірт реңктері) метриканың жоғары мәндерін, ал төменгі бөлігі (көк түстің ашық реңктері) төменгі мәндерді көрсетеді.

Мақсатты көрсеткіштер мен жоба талаптарына сүйене отырып, белгілі бір тапсырма үшін қай модельдер жақсы жұмыс істейтінін анықтай алады. Мысалы, егер дәлдік жоба үшін ең маңызды болса, ең жоғары дәлдік көрсеткіші бар модельді таңдау керек. Егер дәлдік пен толықтықтың теңдестірілген үйлесімі қажет болса, F1-баллға қарау керек.



2 сурет – Модельдерге арналған метрика картасы

2 суретте бес метрика бойынша машиналық оқытудың тоғыз түрлі моделінің өнімділігін бейнелейтін жылу картасы көрсетілген: Accuracy, Precision, Recall, F1-Score және ROC-AUC.

Көрсеткіштерді талдау:

Accuracy (дәлдік): бұл дұрыс болжамдар санының (шын оң + шын теріс) болжамдардың жалпы санына қатынасы. Ең жоғары дәлдіктегі Модель 0,97, ал ең төменгісі 0,85.

Болжамдардың дәлдігі(Precision): оң нәтижелер ретінде болжанғандардың қай бөлігі дұрыс болғанын көрсетеді. Мұнда ең жақсы мән – 0,98, ал ең жаманы – 0,79.

Толықтығы(Recall): модель оң деп дұрыс анықтаған шынайы оң жағдайлардың қаншалықты үлесін өлшейді. Толықтықтың ең жоғары мәні – 0.98, бұл модель оң жағдайларды жақсы анықтайтындығын көрсетеді.

F1-балл (F1-Score): дәлдік пен толықтықтың Гармоникалық орташа мәні, осы екі көрсеткіштің де теңдестірілген бағасын береді. F1 ұпайлары 0.87-ден 0.97-ге дейін өзгереді.

ROC-AUC: Roc қисығының астындағы аймақ-бұл модельдің әртүрлі сыныптарды ажырата білу қабілетінің өлшемі. ROC-AUC-тің ең жақсы мәні – 0.99, ең жаманы – 0.93.

Талдау жүргізу үшін көрсеткіштердің идеалды мәндері тапсырманың нақты шарттарына байланысты әр түрлі болуы мүмкін екенін ескеру қажет. Мысалы, медициналық диагностикада оң жағдайларды жіберіп алмау үшін жоғары толықтық өте маңызды болуы мүмкін, тіпті егер бұл дәлдіктің төмендеуіне әкелсе де.

Жылу картасына сүйене отырып, талдаушы мақсатты көрсеткіштер мен жоба талаптарына сүйене отырып, белгілі бір тапсырма үшін қай модельдер жақсы жұмыс істейтінін анықтай алады. Мысалы, егер дәлдік жоба үшін ең маңызды болса, ең жоғары дәлдік көрсеткіші бар модельді таңдау керек. Егер дәлдік пен толықтықтың теңдестірілген үйлесімі қажет болса, F1-баллға қарау керек.

3-суреттегі график жіктеу тапсырмасында машиналық оқытудың әртүрлі модельдерінің өнімділігін көрсетеді, мүмкін зиянды бағдарламаны анықтауға байланысты, оны график қолданылатын контекстке сүйене отырып растауға немесе жоққа шығаруға болады.

– Логистикалық регрессия: F1-Score-дан басқа барлық көрсеткіштер бойынша тұрақты жоғары нәтижелерді көрсетеді, онда ол аздап төмендейді. Бұл модель дәлдік пен толықтық арасында жақсы теңдестірілген деп болжайды [6].

– Кездейсоқ орман: ROC-AUC метрикасы бойынша ең жоғары мәндерге ие, бұл оның сыныптарды ажырата білу қабілетін көрсетеді. Дәлдік пен толықтық көрсеткіштері де өте жоғары.

