КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (34), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.34.1237
УДК 004.89

Світличний Віталій Анатолійович
кандидат технічних наук, доцент, доцент кафедри протидії кіберзлочинності
Харківський національний університет внутрішніх справ, Харків, Україна
ORCID:0000-0003-3381-3350
vit.svet@ukr.net

Клімушин Петро Сергійович
кандидат технічних наук, доцент, доцент кафедри протидії кіберзлочинності
Харківський національний університет внутрішніх справ,Харків, Україна
ORCID:0000-0002-1020-9399
klimushyn@ukr.net

Онищенко Юрій Миколайович
кандидат наук з державного управління, доцент, заступник директора інституту з освітньої та науково-дослідної діяльності навчально-наукового інституту № 4
Харківський національний університет внутрішніх справ, Харків, Україна
ORCID:0000-0002-7755-3071
onischenko1980@gmail.com

ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ СПЕЦІАЛІЗОВАНИХ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ В СИСТЕМІ КІБЕРЗАХИСТУ УКРАЇНИ: ТЕХНІКО-ПРАВОВИЙ АСПЕКТ

Анотація. У статті досліджуються можливості використання спеціалізованих моделей штучного інтелекту (AI) в галузі кібербезпеки, зокрема, на прикладі новітньої моделі GPT-5.4-Cyber. Сучасні кіберзагрози характеризуються високим рівнем складності, динамічності та транскордонності, що обумовлює гостру необхідність застосування інтелектуальних систем автоматизованого аналізу та оперативного реагування. Особливої актуальності це питання набуває в умовах воєнного стану та ведення гібридної війни, де цифрова інфраструктура України стає об'єктом постійних атак. У дослідженні детально аналізуються функціональні особливості моделі, її роль у забезпеченні національного кіберзахисту, а також потенційні ризики та перспективи її впровадження в повсякденну діяльність правоохоронних органів України. Окрему увагу приділено діяльності підрозділів Національної поліції України, які виконують ключові функції з протидії кіберзлочинності, захисту державних інформаційних ресурсів та підтримки правопорядку в цифровому середовищі. Запропоновано авторську математичну модель для комплексної оцінки ефективності інтелектуальних систем безпеки (E_{AI}), що базується на розрахунку інтегрального показника. Даний показник поєднує в собі такі критичні параметри, як точність виявлення аномалій (Detection), швидкість автоматизованого реагування (Response), рівень автономності системи (Automation) та стійкість до зовнішніх маніпуляцій чи отруєння даних (Stability). Удосконалено науково-методичний підхід до використання нейромережевих архітектур шляхом обґрунтування доцільності комбінованого застосування вузькоспеціалізованих та універсальних моделей. Результати проведених розрахунків та апробації демонструють суттєву перевагу спеціалізованих моделей над рішеннями загального призначення завдяки їхній попередній підготовці на цільових масивах даних: міжнародних реєстрах вразливостей (CVE) та актуальних зразках шкідливого програмного коду. Аналіз нормативно-правового забезпечення свідчить, що впровадження таких технологій має суворо базуватися на міжнародних стандартах ISO/IEC 27001:2022 та NIST CSF 2.0. Визначено, що високий рівень автоматизації прийняття рішень вимагає системного оновлення законодавчої бази України для чіткого визначення правового статусу та відповідальності за рішення, прийняті за допомогою алгоритмів штучного інтелекту. Подальші наукові розвідки мають бути спрямовані на розробку та впровадження закритих відомчих контурів для розгортання подібних моделей без ризику витоку службової та конфіденційної інформації через публічні хмарні сервіси.

Ключові слова: кібербезпека; штучний інтелект; GPT-5.4-Cyber; Anthropic Mythos; кіберзагрози; інформаційна безпека; правоохоронні органи; Національна поліція України; модель EAI; гібридна війна.

ВСТУП

Постановка проблеми. Сучасні кіберзагрози характеризуються високим рівнем складності та динамічності, що обумовлює необхідність застосування інтелектуальних систем аналізу та реагування [1-3]. В умовах воєнного стану в Україні проблема забезпечення кібербезпеки набуває критичного значення [4; 8]. Особливу увагу привертають спеціалізовані моделі штучного інтелекту, зокрема GPT-5.4-Cyber та Anthropic Mythos, які орієнтовані на задачі кіберзахисту [7], на забезпечення кібербезпеки суспільства, на діяльності Національної поліції України, яка виконує функції протидії кіберзлочинності, захисту інформаційних ресурсів та підтримки правопорядку в цифровому середовищі держави [13].

