УДК 004.42:004.9:681.3 DOI https://doi.org/10.35546/kntu2078-4481.2026.1.38

С. О. ПРИМИСЬКА
кандидат технічних наук, доцент,
доцент кафедри технічних та програмних засобів автоматизації
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0002-5832-0686

В. М. КУЛІКОВ
кандидат технічних наук, доцент
Інститут спеціального зв'язку та захисту інформації
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: 0000-0002-1015-5802

О. О. БАЗИЛЬ
кандидат фізико-математичних наук,
старший викладач кафедри комп'ютерних наук
Сумський державний університет
ORCID: 0000-0002-2644-5361

ЗАСТОСУВАННЯ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ПРОГНОЗУВАННЯ ВІДМОВ У ХМАРНИХ ІНФРАСТРУКТУРАХ

У сучасному світі хмарні інфраструктури стали невіддільною частиною бізнес-процесів, що зумовлює необхідність забезпечення їхньої безперебійності та стійкості. В умовах швидкого розвитку технологій штучного інтелекту (далі – ШІ) виникає потреба їх застосування для прогнозування відмов у хмарних системах. Метою статті є дослідження різних моделей і методів ШІ, що застосовуються для прогнозування відмов, а також оцінка їхньої ефективності, переваг і викликів, які постають перед організаціями під час упровадження таких технологій. У дослідженні застосовано теоретичні (аналіз, синтез, абстрагування, індукція, дедукція) та емпіричні методи, зокрема опис. Застосування технологій ШІ для прогнозування відмов у хмарних інфраструктурах охоплює аналіз технічних збоїв, аномалій у продуктивності, кіберзагроз, відмов в обслуговуванні, порушень інтеграції та чинників, пов'язаних із людськими помилками, що дозволяє підвищити надійність систем за допомогою моделей і методів, таких як Random Forest, XGBoost, LSTM, GRU та технологій AIOps. У статті здійснено огляд різних моделей, а також методів прогнозної аналітики часових рядів та виявлення аномалій. Розглянуто AIOps-технології, які забезпечують автоматизацію процесів моніторингу та управління ризиками. Проаналізовано потенційні переваги впровадження ШІ в хмарних системах, зокрема підвищення точності прогнозування, можливість обробки великих обсягів даних і раннє виявлення аномалій. Водночас розглядаються й виклики, які постають перед організаціями, як-от потреба у великих наборах даних, висока обчислювальна вартість та ризики кібербезпеки. На основі аналізу надано практичні рекомендації для інтеграції технологій ШІ в процеси моніторингу та управління ризиками. Рекомендовано застосовувати цифрові двійники, моделі ШІ для раннього виявлення аномалій, графові нейронні мережі та MLOps-інструменти для автоматизації тестування хмарної інфраструктури. Це дасть змогу мінімізувати ризики, пов'язані з оновленням, і підвищити стійкість системи. Отже, застосування технологій ШІ для прогнозування відмов у хмарних інфраструктурах може значно підвищити їхню надійність і продуктивність, проте потребує ретельного планування та врахування можливих викликів.

Ключові слова: *глибинне навчання, машинне навчання, виявлення аномалій, AIOps, MLOps, Random Forest, XGBoost, LSTM, GRU.*

S. O. PRYMYSKA
Candidate of Technical Sciences, Associate Professor,
Associate Professor at the Automation Hardware and Software Department
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"
ORCID: 0000-0002-5832-0686

V. M. KULIKOV
Candidate of Technical Sciences, Associate Professor
Institute of Special Communications and Information Protection
of National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"
ORCID: 0000-0002-1015-5802

O. O. BAZYL
Candidate of Physical and Mathematical Sciences,
Senior Lecturer at the Department of Computer Science
Sumy State University
ORCID: 0000-0002-2644-5361

