

Розломій Інна, Хотунов Владислав, Науменко Сергій. Розділ XVII. Штучний інтелект у виявленні загроз інформаційній безпеці в хмарних освітніх середовищах. Штучний інтелект в освіті. Частина 1 : монографія / [авт. колектив]; за ред. Яцишин А. – Київ: ІЦО НАПН України, 2025. – С. 257-270. ISBN 978-617-8330-53-8

РОЗДІЛ XVII. ШТУЧНИЙ ІНТЕЛЕКТ У ВИЯВЛЕННІ ЗАГРОЗ ІНФОРМАЦІЙНІЙ БЕЗПЕЦІ В ХМАРНИХ ОСВІТНІХ СЕРЕДОВИЩАХ

DOI: 10.33407/lib.NAES.id/748129

Розломій Інна¹[0000-0001-5065-9004], Хотунов Владислав²[0000-0002-2093-1270],

Науменко Сергій³[0000-0002-6337-1605],

¹ Черкаський державний технологічний університет, Черкаси, Україна

² Черкаський державний фаховий бізнес-коледж, Черкаси, Україна

³ Черкаський національний університет імені Богдана Хмельницького,

Черкаси, Україна

inna-roz@ukr.net

Анотація. У статті розглянуто підхід до виявлення загроз інформаційній безпеці в хмарних освітніх середовищах на основі методів штучного інтелекту. Описано архітектуру типової хмарної платформи, класифіковано основні загрози, включаючи автентифікаційні атаки, порушення цілісності, витоки конфіденційної інформації та внутрішні ризики. Запропоновано інтелектуальну модель виявлення аномальної активності користувачів із використанням алгоритмів машинного навчання – дерев рішень, багатошарового перцептрона та градієнтного бустингу. Проведено підготовку датасету з логів активності, реалізовано попередню обробку, балансування класів та нормалізацію. Оцінювання точності, повноти, специфічності й F1-міри засвідчило перевагу запропонованого підходу порівняно з традиційними сигнатурними IDS. Наведено рекомендації щодо інтеграції моделі в існуючі освітні платформи, зокрема Moodle та Google Classroom, а також технічні й організаційні умови для її впровадження. Запропоноване рішення є масштабованим, адаптивним і придатним для практичного застосування в умовах обмежених ресурсів закладів освіти.

Ключові слова: штучний інтелект, інформаційна безпека, хмарні освітні середовища, машинне навчання, виявлення загроз, поведінковий аналіз.

ARTIFICIAL INTELLIGENCE IN DETECTION OF INFORMATION SECURITY THREATS IN CLOUD EDUCATIONAL ENVIRONMENTS

Rozlomii Inna ¹[0000-0001-5065-9004], Khotunov Vladyslav ²[0000-0002-2093-1270],

Naumenko Serhii ³[0000-0002-6337-1605],

¹ Cherkasy State Technological University, Cherkasy, Ukraine

² Cherkasy State Business College, Cherkasy, Ukraine

³ Bohdan Khmelnytsky National University of Cherkasy, Cherkasy, Ukraine

АНОТАЦІЯ. The article presents an approach to detecting information security threats in cloud-based educational environments using artificial intelligence methods. The architecture of a typical cloud platform is described, and key threat categories are classified, including authentication attacks, data integrity violations, confidentiality breaches, and internal risks. An intelligent model for detecting anomalous user behavior was proposed, applying machine learning algorithms such as decision trees, multilayer perceptrons, and gradient boosting. A dataset of user activity logs was prepared, followed by preprocessing, class balancing, and feature normalization. Evaluation of accuracy, recall, specificity, and F1-score demonstrated the superiority of the proposed approach compared to traditional signature-based IDS. The paper provides recommendations for integrating the model into existing learning management systems such as Moodle and Google Classroom, along with outlining technical and organizational conditions for deployment. The proposed solution is scalable, adaptive, and suitable for practical implementation under the resource constraints of educational institutions.

Ключові слова: artificial intelligence, information security, cloud-based education, machine learning, threat detection, behavioral analysis.