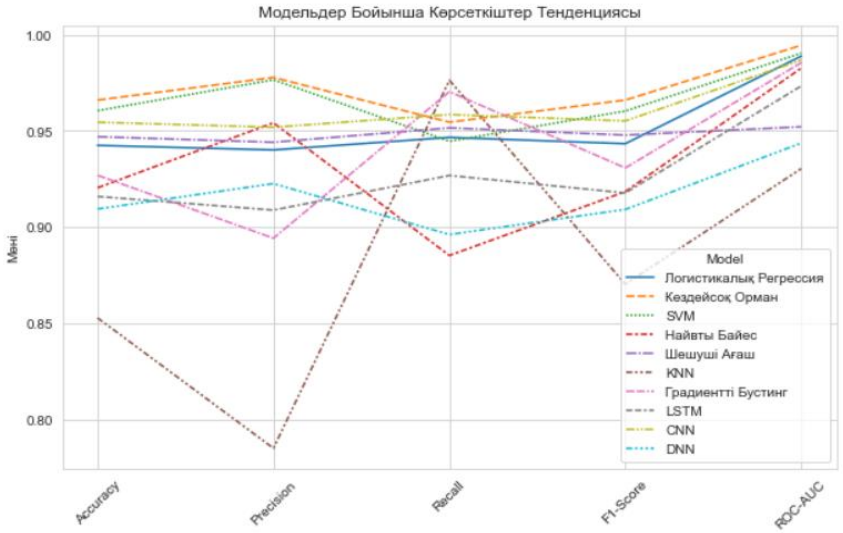
– SVM (анықтамалық векторлық Машина): барлық көрсеткіштер бойынша салыстырмалы түрде тұрақты, бірақ F1-Score және ROC-AUC-та айтарлықтай құлдырау бар, бұл оның дәлдік пен толықтық арасындағы тепе-теңдікті нашар басқаратынын көрсетеді.

– Аңғал Байес: дәлдік пен толықтықта жақсы нәтиже көрсетеді, бірақ оның F1-Score және ROC-AUC жақсарту мүмкіндігі туралы ойлануға мәжбүр етеді.

– Шешім ағашы және KNN (K әдісі-жақын көршілер): екі модель де ROC-AUC метрикасы бойынша төмен көрсеткіштерді көрсетеді, бұл зиянды бағдарламаны анықтауда маңызды.

– Градиентті бустинг: жалпы өнімділігі жақсы, бірақ оның сызығы ROC-AUC метрикасында аздап төмендеуді көрсетеді, бұл сыныптарды ажырату қабілетінің төмендеуін көрсетеді.

– LSTM, CNN (Конволюциялық нейрондық желі) және DNN (терең нейрондық желі): әдетте үлкен деректері бар күрделі тапсырмалар үшін немесе уақыт серияларын талдау (LSTM) үшін қолданылатын бұл үш модель әртүрлі өнімділікті көрсетеді. Әсіресе LSTM барлық көрсеткіштер бойынша қатты тербелістерді көрсетеді, бұл оқу деректер жиынтығының ерекшеліктеріне немесе модельдің параметрлеріне байланысты болады [7].



3 сурет – Модельдер бойынша көрсеткіштер тенденциясы

3-сурет бойынша зиянды бағдарламаны анықтаудың ең жақсы моделін таңдау үшін тапсырманың нақты талаптарын ескеру қажет. Егер жалған оң позитивтерді азайту маңызды болса (мысалы, пайдаланушы қауіпсіз бағдарламаларды бұғаттауға байланысты қолайсыздықты сезінбеуі керек), дәлдікке назар аудару керек.

Терең оқыту алгоритмдері-бұл оқуға және болжам жасауға немесе деректерге негізделген шешім қабылдауға қабілетті модельдер жасау үшін қолданылатын машиналық оқытудың жетілдірілген әдістері. Олар адам миының жұмысын имитациялайтын нейрондық желілерге негізделген. Мұнда терең оқытудың бірнеше негізгі алгоритмдері берілген:

Конволюциялық нейрондық желілер (CNN): негізінен кескінді тану, бейнені талдау және кескінді өңдеу сияқты компьютерлік көру тапсырмалары үшін қолданылады [8].

Ұзақ мерзімді және қысқа мерзімді жад желілері (LSTM): жад күйлерін енгізу арқылы жойылып бара жатқан градиент мәселесін шешу үшін арнайы әзірленген RNN түрі. Бұл оларға ұзақ уақыт бойы ақпаратты тиімді өңдеуге және есте сақтауға мүмкіндік береді [9].

Бұл алгоритмдердің әрқайсысының өзіндік ерекшеліктері бар және тапсырманың ерекшелігіне және өңделетін деректер түріне байланысты қолданылады. Терең оқыту автоматтандырылған жүргізуден бастап дәрі-дәрмектерді әзірлеуге және қаржылық

тенденцияларды болжауға дейінгі әртүрлі салаларға инновациялар әкелу арқылы дамуын жалғастыруда [10].