USING ARTIFICIAL INTELLIGENCE TECHNOLOGIES TO PREDICT FAILURES IN CLOUD INFRASTRUCTURES

In the modern world, cloud infrastructures have become an integral part of business processes, which necessitates the need to ensure their continuity and stability. In the context of the rapid development of artificial intelligence (hereinafter referred to as AI), there is a need to use them to predict failures in cloud systems. The purpose of the article is to study various AI models and methods used for predicting failures, as well as to assess their effectiveness, advantages and challenges that organizations face when implementing such technologies. The study uses theoretical (analysis, synthesis, abstraction, induction, deduction) and empirical methods, including description. The use of AI technologies to predict failures in cloud infrastructures includes analysis of technical failures, performance anomalies, cyber threats, service failures, integration issues, and human error factors, which helps improve system reliability using models and methods such as Random Forest, XGBoost, LSTM, GRU, and AIOps technologies. The article reviews various models and methods for predictive time series analytics and anomaly detection. It examines AIOps technologies that automate monitoring and risk management processes. It analyzes the potential benefits of implementing AI in cloud systems, including improved forecasting accuracy, the ability to process large amounts of data, and early detection of anomalies. At the same time, it also examines the challenges facing organizations, such as the need for large data sets, high computational costs, and cybersecurity risks. Based on the analysis, practical recommendations are provided for integrating AI technologies into monitoring and risk management processes. It is recommended to use digital twins, AI models for early anomaly detection, graph neural networks, and MLOps tools to automate testing of cloud infrastructure. This will minimize the risks associated with updates and increase the resilience of the system. Therefore, the use of AI technologies for failure prediction in cloud infrastructures can significantly increase their reliability and performance, but requires careful planning and consideration of possible challenges.

Key words: *deep learning, machine learning, anomaly detection, AIOps, MLOps, Random Forest, XGBoost, LSTM, GRU.*

Постановка проблеми

Сучасні хмарні інфраструктури є невіддільною частиною інформаційних технологій, забезпечуючи організаціям гнучкість, масштабованість та ефективне управління ресурсами. Проте зростання залежності від хмарних рішень підвищує ризик виникнення відмов, які можуть призвести до фінансових збитків, втрати інформації та зниження довіри користувачів. У цьому контексті своєчасне виявлення та прогнозування можливих проблем є важливим для забезпечення стійкості та безперебійності бізнес-процесів.

Сучасні методи моніторингу та управління хмарними інфраструктурами нерідко виявляються недостатньо дієвими для своєчасного реагування на потенційні загрози. Використання статичних правил чи простих алгоритмів не забезпечує належної чутливості до динаміки процесів, унаслідок чого підвищується ризик раптових відмов. З огляду на це, постає потреба застосування більш адаптивних технологій, здатних прогнозувати критичні стани на основі аналізу великих масивів даних.

У цьому контексті технології ШІ відкривають нові можливості для розв'язання зазначених проблем. Алгоритми машинного навчання (далі – МН) та глибокого навчання (далі – ГН) здатні аналізувати історичні дані про роботу системи, виявляти патерни та аномалії, що можуть свідчити про потенційні відмови. Це не лише підвищує точність прогнозування, але й автоматизує процеси моніторингу та управління ризиками, що є важливим аспектом для підтримки стабільності хмарних інфраструктур.

Актуальність дослідження зумовлена стрімким розвитком інформаційних технологій і зростанням обсягів даних, що обробляються в хмарних системах, унаслідок чого традиційні підходи до управління вже не забезпечують належної ефективності. Інтеграція технологій ШІ в процеси прогнозування відмов дає змогу підвищити надійність хмарних інфраструктур, скоротити витрати на їхнє обслуговування та зміцнити загальний рівень безпеки.

Аналіз останніх досліджень і публікацій

Хмарні технології активно розвиваються, що посилює потребу в ефективному моніторингу та управлінні ризиками. Аналіз сучасних досліджень дає змогу виявити наявні прогалини та формує підґрунтя для подальших наукових розробок. Так, Ф. Воутсас (F. Voutsas), Х. Віолос (J. Violos), А. Лівадіас (A. Leivadeas) представляють механізм подолання надмірності сповіщень у моніторингу хмарних інфраструктур на основі методів машинного навчання. Запропонована система фільтрує нерелевантні події, підвищуючи точність реагування адміністраторів на понад 90 %. Використання реальних даних компанії Netdata підкреслює практичну цінність дослідження [1].