Вступ. Цифровізація освітнього процесу стрімко трансформує традиційні підходи до навчання, комунікації та управління освітніми установами. Одним із ключових рушіїв цієї трансформації є хмарні технології, які забезпечують централізоване зберігання навчальних матеріалів, інтегровану взаємодію між

учасниками освітнього процесу та доступ до даних у режимі реального часу з будь-якої точки світу. Платформи на зразок Google Workspace for Education, Microsoft 365 Education, MoodleCloud та інші демонструють зростаючу популярність як серед закладів вищої освіти, так і серед шкіл та центрів підвищення кваліфікації.

За останні роки кількість користувачів хмарних освітніх середовищ зросла в геометричній прогресії. Цьому сприяють як вимоги до гнучкості навчання (зокрема в умовах пандемії та воєнного стану), так і прагнення до економії ресурсів та підвищення ефективності освітнього процесу. У хмарних середовищах здійснюється зберігання персональних даних учасників освітнього процесу, управління результатами навчання, обробка адміністративної документації, а також надається доступ до інструментів тестування, відеоконференцій, віртуальних лабораторій та колективної роботи.

Разом із тим, така відкритість і масштабованість хмарних сервісів створює значні виклики в контексті інформаційної безпеки. Освітні установи дедалі частіше стають об'єктами кіберзлочинців, які прагнуть отримати несанкціонований доступ до масивів персональних або службових даних. Типовими загрозами виступають фішинг-атаки, зловмисне програмне забезпечення, несанкціоноване проникнення, внутрішні витоки інформації та вразливості в конфігурації хмарних сервісів. Складність ситуації поглиблюється тим, що в освітньому середовищі користувачами часто є малодосвідчені з безпекового погляду особи — студенти, викладачі, адміністративний персонал, які не завжди дотримуються політик безпеки або не мають навичок реагування на підозрілу активність.

У відповідь на ці виклики все частіше залучаються інструменти штучного інтелекту, які здатні виявляти аномальну поведінку в системі, аналізувати великі обсяги логів, виявляти нові типи атак та адаптувати захисні механізми відповідно до поточних умов. Методи машинного навчання, зокрема кластеризація, класифікація, нейронні мережі та байєсівські моделі, застосовуються для створення інтелектуальних систем виявлення загроз (IDS) і прогнозування кіберінцидентів у

реальному часі. В освітньому контексті особливої уваги набувають також питання ресурсоефективності таких рішень, адаптивності до змінних навчальних середовищ та сумісності з уже впровадженими платформами.

Таким чином, використання штучного інтелекту в системах безпеки хмарних освітніх середовищ відкриває нові можливості для проактивного виявлення загроз, однак вимагає глибокого аналізу технічних, організаційних і педагогічних аспектів впровадження таких рішень. Саме ці питання лежать в основі подальшого дослідження.

Аналіз літератури та постановка проблеми. Тема застосування методів штучного інтелекту для забезпечення інформаційної безпеки в освітніх середовищах перебуває у фокусі наукових досліджень останніх років. З огляду на динамічне впровадження хмарних технологій в освіту, значна частина праць присвячена аналізу вразливостей таких платформ і розробці інтелектуальних систем виявлення загроз.

У роботі A. Alsghaier та співавт. [1] розглянуто моделі загроз, що виникають у хмарних освітніх середовищах, зокрема загрози аутентифікації, витоку даних і атак з боку внутрішніх користувачів. Автори наголошують на необхідності динамічного моніторингу активності користувачів і вказують на перспективність використання алгоритмів машинного навчання для розпізнавання аномальної поведінки.

Дослідження S. Shakeel, M. Baskar та ін. [2] демонструє, як методи класифікації на основі нейронних мереж можуть ефективно виявляти підозрілу активність у хмарних системах. Зокрема, в моделі використовуються багатошарові перцептрони для аналізу логів доступу та запитів до серверів, що дозволяє виявляти нетипові патерни.

У праці R. Vinayakumar et al. [3] запропоновано глибоку згорткову нейронну мережу (CNN) для виявлення кіберзагроз у хмарних середовищах, яка демонструє високу точність при виявленні атак типу DoS, R2L та Probe. Ця модель може бути адаптована для потреб освітніх платформ із великим потоком даних.