Қорытынды

Веб-беттердегі зиянды жарнамалық бағдарламаларды анықтау үшін машиналық оқыту алгоритмдерін қолдануға бағытталған зерттеу цифрлық ортаның қауіпсіздігін жақсартудағы деректерге негізделген тәсілдердің айтарлықтай әлеуетін көрсетті. Жұмыс аясында логистикалық регрессия моделі жасалды, оның жоғары дәлдігі мен толықтығы жылу картасы түріндегі тиісті көрсеткіштермен расталған, оның заңды және зиянды мазмұн арасындағы жіктеу мен саралаудағы тиімділігін көрсетеді.

Машиналық оқыту әдістері арқылы зиянды бағдарламалық құралды анықтау алдын алу киберқауіпсіздігінде жаңа көкжиектерді ашады. Атап айтқанда, зиянды қауіптерді анықтауда дәлдік пен толықтықтың теңдестірілген арақатынасын көрсеткен логистикалық регрессияны қолдану неғұрлым күрделі қорғаныс жүйелерін дамытуға негіз бола алады. Мұндай жүйелер жаңа қауіптерге бейімделе алады, олардың дерекқорларын жедел жаңартады және анықталған қауіптерді бейтараптандыру шараларын автоматты түрде қолдана алады.

Сонымен қатар, зерттеу үнемі өзгеріп отыратын киберқауіптер ландшафтында жүйенің өзектілігін сақтау үшін ауқымды оқыту үлгілерін жинау және үлгілерді үнемі жаңарту қажеттілігі туралы маңызды сұрақтарды көтереді. Болашақ жұмыс модельдердің күрделі және жасырын зиянды бағдарламаларды тану қабілетін жақсарту үшін терең оқыту мен нейрондық архитектураны біріктіруге бағытталуы мүмкін.

Осы зерттеудің нәтижелері ғылыми қоғамдастыққа киберқауіпсіздік контекстінде машиналық оқытудың теориялық негіздерін одан әрі тереңдетуге және цифрлық ортаны барлық пайдаланушылар үшін қауіпсіз ете отырып, нақты уақыт режимінде киберқауіптерге төтеп бере алатын практикалық шешімдерді әзірлеуге байланысты жаңа міндеттер қояды.

References

1. Oshingbesan A. Detection of Malicious Websites Using Machine Learning Techniques. Cryptography and Security (cs.CR) / A. Oshingbesan, et al // Machine Learning (cs.LG)/ – 2022. Vol. 1. http://dx.doi.org/10.13140/RG.2.2.30165.14565.
2. Akhtar M.S. Malware Analysis and Detection Using Machine Learning Algorithms. / M. S. Akhtar & T. Feng // Symmetry. – 2022. – № 14(11). – P. 2304. http://dx.doi.org/10.3390/sym14112304.
3. Saxe J. Deep Neural Network Based Malware Detection Using Two Dimensional Binary Program Features / J. Saxe & K. Berlin // 10th International Conference on Malicious and Unwanted Software (MALWARE). – 2015. http://dx.doi.org/10.1109/MALWARE.2015.7413680.
4. Shijo P.V. Integrated Static and Dynamic Analysis for Malware Detection / P.V. Shijo & A. Salim // Procedia Computer Science. – 2015. – № 46. P. 804-811. http://dx.doi.org/10.1016/j.procs. 2015.02.149.
5. Alazab M. Malware Detection Systems: The Need for Machine Learning Algorithms / M. Alazab M. et al // Springer Nature. – 2020. http://dx.doi.org/10.1007/978-981-16-7618-5_53.
6. Shin S.-S. A Heterogeneous Machine Learning Ensemble Framework for Malicious Webpage Detection / S.-S. Shin, S.-G. Ji & S.-S. Hong // Applied Sciences. – 2022. – № 12(23). P. 12070. http://dx.doi.org/10.3390/app10030936.
7. Anis F.M. Interpretable Machine Learning Models for Malicious Domains Detection Using Explainable Artificial Intelligence (XAI) / F.M. Anis, R.M. Aljuaid & R. Baageel // Sustainability. – 2022. – № 14(12). P. 7375. http://dx.doi.org/10.3390/su14127375.
8. Khan T.A. Significance of Machine Learning for Detection of Malicious Websites on an Unbalanced Dataset / T.A Khan & R. Kouatly // Digital. – 2020. № 2(4). P. 501-519. http://dx.doi.org/10.3390/digital2040027.
9. Akhtar M.S. Malware Analysis and Detection Using Machine Learning Algorithms / M.S. Akhtar & T. Feng // Symmetry. – 2022. № 14(11). P. 2304. http://dx.doi.org/10.3390/sym14112304.
10. Detection of Malicious Websites Using Machine Learning Techniques / A. Oshingbesan et al. – 2022. http://dx.doi.org/10.13140/RG.2.2.30165.14565.