Стратегії оцінки та мінімізації ризиків у хмарних середовищах розглядають О. Агбула та колеги (O. Agboola et al.), підкреслюючи. роль інноваційних технологій (ШІ, Zero Trust моделей, автоматизації відповідності вимогам) у забезпеченні стійкості хмарних інфраструктур до сучасних загроз [2].

Управління ризиками хмарної безпеки в контексті Індустрії 4.0 аналізують С. Дріс (S. Drissi), М. Шарджі (M. Chergui), З. Хатер (Z. Khatar), акцентуючи на важливості застосування ШІ та МН для виявлення загроз і необхідності адаптивних моделей управління ризиками для складних хмарних середовищ [3].

На ролі даних і методів МН у хмарних та туманних обчисленнях наголошує Ю. Козак, розглядаючи специфіку fog-платформ, їхні сценарії та підходи до прогнозування, зокрема статистичні моделі, дерева рішень та нейронні мережі (далі – НМ) [4].

Вплив нейронних мереж на кібербезпеку досліджують О. Кушнерьов, І. Позовна, В. Сокол. Автори зазначають, що НМ здатні автоматизувати виявлення загроз та ефективно інтегруватися з сучасними технологіями безпеки [5].

Значення ШІ та МН для оптимізації DevOps-процесів підкреслюють О. Дуда, І. Шаклейна, М. Лучкевич. Науковці вивчають проблеми традиційних підходів, можливості аналізу логів, прогнозування технічних перешкод і перспективи створення автономних DevOps-агентів [6].

Трансформацію кібербезпеки США під впливом ШІ аналізують А. Адевус та співавтори (A. O. Adewus et al.), висвітлюючи переваги ШІ в захисті важливих національних систем та акцентуючи на етичних і регуляторних аспектах його впровадження [7].

Застосування систем ШІ для підвищення продуктивності та адаптивності систем безпеки досліджують І. Буров (Y. Burov), І. Жовнір (Y. Zhovnir), О. Захарія (O. Zakharya). Запропонована ними система ситуаційно-орієнтованої обробки інцидентів у реальному часі використовує онтологію GFO та базу знань зі сценаріями типових загроз, що важливо для хмарних інфраструктур [8].

На значенні ШІ в кібербезпеці в умовах Четвертої промислової революції наголошують І. Саркер (I. Sarker), М. Фурхад (M. Furhad), Р. Новроз (R. Nowrozy). Науковці визначають роль МН, гібридних моделей, NLP та експертних систем в автоматизації захисту та протидії кібератакам [9].

На прогнозуванні навантаження хмарних ресурсів за допомогою нейронних мереж зосереджуються В. Давидов (V. Davydov) та Д. Гребенюк (D. Hrebeniuk). Експерименти підтверджують можливість значного підвищення ефективності використання ресурсів через удосконалені моделі прогнозування [10].

Загалом, аналіз літератури демонструє швидкий розвиток досліджень, присвячених застосуванню ШІ для прогнозування відмов та збереження безпеки в хмарних системах. Інтеграція інтелектуальних технологій дає змогу підвищувати ефективність, але залишається низка викликів, зокрема якість даних, стандартизація та складність системної інтеграції. Поточне дослідження спрямоване на подолання цих проблем шляхом розроблення нових підходів, що покращать точність прогнозування та знизять ризики для організацій, які застосовують хмарні технології.

Формулювання мети дослідження

Мета статті – дослідити застосування технологій ШІ для прогнозування відмов у хмарних інфраструктурах, оцінити їхню ефективність, переваги та виклики, що виникають під час упровадження таких технологій в організаціях; розробити практичні рекомендації для інтеграції ШІ в процеси моніторингу та управління ризиками, щоб забезпечити безперебійність і стійкість роботи хмарних систем.

Завдання роботи:

1) дослідити різні моделі та методи ШІ для прогнозування відмов, зосереджуючись на їхній ефективності та застосуванні в контексті хмарних технологій;

2) проаналізувати потенційні переваги (зокрема, підвищення точності прогнозування, зниження операційних витрат та покращення загальної продуктивності систем) та виклики, що постають перед організаціями під час упровадження таких технологій;

3) розробити практичні рекомендації для організацій щодо інтеграції технологій ШІ в процеси моніторингу та управління ризиками для забезпечення стійкості та безперебійності функціонування хмарних систем.