Аналіз останніх досліджень і публікацій. Проблематика застосування штучного інтелекту у сфері кібербезпеки активно досліджується в сучасній науковій літературі. Роль штучного інтелекту як стратегічного інструменту в сучасній кібервійні та методи оцінювання його функціональності розглядаються у роботах G. Apruzzese, де акцентується увага на необхідності створення нових метрик для вимірювання стійкості AI-систем [1]. A. Buczak та E. Guven здійснили ґрунтовний огляд методів інтелектуального аналізу даних, підкресливши переваги нейромережевих підходів у задачах класифікації кібератак та виявлення вторгнень,що заклало основу для подальших досліджень у цій сфері [2].

Сучасний етап розвитку технологій пов’язаний із появою великих мовних моделей LLM (від англ Large Language Model). Зокрема, M. Ferrag та співавтори у своєму дослідженні 2024 року аналізують революційний вплив LLM на автоматизацію виявлення загроз, зазначаючи, що спеціалізовані моделі демонструють вищу ефективність порівняно з універсальними рішеннями [3]. Важливість дотримання міжнародних стандартів при впровадженні таких систем підтверджується нормами ISO/IEC 27001:2022 [4], а використання бази тактик та технік атак MITRE ATT&CK [5] дозволяє створювати релевантні масиви даних для навчання вузькоспеціалізованих моделей.

Системний підхід до управління ризиками в цифрову епоху також базується на фреймворку NIST CSF 2.0 [6], який визначає критичні аспекти захисту інфраструктури. Окремі прикладні можливості новітніх моделей, таких як GPT-5.4-Cyber, для вирішення практичних завдань кібербезпеки висвітлені у технічних оглядах галузевих експертів [7].

В українському науковому дискурсі правові та організаційні засади інформаційної та кібернетичної безпеки досліджуються через призму Доктрини інформаційної безпеки України [8] та Стратегії кібербезпеки України [9]. Фундаментальні аспекти використання машинного навчання безпосередньо для виявлення мережевих вторгнень у «відкритому світі» детально проаналізовані у працях R. Sommer та V. Paxson [10]. Нормативне регулювання відносин у цій сфері базується на базових законах України: «Про інформацію» [11], «Про захист інформації в інформаційно-комунікаційних системах» [12] та «Про основні засади забезпечення кібербезпеки України» [13].

Незважаючи на значну кількість публікацій, питання розробки інтегральних математичних моделей для комплексної оцінки ефективності саме спеціалізованих нейромережевих архітектур, що враховували б специфіку правоохоронної діяльності (зокрема НПУ), залишається недостатньо висвітленим, тому на сьогодні залишаються невирішеними такі аспекти:

- відсутність формалізованих моделей оцінки ефективності застосування генеративного AI у кібербезпеці;
- недостатнє дослідження спеціалізованих AI-моделей (типу GPT-5.4-Cyber) у контексті практичного застосування;
- відсутність науково обґрунтованих підходів до комбінованого використання різних типів AI;
- недостатня адаптація існуючих підходів до умов воєнного стану та гібридних загроз в Україні.

З огляду на зазначене, дане дослідження спрямовано на:

- розроблення моделі оцінки ефективності застосування AI у кібербезпеці;
- проведення порівняльного аналізу сучасних AI-моделей;
- обґрунтування можливостей їх використання у діяльності правоохоронних органів України.

Мета статті полягає у науковому обґрунтуванні доцільності використання спеціалізованих моделей штучного інтелекту у сфері кібербезпеки, розробці інтегральної математичної моделі для оцінювання їхньої ефективності в умовах воєнного стану, а також у визначенні нормативно-правових засад їх впровадження в діяльність правоохоронних органів України.

Для досягнення поставленої мети визначено такі завдання дослідження:

- Проаналізувати сучасні наукові підходи до використання штучного інтелекту у сфері кібербезпеки.

– Дослідити функціональні можливості спеціалізованих моделей AI, зокрема GPT-5.4-Cyber та Anthropic Mythos.
– Провести порівняльний аналіз спеціалізованих та універсальних AI-моделей у контексті кіберзахисту.
– Розробити модель оцінки ефективності застосування штучного інтелекту у сфері кібербезпеки.
– Визначити особливості та перспективи впровадження таких моделей у діяльність правоохоронних органів України.

Об'єктом дослідження є процеси забезпечення кібербезпеки в умовах сучасних цифрових загроз та воєнного стану.