Н.Е.Рахимбай*, К.Б. Тусупова
Казахский национальный университет им. аль-Фараби,
050040, Республика Казахстан, г. Алматы, пр. аль-Фараби, 71
*e-mail: nazerke.rkh@gmail.com

ИСПОЛЬЗОВАНИЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ ВРЕДОНОСНЫХ РЕКЛАМНЫХ ПРОГРАММ НА ВЕБ-СТРАНИЦАХ

В статье рассматривается проблема распространения вредоносных рекламных программ через веб-страницы, которые представляют серьезную угрозу конфиденциальности и безопасности пользователей интернета. Использование алгоритмов машинного обучения для обнаружения и нейтрализации вредоносных рекламных программ, встроенных в Веб-страницы. Сосредоточив внимание на методах обработки данных, извлечения меток и классификации, машинное обучение подробно анализирует, как оно может улучшить процессы обнаружения вредоносных программ. Различные алгоритмы машинного обучения, включая логистическую регрессию, деревья решений, случайный лес, наивные байесовские и ансамблевые методы, изучаются для определения их эффективности в различении вредоносного и законного рекламного контента.

Описана методика построения обучающих и тестовых моделей, включающая данные о вредоносных и безопасных рекламных модулях. Различные подходы к машинному обучению, включая обучение с учителем, обучение без учителя и методы глубокого обучения, анализируются для выявления скрытых моделей вредного поведения. Результаты исследования показывают, что использование алгоритмов машинного обучения позволяет с высокой точностью обнаруживать вредоносные рекламные программы, что может стать основой для разработки более эффективных инструментов кибербезопасности. Также обсуждаются потенциальные проблемы и ограничения существующих методов, а также предлагаются направления для дальнейших исследований по выявлению вредоносных рекламных программ с помощью машинного обучения.

Ключевые слова: *машинное обучение, вредоносное ПО, кибербезопасность, искусственный интеллект, веб-страницы.*

N.E. Rakhimbay*, K.B. Tusupova
Al-Farabi Kazakh National University,
050040, Republic of Kazakhstan, Almaty, al-Farabi Ave., 71
*e-mail: nazerke.rkh@gmail.com

USING MACHINE LEARNING ALGORITHMS TO DETECT MALICIOUS ADVERTISEMENTS ON WEB PAGES

The article examines the problem of the spread of malicious advertising programs through web pages that pose a serious threat to the privacy and security of Internet users. Using machine learning algorithms to detect and neutralize malicious advertising programs embedded in Web pages. By focusing on data processing, tag extraction, and classification techniques, machine learning analyzes in detail how it can improve malware detection processes. Various machine learning algorithms, including logistic regression, decision trees, random forest, naive Bayesian and ensemble methods, are being studied to determine their effectiveness in distinguishing malicious and legitimate advertising content.

A methodology for building training and test models, including data on malicious and secure advertising modules, is described. Various approaches to machine learning, including teacher-led learning, unsupervised learning, and deep learning techniques, are being analyzed to identify hidden patterns of harmful behavior. The results of the study show that the use of machine learning algorithms makes it possible to detect malicious advertising programs with high accuracy, which can become the basis for the development of more effective cybersecurity tools. Potential problems and limitations of existing methods are also discussed, as well as directions for further research on detecting malicious advertising programs using machine learning.

Key words: *machine learning, malware, cybersecurity, artificial intelligence, web pages.*

Авторлар туралы мәліметтер

Назерке Рахимбай – магистрант, «Ақпараттық жүйелер» кафедрасы, әл-Фараби атындағы Қазақ ұлттық университеті, e-mail: nazerke.rkh@gmail.com

Камшат Тусупова – PhD, «Ақпараттық жүйелер» кафедрасының аға оқытушысы, әл-Фараби атындағы Қазақ ұлттық университеті, e-mail: kamshat-0707@mail.ru

Сведения об авторах

Назерке Рахимбай – магистрант кафедры «Информационные системы», Казахский национальный университет им. аль-Фараби, e-mail: nazerke.rkh@gmail.com

Камшат Тусупова – PhD, старший преподаватель кафедры «Информационные системы», Казахский национальный университет им. аль-Фараби, e-mail: kamshat-0707@mail.ru

Information about the authors

Nazerke Rakhimbay – Master's student of the Department of Information Systems, Al-Farabi Kazakh National University, e-mail: nazerke.rkh@gmail.com.