З іншого боку, в роботі S. Sahu et al. [4] акцент зроблено на ефективності виявлення загроз у середовищах з обмеженими обчислювальними ресурсами. Автори пропонують гібридну систему, що поєднує легковагові алгоритми класифікації та попередню обробку даних із метою зменшення навантаження на систему.

У сфері освітніх технологій дослідження A. Ahmed et al. [5] розглядає можливості застосування ШІ для моніторингу академічної доброчесності та забезпечення конфіденційності студентських даних. Автори підкреслюють потребу в створенні систем, які не тільки захищають від зовнішніх атак, а й враховують етичні аспекти роботи з освітніми даними.

Попри наявність широкого спектра досліджень, ряд проблем залишається невирішеним. Зокрема, більшість існуючих рішень розроблено для комерційних хмарних систем і не враховують специфіку освітнього середовища, зокрема його відкритість, непостійність складу користувачів і відсутність у них технічних навичок. Крім того, ускладнює ситуацію обмеженість обчислювальних ресурсів у навчальних закладах, що не дозволяє використовувати складні моделі глибокого навчання без попередньої оптимізації. Актуальним є й питання пояснюваності рішень моделей ШІ, що особливо важливо в контексті освітньої прозорості та дотримання політик конфіденційності.

Проблема дослідження полягає у відсутності універсальних механізмів раннього виявлення загроз інформаційній безпеці в хмарних освітніх середовищах, здатних ефективно функціонувати в умовах високої варіативності атак та обмежених обчислювальних ресурсів.

Метою дослідження є обґрунтування доцільності та ефективності використання методів штучного інтелекту для виявлення загроз інформаційній безпеці в хмарних освітніх середовищах, а також формування підходів до проектування адаптивних систем виявлення атак.

Результати дослідження. Сучасна хмарна освітня платформа є багаторівневою розподіленою інформаційною системою, що поєднує зберігання,

обробку та передавання освітніх даних у режимі реального часу. У такому середовищі передбачена взаємодія великої кількості користувачів – студентів, викладачів, адміністративного персоналу, а також зовнішніх сервісів, що створює широке поле для потенційних загроз. На рисунку 1 наведено узагальнену архітектуру типової хмарної освітньої платформи