Предметом дослідження є методи, моделі та програмно-технічні засоби застосування штучного інтелекту для підвищення ефективності кіберзахисту.

Методологія дослідження. Методологічну основу дослідження становить комплекс загальнонаукових та спеціальних методів пізнання, що забезпечують системний аналіз застосування моделей штучного інтелекту у сфері кібербезпеки.

У процесі дослідження використано такі методи:

- аналіз і синтез – для узагальнення наукових підходів до використання штучного інтелекту у кібербезпеці;
- порівняльний метод – для зіставлення функціональних можливостей моделей GPT–5.4–Cyber та Anthropic Mythos;
- системний підхід – для визначення місця AI у загальній системі забезпечення кібербезпеки держави;
- структурно–функціональний аналіз – для дослідження ролі окремих компонентів AI–моделей у процесах кіберзахисту;
- прогностичний метод – для оцінки перспектив розвитку технологій штучного інтелекту у сфері кібербезпеки.

Емпіричну базу дослідження становлять аналітичні матеріали щодо сучасних кіберзагроз, публікації у сфері штучного інтелекту, а також відкриті дані про функціональні можливості сучасних AI-моделей.

Гіпотеза дослідження. Передбачається, що використання спеціалізованих моделей штучного інтелекту у поєднанні з універсальними аналітичними системами забезпечує підвищення ефективності кіберзахисту, що може бути формалізовано через інтегральний показник, який враховує точність виявлення загроз, швидкість реагування, рівень автоматизації та стійкість системи

Наукова новизна. У роботі вперше запропоновано комплексний підхід до оцінки ефективності застосування спеціалізованих моделей штучного інтелекту у сфері кібербезпеки, який базується на інтегральному показнику, що поєднує параметри точності виявлення загроз, швидкості реагування, рівня автоматизації та стійкості системи. На відміну від існуючих підходів, запропонована модель враховує специфіку функціонування систем кіберзахисту в умовах воєнного стану та може бути використана для порівняльного аналізу сучасних AI-рішень, зокрема спеціалізованих моделей типу GPT-5.4-Cyber та Anthropic Mythos.

Удосконалено науково-методичний підхід до використання моделей штучного інтелекту в діяльності правоохоронних органів України шляхом обґрунтування доцільності комбінованого застосування спеціалізованих (кібербезпекових) та універсальних аналітичних моделей. Запропонований підхід дозволяє підвищити ефективність кіберзахисту за рахунок розподілу функцій між моделями (технічний аналіз – стратегічна аналітика), що забезпечує більш повне охоплення кіберзагроз у цифровому середовищі.

Набуло подальшого розвитку наукове уявлення про роль штучного інтелекту в національній системі кібербезпеки України в умовах гібридної війни, зокрема в частині інтеграції AI у процеси кіберрозвідки, аналізу інцидентів та захисту критичної інформаційної інфраструктури. Розширено підходи до нормативного та організаційного забезпечення впровадження таких технологій з урахуванням сучасних викликів і загроз, характерних для держави, що перебуває в умовах збройної агресії.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Аналіз літератури та існуючих стандартних, практичних, де-факто підходів до оцінкі ефективності AI у сфері кібербезпеки показує, що на сьогодні найбільш розповсюдженими є такі метрики та фреймворки:

– CVSS (Common Vulnerability Scoring System): зосереджена на кількісній оцінці небезпеки окремих вразливостей. Хоча вона є галузевим стандартом, її головним недоліком у контексті ШІ є

статичність – вона не враховує динаміку розгортання атаки в часі та адаптивні можливості захисної системи.

– MTTD (Mean Time to Detect) та MTTR (Mean Time to Respond): класичні часові метрики, що використовуються в SOC (Security Operations Centers). Вони дозволяють оцінити швидкість, але не дають уявлення про рівень автоматизації та стійкість системи до специфічних атак на алгоритми машинного навчання (наприклад, adversarial attacks).

– Метрики точності машинного навчання (Precision, Recall, F1-Score): стандартні показники для оцінки якості роботи класифікаторів. Проте для правоохоронних органів та критичної інфраструктури лише показників точності недостатньо, оскільки вони не відображають операційну ефективність моделі в реальному бойовому середовищі.

Враховуючи обмеженість існуючих підходів, виникає необхідність у створенні комплексної моделі, яка б поєднувала технічну точність із показниками операційної швидкості та автономності. Саме тому було розроблено принципово іншу інтегральну модель оцінки ефективності.