Kamshat Usupova – PhD, Senior Lecturer at the Department of Information Systems, Al-Farabi Kazakh National University, e-mail: kamshat-0707@mail.ru.

Редакцияға енуі 11.04.2024
Жариялауға қабылданды 13.05.2024

DOI: 10.53360/2788-7995-2024-2(14)-7

ҒТАХР: 81.93.29

Ж.К. Абдугулова, М.Н.Тлеген*, Т.М. Төлеубеков
Л.Н. Гумилев атындағы Еуразия ұлттық университеті
10000, Қазақстан Республикасы, Астана қ., Сатпаев к-сі, 2
*e-mail: meruert-0202@mail.ru

МУЛЬТИРОТОР ТИПТІ ҰШҚЫШСЫЗ ҰШУ АППАРАТЫНЫҢ БАСҚАРУ ЖҮЙЕСІН ЗЕРТТЕУ ЖӘНЕ ҚҰРАСТЫРУ

Аңдатпа: *Мультиротор типті ұшқышсыз ұшу аппаратының басқару жүйесін зерттеу және құрастыру қарастырылды. ҰҰА толықтай ақпарат беріліп, қазіргі таңдағы маңыздылығы анықталды. Басқару объектісі ретінде квадрокоптер алынды. ҰҰА зерттеп, олардың түрін ажыратып талдау және басқару объектісі квадрокоптердің құрлымын, жұмыс істеу принципі зертелді және оның қозғалысына байланысты математикалық моделі жасалды. ҰҰА-ның дизайны және формасына, жұмыс істеу принципіне байланысты мультироторлы, бекітілген қанатты, бір роторлы-ұшқышсыз тікұшақ, гибридті түрлері ажыратылды. ҰҰА-ның мультироторлы типі, соның ішінде квадрокоптердің құрлымы мен бөлшектері қарастырылды, олардың түрлеріне және құрлымына байланысты түсініктеме берілді. Жұмыс істеу принципіне талдау жасалды. Квадрокоптерді басқару негізінен жердегі оператор және борттық жүйе болып екі бөлікке бөлінеді бөлінеді. Жердегі оператор өзіне керекті биіктікті, бағытты таратқыш, яғни, басқару пульті арқылы ақпаратты квадрокопптерге береді. Таратқыштар көбінде 4 арналы болып келеді және ол 2.4 GHz жиілігінде жұмыс істейді. Кейбір таратқыштарда бұл арналардың саны 7 немесе 10-ға жетуі мүмкін. Борттық жүйедегі қабылдағыш ақпаратты қабылдап оны ҰК-ға берді. ҰК-і операторға керекті бағытқа бұрылу үшін әр қозғалтқышқа берілетін күштің, кернеудің, жылдамдықтың есептеулерін жасайды. Ал қозғалтқыштарға керекті жылдамдықты реттеу үшін электронды жылдамдықты реттегіш қолданылады. Әр қозғалтқыштың жылдамдығын өзгерту арқылы оның бағытын өзгертеміз. Квадрокоптердің ұшу бағытының өзгеріс түрлеріне бұрандалардың әсері, яғни, әр бұрандаға әсер ететін тарту күшінің әсерінен аппараттың бағытының өзгерісі анықталды. Аппараттың қозғалысының крен, рыскание, тангаж, төмен-жоғары, алдыға-артқа, оңға – солға 6 бағыты болады.*

Түйін сөздер: *Ұшқышсыз ұшу аппараттары, Мультикоптер, Басқару объекті, Дрон.*

Кіріспе. Қазіргі уақытта автоматты техниканы дамытудың арқасында ұшқышсыз авиациялық жүйелерді, сондай-ақ олардың негізіндегі кешенді шешімдерді қолдану айтарлықтай кеңеюде. Ұшқышсыз ұшу аппараттары (ҰҰА) – бұл заманауи, қазіргі таңдағы өте жоғарғы қарқынмен дамып келе жатқан ірі көлемдегі ұшу аппараттарынан (ҰА) пайдалану мүмкін емес жағдайларда немесе адам өміріне қатер төндіретін жағдайларда қолдануға арналған авиациялық техниканың бір түрі.