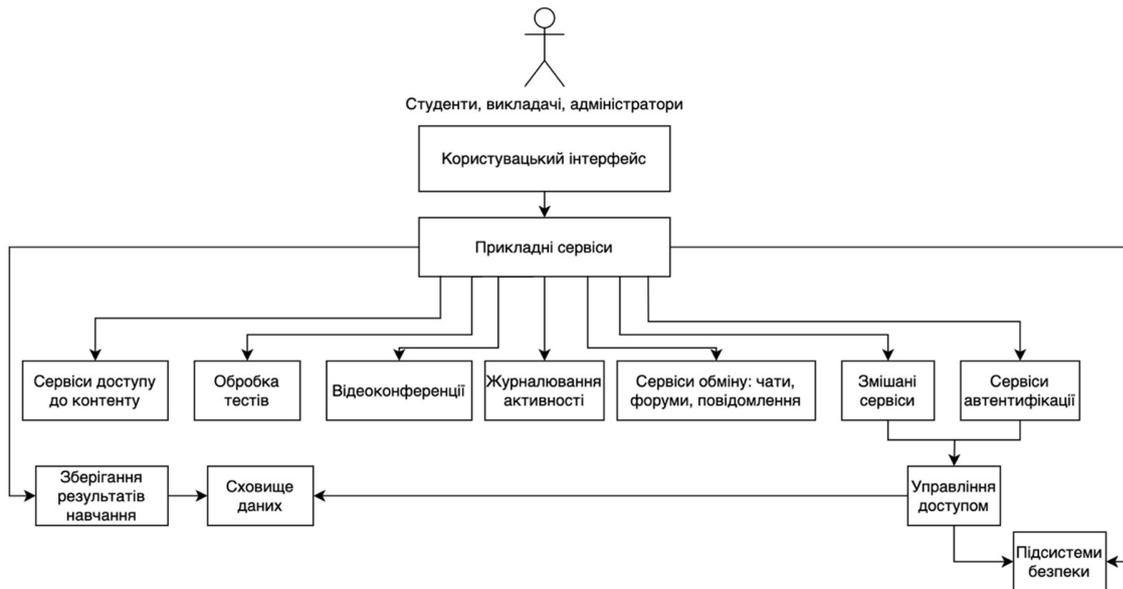


Рис. 1. Архітектура хмарної освітньої платформи

На рисунку представлено архітектуру з поділом на рівні: користувацький інтерфейс, прикладні сервіси, рівень управління доступом, сховище даних, та підсистеми безпеки. Основними елементами є модулі автентифікації, сервіси доступу до контенту, обробки тестів, зберігання результатів навчання, відеоконференцв'язку та журналювання активності.

До основних функціональних компонентів платформи належать:

1. Сервіси зберігання даних – забезпечують централізоване зберігання навчальних матеріалів, електронних журналів, результатів тестування, персональних даних користувачів. Реізуються зазвичай у вигляді хмарних СКБД або об'єктних сховищ.

2. Сервіси автентифікації та авторизації – здійснюють ідентифікацію користувачів, перевірку прав доступу, керування сесіями. Включають механізми двофакторної автентифікації, OAuth, SSO.

3. Сервіси обміну даними – модулі, що забезпечують взаємодію між користувачами та системою, включаючи чати, форуми, відеоконференції (Zoom, Meet), сервіси перевірки знань, надсилання повідомлень та зворотного зв'язку.

Типові сценарії використання хмарної освітньої платформи разом із потенційними точками вторгнення подано в таблиці 1.

Хмарні освітні середовища характеризуються великою кількістю точок потенційного доступу до системи, серед яких – як очевидні, так і менш помітні вектори атаки. Умови постійної взаємодії великої кількості користувачів з різним рівнем технічної підготовки, а також відкритість до зовнішніх сервісів створюють складне безпекове середовище, яке потребує постійного аналізу та контролю. Це вимагає впровадження інструментів, здатних виявляти нетипову активність, адаптуватися до нових форм загроз і враховувати особливості архітектури освітніх платформ.

Таблиця 1

Типові сценарії використання платформи та можливі вектори атак

Сценарій використання	Можлива точка вторгнення	Тип загрози
Студент входить до системи через публічну мережу Wi-Fi	Перехоплення сесії, атака "людина посередині"	Витік облікових даних
Викладач завантажує навчальні матеріали до хмари	Інжекція шкідливого коду через файли	Порушення цілісності системи
Перевірка тестових завдань у вебінтерфейсі	Злом облікового запису викладача	Несанкціонований доступ
Відеозустріч з керівником програми у Zoom або Meet	Підключення зловмисника до відкритого посилання	Порушення конфіденційності
Адміністратор змінює параметри доступу до сховища	Вразливості у панелі адміністрування	Ескалація привілеїв
Масове надсилання повідомлень студентам	Злом системи обміну повідомленнями (SMTP/Push)	Соціальна інженерія, фішинг

У хмарних освітніх середовищах інформаційна безпека формується на перетині технічних, організаційних та поведінкових чинників, а самі загрози мають різноманітний характер і можуть виникати як ззовні, так і зсередини системи. Динаміка взаємодії між користувачами, сервісами та інфраструктурою створює складну та багаторівневу модель ризиків, у якій недостатньо застосовувати традиційні підходи до виявлення загроз. Ефективне моделювання потенційно небезпечних ситуацій потребує попередньої класифікації загроз за типами їх впливу та джерелами походження. На рисунку 2 представлено узагальнену структуру загроз, характерних для хмарних освітніх середовищ.