Запропонована модель оцінки ефективності. Для кількісної оцінки ефективності використання моделей штучного інтелекту у сфері кібербезпеки доцільно застосувати інтегральний показник ефективності, який враховує ключові параметри функціонування систем кіберзахисту.

Інтегральний показник ефективності можна подати у вигляді:

$$E_{AI} = \alpha D + \beta R + \gamma A + \delta S$$

де:
E_{AI} – інтегральна ефективність використання AI у кібербезпеці;
D – точність виявлення загроз;
R – швидкість реагування;
A – рівень автоматизації;
S – стійкість системи;
$\alpha, \beta, \gamma, \delta$ – коефіцієнти, вони визначаються експертно залежно від задачі. Наприклад для правоохоронних органів України в умовах воєнного стану доцільно збільшувати вагу коефіцієнтів які впливають: на швидкість реагування – β, стійкість системи – δ. Для SOC (Security Operations Center) центрів, що забезпечують цілодобовий моніторинг, виявлення та реагування на кіберзагрози бізнеса: критичними є значення коефіцієнтів D та A. Для об'єктів критичної інфраструктури: підвищити коефіцієнт стійкості системи – δ та значення параметра швидкості реагування R, щоб уникнути простою.

Порівняння запропонованої моделі E_{AI} із галузевими стандартами, такими як NIST CSF та CVSS, дозволяє чітко визначити її місце в ієрархії кібербезпеки. Головна відмінність полягає в тому, що NIST – це стратегічна рамка (фреймворк), CVSS – це метрика вразливості, E_{AI} модель – це метрика ефективності конкретного AI. Порівняльний аналіз використання моделей штучного інтелекту у сфері кібербезпеки наведено у таб. 1.

Таблиця 1

Детальний порівняльний аналіз

Параметр	E_{AI}	NIST CSF (2.0)	CVSS (v4.0)
Об'єкт оцінки	продуктивність AI-моделі (дефенсивний аспект)	загальний стан безпеки організації	ступінь небезпеки окремої вразливості
Detection accuracy	математична точність виявлення (True Positive)	функція Detect (DE): якісний стан моніторингу	не оцінюється (оцінюється лише складність експлуатації)
Response speed	час від виявлення до нейтралізації	функція Respond (RS): процеси реагування	Supplemental Metric: Recovery (здатність системи до відновлення)
Automation level	ступінь автономності AI без участі людини	функція Protect (PR): технології захисту	Metric Automatable: чи може атакувальник автоматизувати атаку
Scalability	стійкість до великих обсягів даних	масштабованість процесів управління	не оцінюється

Запропонована модель E_{AI} є операційною метрикою, яка доповнює стратегічний рівень NIST та технічний рівень CVSS. Вона дозволяє керівнику підрозділу кібербезпеки отримати конкретну цифру, яка характеризує "інтелектуальну потужність" захисту, а не просто перелік вразливостей у системі. Наприклад CVSS надає статичну високу оцінку вразливості, а модель E_{AI} показує, наскільки ефективно

AI "закриває" цю вразливість у реальному часі. Якщо AI має високий коефіцієнт швидкості, то навіть критична вразливість стає менш небезпечною. CVSS зосереджений на тому, як важко зламати систему. Модель E_{AI} зосереджений на тому, наскільки розумно система захищається. E_{AI} оцінює автоматизацію захисту, тоді як CVSS оцінює автоматизацію нападу.

Таким чином модель E_{AI} дозволяє здійснювати порівняльний аналіз AI-рішень (наприклад, GPT-5.4-Cyber, Anthropic Mythos), обґрунтовувати вибір технологій, формалізувати оцінку кіберзахисту.

Практичне застосування моделі E_{AI}. Аналіз джерела [7] показав, що для умовного оцінювання властивостей GPT-5.4-Cyber та Anthropic Mythos, порівняння здійснено на основі експертного підходу. Були прийняти якісні, але суб'єктивні, не формалізовані критерії. Результати зведені в таб. 2.

Таблиця 2

Порівняльний аналіз моделей

Критерій	GPT-5.4-Cyber	Anthropic Mythos
Тип моделі	Спеціалізована (кібербезпека)	Універсальна аналітична
Основне призначення	Defensive cybersecurity	Аналітика та оцінка ризиків
Аналіз вразливостей	Високий рівень (глибокий технічний аналіз)	Середній рівень
Реверс-інжиніринг	Підтримується	Обмежено
Робота зі шкідливим ПЗ	Детальний аналіз	Узагальнений аналіз
Автоматизація SOC/IR	Висока	Середня
Пояснюваність рішень	Середня	Висока
Обмеження доступу	Жорсткий контроль (Trusted Access)	Менш обмежений
Ризик зловживання	Високий (через технічні можливості)	Середній
Найбільш ефективна роль	Технічний кіберзахист	Стратегічна аналітика

З результатів аналізу слідує, що найбільш ефективним є комбіноване використання моделей, де:

– GPT-5.4-Cyber виконує глибокий технічний аналіз,
– Anthropic Mythos – оцінку ризиків та стратегічне прогнозування.

Для перевірки практичної застосовності запропонованої моделі, нами проведено умовну апробацію шляхом порівняльного оцінювання ефективності використання спеціалізованої моделі GPT-5.4-Cyber та універсальної аналітичної моделі Anthropic Mythos. Оцінювання характеристик моделей було здійснено на основі експертного підходу з урахуванням функціональних можливостей, описаних у сучасних наукових та аналітичних джерелах, що відповідає практиці досліджень у сфері кібербезпеки [1, 2].

Технічно оцінювання здійснено на з використанням шкали від 0 до 1, де 1 відповідає максимальному рівню показника.

Прийнято такі значення вагових коефіцієнтів (для умов діяльності правоохоронних органів України в умовах воєнного стану):

$\alpha = 0{,}35$ (точність виявлення загроз)
$\beta = 0{,}30$ (швидкість реагування)
$\gamma = 0{,}20$ (рівень автоматизації)
$\delta = 0{,}15$ (стійкість системи)

Результати розрахунку для моделей:

1. GPT-5.4-Cyber:
$D = 0{,}90$
$R = 0{,}85$
$A = 0{,}80$
$S = 0{,}75$
$E_{AI}^{GPT} = 0.35 \cdot 0.90 + 0.30 \cdot 0.85 + 0.20 \cdot 0.80 + 0.15 \cdot 0.75$
Результат $E_{AI}^{GPT} \approx 0{,}835$
2. Anthropic Mythos:
$D = 0{,}75$
$R = 0{,}70$
$A = 0{,}65$
$S = 0{,}80$
$E_{AI}^{Mythos} = 0.35 \cdot 0.75 + 0.30 \cdot 0.70 + 0.20 \cdot 0.65 + 0.15 \cdot 0.80$
Результат $E_{AI}^{Mythos} \approx 0{,}715$

Аналіз розрахунків підтверджують гіпотезу про те, що вузькопрофільне навчання на масивах кіберзагроз суттєво підвищує точність *D* та швидкість *R* системи. Отримані результати свідчать, що GPT-5.4-Cyber демонструє вищий рівень ефективності у задачах технічного кіберзахисту, тоді як Anthropic Mythos є більш придатною для стратегічного аналізу, що повністю співпадаєз результатами таблиці №1. Таким чином доведено, що комбіноване використання моделей дозволяє підвищити загальний рівень кібербезпеки за рахунок поєднання їх функціональних переваг.

Слід зазначити, що у зв'язку з високим рівнем новизни спеціалізованих моделей штучного інтелекту у сфері кібербезпеки, зокрема GPT-5.4-Cyber, їх відображення у наукових публікаціях є обмеженим, що обумовлює необхідність використання комплексного підходу із залученням як академічних, так і аналітичних джерел.

Отримані результати математичного моделювання підтверджують високу технологічну ефективність спеціалізованих моделей, проте їхнє практичне впровадження в оперативну діяльність підрозділів кіберполіції не може базуватися виключно на показниках продуктивності. Високий рівень автоматизації, зафіксований у моделі E_{AI}, вимагає чіткого визначення меж відповідальності та правового статусу рішень, прийнятих за допомогою AI. Це зумовлює необхідність аналізу чинного законодавчого поля, яке має стати фундаментом для безпечної інтеграції таких технологій у державну систему кіберзахисту.

Нормативно-правове забезпечення застосування AI у сфері кібербезпеки в Україні регулюється нормативно-правовими актами [8-9, 11-13], зокрема:

- законодавством у сфері кібербезпеки;
- актами щодо захисту інформації;
- стратегічними документами державного рівня.