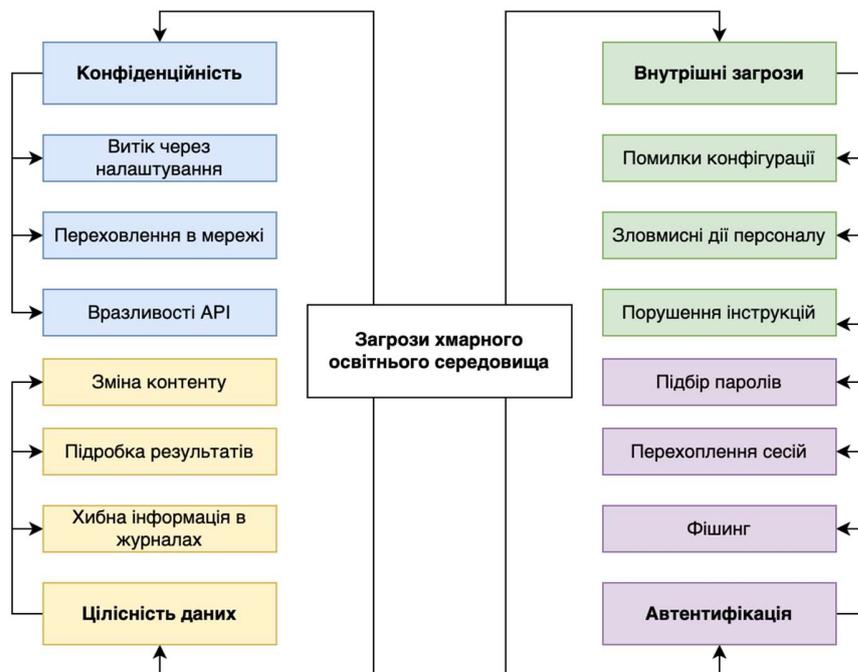


Рис. 2. Класифікація загроз інформаційній безпеці хмарного освітнього середовища

Схема представляє чотири основні категорії загроз: пов'язані з автентифікацією, цілісністю, конфіденційністю та внутрішніми ризиками. Для кожної категорії наведено приклади типових векторів атак або порушень, а також вказано, на які підсистеми освітнього середовища вони можуть впливати: користувачів, сховища даних, канали зв'язку або інтерфейси адміністрування.

Такий підхід до класифікації дозволяє структурувати ризики й сформувати основу для побудови інтелектуальних систем, здатних виявляти загрози з урахуванням їх типових ознак і поведінкових моделей.

У межах дослідження було запропоновано використання методів машинного навчання для виявлення загроз інформаційній безпеці в хмарних освітніх середовищах. Ураховуючи обмеженість обчислювальних ресурсів, варіативність поведінки користувачів та потребу в пояснюваності рішень, пріоритет було надано класичним алгоритмам класифікації – таким як дерева рішень (Decision Tree) та стохастичний градієнтний бустинг. У процесі експериментів також було протестовано багатошаровий перцептрон (MLP) як представника нейромережевого підходу, орієнтованого на виявлення складніших залежностей у даних.

Для навчання моделей сформовано набір ознак, що відображають типові дії користувачів у хмарному освітньому середовищі. До таких ознак належать: час входу в систему, географічне положення користувача, кількість спроб авторизації, частота звернень до ресурсів, зміни в навчальному контенті, характер взаємодії з файлами, інтенсивність активності в системі тощо. Значну увагу приділено очищенню, нормалізації та балансуванню даних, що дозволило зменшити вплив шуму та зміщень у вибірці.

Архітектура моделі залежала від обраного підходу. У випадку дерева рішень було реалізовано ієрархічну структуру з розгалуженням за найбільш значущими ознаками, визначеними за допомогою критерію Джині. Для MLP використано три шари з функціями активації ReLU, що дозволили досягти прийняттого співвідношення між точністю і швидкістю навчання. Модель було навчено у середовищі Python з використанням бібліотек Scikit-learn, Keras та TensorFlow, що забезпечили гнучкість налаштувань, візуалізацію процесів та автоматизацію підбору гіперпараметрів.

Розроблений підхід орієнтований на виявлення нетипової поведінки користувачів, що потенційно вказує на загрозу. Він дозволяє виявляти як відомі типи атак, так і аномалії, що не були явно визначені на етапі навчання, завдяки здатності моделі до генералізації та виявлення прихованих закономірностей у поведінкових патернах.