Це створює правову основу для інтеграції інноваційних технологій у діяльність правоохоронних органів та забезпечує правове регулювання застосування технологій кібербезпеки в Україні, яке здійснюється на основі таких нормативно-правових актів:

- Закон України «Про основні засади забезпечення кібербезпеки України» від 05.10.2017 № 2163-VIII. Визначає правові та організаційні основи забезпечення кібербезпеки держави, а також повноваження суб'єктів сектору безпеки і оборони.
- Закон України «Про інформацію» від 02.10.1992 № 2657-XII. Регулює відносини щодо створення, збирання, зберігання та захисту інформації.
- Закон України «Про захист інформації в інформаційно-телекомунікаційних системах» від 05.07.1994 № 80/94-ВР. Визначає вимоги до технічного та криптографічного захисту інформації.
- Доктрина інформаційної безпеки України, затверджена Указом Президента України від 25.02.2017 № 47/2017. Визначає стратегічні пріоритети державної політики у сфері інформаційної безпеки.
- Стратегія кібербезпеки України (нова редакція, 2021 р.). Передбачає розвиток національної системи кібербезпеки та впровадження інноваційних технологій, зокрема штучного інтелекту.

Аналіз нормативно-правого забезпечення застосування AI у сфері кібербезпеки у міжнародної практиці свідчить, що впровадження інтелектуальних систем у сфері кіберзахисту базується на поєднанні жорстких стандартів управління інформаційною безпекою та гнучких фреймворків, що адаптуються під можливості нових технологій:

- Оновлений міжнародний стандарт ISO/IEC 27001:2022 є фундаментом для легітимізації AI-рішень у корпоративному та державному секторах. Він встановлює вимоги до створення Систем управління інформаційною безпекою.
- Стратегічне управління забезпечено NIST (від англ. National institute of standards and technology, Національний інститут стандартів і технологій США) Cybersecurity Framework 2.0. Нова редакція якого розширює сферу застосування кіберзахисту у загальній стратегії безпеки організації.
- Спеціалізовані рішення та правові виклики (на прикладі GPT-5.4-Cyber). Поява таких моделей, як GPT-5.4-Cyber, демонструє технологічний розрив між можливостями AI та наявним правовим регулюванням.

Таким чином вітчизняний та закордонний досвід показує, що нормативне забезпечення рухається шляхом інтеграції AI в існуючі системи управління безпекою ISO та стратегічні рамки NIST. Для України важливо використовувати ці стандарти як базис для сертифікації спеціалізованих моделей AI, що використовуються у правоохоронній діяльності та захисті критичної інфраструктури.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

У цієї оригінальній статті проаналізовано та доведено, що використання вузькопрофільних моделей штучного інтелекту, таких як GPT-5.4-Cyber, забезпечує значно вищу точність детекції та

швидкість аналізу специфічних кіберзагроз порівняно з універсальними LLM AI. Висока ефективність функціонування спеціалізованих систем AI в галузі кібербезпеки зумовлена їхньою попередньою підготовкою на цільових масивах даних. Зокрема, процес навчання охоплює роботу з міжнародними реєстрами загальновідомих вразливостей та експозицій програмного забезпечення, що дозволяє моделі розпізнавати технічні недоліки в архітектурі цифрових систем. Крім того, критично важливим етапом є аналіз великих обсягів вихідного коду шкідливих програмних продуктів, таких як комп'ютерні віруси, мережеві хробаки та програми-вимагачі. Такий підхід дає змогу алгоритмам штучного інтелекту вивчати логіку дій зловмисників, виявляти приховані загрози та прогнозувати потенційні методи атак ще до моменту їхнього фактичного здійснення

Запропонована інтегральна модель оцінки ефективності $E_{AI}=\alpha D+\beta R+\gamma A+\delta S$ дозволяє правоохоронним органам та установам критичної інфраструктури здійснювати математично обґрунтований вибір AI-рішень, виходячи з пріоритетів безпеки.

Встановлено, що в умовах гібридної війни впровадження AI створює нові вектори загроз (отруєння даних, змагальні атаки). Це потребує розробки проактивних стратегій захисту самих AI-систем, що є критичним для забезпечення стійкості правоохоронного сектору.

Визначено, що високий рівень автоматизації інтелектуальних систем випереджає чинне нормативно-правове регулювання в Україні. Для легітимізації рішень, прийнятих за допомогою AI, необхідне оновлення законодавчої бази, зокрема в частині визначення доказової сили результатів автоматизованого кібермоніторингу [13].