Побудова ефективної моделі виявлення загроз потребує ретельної роботи з даними на всіх етапах – від їх збору до оцінки якості навчання. У межах дослідження було реалізовано послідовний підхід до підготовки датасету та налаштування процесу навчання, який включав такі кроки:

1. Джерела даних. Для навчання моделей використано знеособлені лог-файли активності користувачів, отримані з хмарних освітніх платформ, які містили інформацію про типи взаємодій, час подій, параметри авторизації та використання ресурсів. У разі обмеженого доступу до реальних даних застосовано відкриті набори, зокрема *UNSW-NB15* та *CICIDS2017*, що були адаптовані до особливостей освітніх середовищ шляхом фільтрації відповідних типів активності.

2. Предобробка даних. Після первинного очищення від дублікатів і неповних записів виконано нормалізацію числових ознак для зменшення впливу масштабних відмінностей між параметрами. Було враховано проблему незбалансованості класів – більшість прикладів відповідала нормальній поведінці користувачів. Для забезпечення коректного навчання моделей застосовано методи балансування, зокрема SMOTE, а також кодування категоріальних ознак через one-hot перетворення.

3. Умови навчання. Дані поділено на тренувальну та тестову вибірки у співвідношенні 80:20. Для підвищення надійності оцінки моделей використано п'ятиразову крос-валідацію. Оцінювання ефективності здійснювалося за метриками точності, повноти, точності позитивного класу (precision), F1-міри та AUC, що дозволило всебічно проаналізувати здатність моделей до класифікації нормальної й аномальної активності.

Підхід до підготовки даних був спрямований на досягнення максимальної достовірності результатів моделювання та забезпечення адаптивності алгоритмів до різноманітних сценаріїв використання хмарних освітніх платформ.

Для об'єктивного аналізу результатів було використано поширені метрики оцінювання якості класифікації, які дозволяють визначити здатність моделі правильно виявляти загрози та мінімізувати кількість помилкових спрацювань. У процесі тестування класифікатора було отримано такі значення компонентів матриці помилок:

- TP (істинно позитивні, коректно виявлені загрози): 865;
- TN (істинно негативні, коректно розпізнана нормальна активність): 3710;
- FP (хибно позитивні, помилкові тривоги): 215;
- FN (хибно негативні, пропущені загрози): 90.

На основі цих даних обчислимо ключові показники:

$$1. \text{ Точність (Accuracy) } - \text{ Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{865+3710}{865+3710+215+90} \approx 0.9375.$$

$$2. \quad \text{Повнота (Recall) – Recall} = \frac{TP}{TP+FN} = \frac{865}{865+90} \approx 0.9058.$$

$$3. \quad \text{Специфічність (Specificity) – Specificity} = \frac{TN}{TN+FN} = \frac{3710}{3710+215} \approx 0.9442.$$

$$4. \quad \text{F1-міра – Precision} = \frac{TP}{TP+FN} = \frac{865}{865+215} \approx 0.8009,$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \cdot \frac{0.8009 \cdot 0.9058}{0.8009 + 0.9058} \approx 0.8511.$$

Отримані результати демонструють високу загальну точність класифікації (93,75%) та чутливість до загроз (90,58%), що свідчить про здатність моделі виявляти більшість потенційно небезпечних дій користувачів у хмарному освітньому середовищі. F1-міра в межах 0,85 вказує на збалансованість між повнотою та точністю навіть за наявності неідеального співвідношення класів.

Для порівняння було проведено тестування базової системи виявлення загроз, яка реалізує сигнатурний підхід на основі фіксованого набору відомих патернів атак. Така система виявила 720 істинно позитивних випадків і продемонструвала такі характеристики:

– Accuracy ≈ 0.898 ;

– Recall ≈ 0.754 ;

– Specificity ≈ 0.931 ;

– F1 ≈ 0.768 .

Як видно з порівняння, сигнатурний підхід поступається інтелектуальній моделі за всіма основними показниками, особливо в частині повноти та F1-міри. Це підтверджує обмеження традиційних рішень, які не здатні адаптуватися до нових або модифікованих типів атак і не враховують контекст поведінки користувачів.

Таким чином, результати оцінювання свідчать про перевагу інтелектуального підходу, який не лише забезпечує вищу точність виявлення загроз, а й демонструє кращу адаптивність до змінного безпекового середовища хмарних освітніх платформ.