Подальші дослідження мають бути спрямовані на створення закритих відомчих контурів для розгортання подібних моделей, що дозволить використовувати переваги GPT-5.4-Cyber без ризику витоку конфіденційної інформації через хмарні сервіси.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Apruzzese, G., et al. (2023). The role of artificial intelligence in cybersecurity: A survey. *ACM Computing Surveys, 55*(4), Article 58. https://doi.org/10.1145/3546933
2. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials, 18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502
3. Ferrag, M. A., et al. (2024). Revolutionizing cybersecurity with large language models: A comprehensive survey. *IEEE Access, 12*, 10415–10444. https://doi.org/10.1109/ACCESS.2024.3354356
4. International Organization for Standardization. (2022). *Information security, cybersecurity and privacy protection — Information security management systems — Requirements* (ISO/IEC Standard No. 27001:2022). https://www.iso.org/standard/27001
5. MITRE. (2024). *MITRE ATT&CK® v15*. https://attack.mitre.org/
6. National Institute of Standards and Technology. (2024). *The NIST Cybersecurity Framework (CSF) 2.0*. U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.29
7. NV Techno. (2024, May 14). *GPT-5.4-Cyber: Shcho vmiie nova AI-model dlia kiberbezpeky* [GPT-5.4-Cyber: What the new AI model for cybersecurity can do]. https://techno.nv.ua/ukr/it-industry/gpt-5-4-cyber-shcho-vmiye-nova-shi-model-vid-openai-dlya-kiberbezpeki-50600407.html
8. President of Ukraine. (2017). *Pro Doktrynu informatsiinoi bezpeky Ukrainy* [On the Doctrine of Information Security of Ukraine] (Decree No. 47/2017). https://zakon.rada.gov.ua/laws/show/47/2017
9. President of Ukraine. (2021). *Pro rishennia Rady natsionalnoi bezpeky i oborony Ukrainy vid 14 travnia 2021 roku "Pro Stratehiiu kiberbezpeky Ukrainy"* [On the decision of the National Security and Defense Council of Ukraine of May 14, 2021 "On the Cybersecurity Strategy of Ukraine"] (Decree No. 447/2021). https://zakon.rada.gov.ua/laws/show/447/2021
10. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In *2010 IEEE Symposium on Security and Privacy* (pp. 305–316). https://doi.org/10.1109/SP.2010.25
11. Verkhovna Rada of Ukraine. (1992). *Pro informatsiiu* [On information] (Law No. 2657-XII). https://zakon.rada.gov.ua/laws/show/2657-12
12. Verkhovna Rada of Ukraine. (1994). *Pro zakhyst informatsii v informatsiino-komunikatsiinykh systemakh* [On the protection of information in information and communication systems] (Law No. 80/94-VR). https://zakon.rada.gov.ua/laws/show/80/94-вр
13. Verkhovna Rada of Ukraine. (2017). *Pro osnovni zasady zabezpechennia kiberbezpeky Ukrainy* [On the basic principles of ensuring cybersecurity of Ukraine] (Law No. 2163-VIII). https://zakon.rada.gov.ua/laws/show/2163-19

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (34), 2026
ISSN 2663 – 4023

Vitaliy Svitlychny
PhD, Associate Professor of the Department of Counteracting Cybercrime
Kharkiv National University of Internal Affairs, Kharkiv, Ukraine
ORCID:0000-0003-3381-3350
vit.svet@ukr.net

Petro Klimushyn
PhD, Associate Professor of the Department of Counteracting Cybercrime
Kharkiv National University of Internal Affairs, Kharkiv, Ukraine
ORCID:0000-0002-1020-9399
klimushyn@ukr.net

Yuriy Onyshchenko
PhD, Associate Professor, Deputy Director of the Institute for Educational andResearch Activities of Educational and Scientific Institute No. 4,
Kharkiv National University of Internal Affairs, Kharkiv, Ukraine
ORCID:0000-0002-7755-3071
onischenko1980@gmail.com

ASSESSMENT OF THE EFFECTIVENESS OF SPECIALIZED ARTIFICIAL INTELLIGENCE MODELS IN THE CYBER DEFENSE SYSTEM OF UKRAINE: TECHNICAL AND LEGAL ASPECT