Висновки. Запропонований у дослідженні інтелектуальний підхід до виявлення загроз інформаційній безпеці в хмарних освітніх середовищах продемонстрував високу ефективність за ключовими метриками класифікації та суттєво перевищив точність традиційних сигнатурних систем. Розроблена модель виявлення загроз базується на аналізі поведінкових ознак користувачів, що дозволяє фіксувати як відомі, так і нові типи аномальної активності, пов'язані з порушенням автентифікації, цілісності або конфіденційності даних.

Для успішного впровадження такого підходу в освітніх установах доцільно забезпечити низку технічних та організаційних умов. На технічному рівні важливою є наявність логів взаємодії користувачів із системою, їх централізований збір, знеособлення та зберігання у форматі, придатному для подальшого аналізу. Окрему увагу слід приділити конфігурації хмарних середовищ з точки зору підтримки API-доступу для отримання подій у режимі реального часу. З організаційного боку варто передбачити регулярний аудит безпеки, підготовку персоналу до роботи з аналітичними інструментами, а також встановлення чітких політик щодо обробки інцидентів.

Розгортання моделі в межах уже функціонуючих освітніх платформ можливе за умови інтеграції з існуючими системами управління навчанням (LMS), такими як Moodle або Google Classroom. У першому випадку інтеграція може реалізовуватись через модулі спостереження за активністю користувачів, а в другому – за допомогою інструментів Google Workspace Admin API або сторонніх сервісів безпеки з підтримкою аналітики. Універсальність архітектури розробленої моделі дозволяє адаптувати її до конкретного програмного середовища без суттєвих змін структури.

Враховуючи постійне зростання ролі хмарних технологій в освіті, запропонований підхід може стати основою для створення адаптивних систем виявлення загроз, що працюють у режимі реального часу та підвищують загальний рівень інформаційної безпеки цифрових навчальних середовищ.

Список джерел

1. Розломій, І. О., & Ярмілко, А. В. (2022). Інтегрований криптографічно-стеганографічний захист мультимедійного контенту і даних в системах e-learning. Вісник Черкаського національного університету. Серія «Прикладна математика. Інформатика», (1), 50-60.
2. Rozlomii, I., Yehorchenkova, N., Yarmilko, A., & Naumenko, S. (2023). Data Protection in the Utilization of Natural Language Processors for Trend Analysis and Public Opinion: rypographic Aspect. In Proceedings of the 2nd International Workshop on Social Communication and Information Activity in Digital Humanities (SCIA-2023) (pp. 1-11).
3. Alsghaier, A., Alenezi, A., & Mahmoud, Q. H. (2021). Security threats in cloud-based learning environments: a comprehensive review. *Computers & Security*, 105, 102241. <https://doi.org/10.1016/j.cose.2021.102241>
4. Shakeel, S., Baskar, M., & Dhulipala, V. S. (2020). Neural network based anomaly detection in cloud environments for security enhancement. *Journal of Intelligent & Fuzzy Systems*, 38(2), 1935–1943. <https://doi.org/10.3233/JIFS-179271>
5. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2019). Applying convolutional neural network for network intrusion detection. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 1222–1228. <https://doi.org/10.1109/ICACCI.2017.8126056>
6. Sahu, S., Dash, R., & Behera, H. S. (2021). A hybrid lightweight intrusion detection system for cloud environments. *Future Generation Computer Systems*, 115, 346–359. <https://doi.org/10.1016/j.future.2020.09.020>
7. Ahmed, A., Alenezi, M., & Zain, J. M. (2022). Artificial intelligence applications in education: security, privacy and ethics. *Education and Information Technologies*, 27(3), 3255–3273. <https://doi.org/10.1007/s10639-021-10704-1>
8. Zabolotnii, S., Yarmilko, A., Rozlomii, I., & Mysiura, Y. (2023). Applying the Arithmetic Compression Method in Digital Speech Data Processing. In Selected Papers of the X International Scientific Conference «Information Technology and Implementation» (IT&I-2023). (pp. 170-179).