Abstract. The article explores the possibilities of using specialized artificial intelligence (AI) models in the field of cybersecurity, in particular, using the latest GPT-5.4-Cyber model as an example. Modern cyber threats are characterized by a high level of complexity, dynamism, and cross-border nature, which creates an urgent need for the use of intelligent systems for automated analysis and rapid response. This issue becomes particularly relevant in the context of martial law and hybrid warfare, where Ukraine's digital infrastructure is becoming the object of constant attacks. The study analyzes in detail the functional features of the model, its role in ensuring national cyber defense, as well as potential risks and prospects for its implementation in the daily activities of law enforcement agencies of Ukraine. Special attention is paid to the activities of units of the National Police of Ukraine, which perform key functions in combating cybercrime, protecting state information resources, and maintaining law and order in the digital environment. An author's mathematical model for a comprehensive assessment of the effectiveness of intelligent security systems (E_{AI}) is proposed, which is based on the calculation of an integral indicator. This indicator combines such critical parameters as the accuracy of anomaly detection (Дetection), the speed of automated response (Response), the level of system autonomy (Automation), and resistance to external manipulation or data poisoning (Stability). The scientific and methodological approach to the use of neural network architecture by substantiating the feasibility of combined use of highly specialized and universal models. The results of the calculations and testing demonstrate a significant advantage of specialized models over general-purpose solutions due to their preliminary training on target data sets: international vulnerability registers (CVE) and current samples of malicious code. The analysis of regulatory support shows that the implementation of such technologies should be strictly based on international standards ISO/IEC 27001:2022 and NIST CSF 2.0. It was determined that a high level of automation of decision-making requires a systematic update of the legislative framework of Ukraine to clearly define the legal status and responsibility for decisions made using artificial intelligence algorithms. Further scientific research should be aimed at developing and implementing closed departmental circuits for the deployment of such models without the risk of leakage of official and confidential information through public cloud services.

Keywords: cybersecurity; artificial intelligence; GPT-5.4-Cyber; Anthropic Mythos; cyber threats; information security; law enforcement; National Police of Ukraine; E_{AI} model; hybrid warfare.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Apruzzese, G., et al. (2023). The role of artificial intelligence in cybersecurity: A survey. *ACM Computing Surveys, 55*(4), Article 58. https://doi.org/10.1145/3546933
2. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials, 18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502
3. Ferrag, M. A., et al. (2024). Revolutionizing cybersecurity with large language models: A comprehensive survey. *IEEE Access, 12*, 10415–10444. https://doi.org/10.1109/ACCESS.2024.3354356
4. International Organization for Standardization. (2022). *Information security, cybersecurity and privacy protection — Information security management systems — Requirements* (ISO/IEC Standard No. 27001:2022). https://www.iso.org/standard/27001
5. MITRE. (2024). *MITRE ATT&CK® v15*. https://attack.mitre.org/
6. National Institute of Standards and Technology. (2024). *The NIST Cybersecurity Framework (CSF) 2.0*. U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.29
7. NV Techno. (2024, May 14). *GPT-5.4-Cyber: Shcho vmiie nova AI-model dlia kiberbezpeky* [GPT-5.4-Cyber: What the new AI model for cybersecurity can do]. https://techno.nv.ua/ukr/it-industry/gpt-5-4-cyber-shcho-vmiye-nova-shi-model-vid-openai-dlya-kiberbezpeki-50600407.html
8. President of Ukraine. (2017). *Pro Doktrynu informatsiinoi bezpeky Ukrainy* [On the Doctrine of Information Security of Ukraine] (Decree No. 47/2017). https://zakon.rada.gov.ua/laws/show/47/2017
9. President of Ukraine. (2021). *Pro rishennia Rady natsionalnoi bezpeky i oborony Ukrainy vid 14 travnia 2021 roku "Pro Stratehiiu kiberbezpeky Ukrainy"* [On the decision of the National Security and Defense Council of Ukraine of May 14, 2021 "On the Cybersecurity Strategy of Ukraine"] (Decree No. 447/2021). https://zakon.rada.gov.ua/laws/show/447/2021
10. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In *2010 IEEE Symposium on Security and Privacy* (pp. 305–316). https://doi.org/10.1109/SP.2010.25
11. Verkhovna Rada of Ukraine. (1992). *Pro informatsiiu* [On information] (Law No. 2657-XII). https://zakon.rada.gov.ua/laws/show/2657-12
12. Verkhovna Rada of Ukraine. (1994). *Pro zakhyst informatsii v informatsiino-komunikatsiinykh systemakh* [On the protection of information in information and communication systems] (Law No. 80/94-VR). https://zakon.rada.gov.ua/laws/show/80/94-вр
13. Verkhovna Rada of Ukraine. (2017). *Pro osnovni zasady zabezpechennia kiberbezpeky Ukrainy* [On the basic principles of ensuring cybersecurity of Ukraine] (Law No. 2163-VIII). https://zakon.rada.gov.ua/laws/show/2163-19

Отримано редакцією журналу / Received: 25.02.26
Прорецензовано / Revised: 05.03.26
Схвалено до друку / Accepted: 24.09.26