Викладення основного матеріалу дослідження

Прогнозування відмов у хмарних інфраструктурах – важливе завдання для мінімізації ризиків, пов'язаних із простоєм систем, та підвищення їхньої надійності. Для досягнення цієї мети застосовуються різні технології ШІ, серед яких основними є МН, НМ та аналіз даних. Огляд технологій ШІ для прогнозування відмов у хмарних інфраструктурах наведено в таблиці 1.

У табл. 1 представлено різні категорії технологій ШІ, які застосовуються для прогнозування відмов у хмарних інфраструктурах. Кожна категорія містить специфічні методи та моделі, а також їхні характеристики, що демонструє важливість цих технологій для підвищення надійності систем. У категорії МН представлено такі моделі, як Random Forest, XGBoost, LightGBM, SVM, KNN, Naive Bayes і логістична регресія. Ці методи здатні аналізувати історичні дані для прогнозування відмов, що дає змогу класифікувати стани системи та оцінювати ризики [11, с. 20073; 12, с. 13]. Застосування МН у цьому контексті забезпечує проактивне виявлення потенційних проблем, що необхідно для підтримки безперебійності бізнес-процесів.

Таблиця 1

Технології ШІ для прогнозування відмов у хмарних інфраструктурах

Категорія технологій	Методи / моделі	Характеристика застосування
МН	Random Forest, XGBoost, LightGBM, SVM, KNN, Naive Bayes, Логістична регресія	Прогнозування відмов за історичними даними, класифікація станів системи, оцінка ризику
ГН	LSTM, GRU, CNN, Time Series Transformers, Autoencoders	Передбачення аномалій у часових рядах, виявлення прихованих патернів, раннє попередження збоїв
Прогностична аналітика часових рядів	ARIMA, SARIMA, Prophet, VAR	Прогноз навантаження, ресурсів, піків, трендів, деградації продуктивності
Виявлення аномалій	Isolation Forest, LOF, One-Class SVM, Gaussian Mixture Models, Reconstruction Error (Autoencoders)	Виявлення відхилень від нормальної поведінки, що передують відмові
Підкріплювальне навчання	Q-Learning, DQN, Policy Gradient Methods	Адаптивне масштабування, автоматичне реагування, мінімізація ризику збоїв
Аналіз логів і подій	LogBERT, LogGPT, BERT for Logs, Seq2Seq, K-Means, DBSCAN	Виявлення аномальних подій у логах, прогноз наступних лог-записів, кореляція подій
Графові нейронні мережі	GCN, GraphSAGE, GAT	Прогноз каскадних відмов, аналіз взаємозалежностей у мікросервісах, побудова графів доступності
Платформи та фреймворки ШІ/МН для хмари	AWS AI/CloudWatch, Azure ML, Vertex AI, Prometheus ML, KubeFlow, Dynatrace Davis AI, Datadog AIOps	Готові інструменти для моніторингу, прогнозування відмов, автоматизації операцій
AIOps-технології	Moogsoft, BigPanda, IBM Watson AIOps, Splunk ITSI, Dynatrace AI Engine	Автоматична кореляція подій, прогноз інцидентів, аналіз причин відмов, усунення збоїв

Джерело: створено автором на основі [11, с. 20073; 12, с. 13; 13, с. 262; 14, с. 3942; 15, с. 747; 16, с. 104;17, с. 6]

ГН, представлене моделями LSTM, GRU, CNN, Time Series Transformers і Autoencoders, має значний потенціал у прогнозуванні аномалій у часових рядах. Ці моделі можуть виявляти приховані патерни в даних, що сприяє ранньому запобіганню порушень [13, с. 262]. Застосування ГН дає змогу системам адаптуватися до динамічних змін навантаження та умов експлуатації, що підвищує їхню загальну надійність.

Методи прогностичної аналітики часових рядів, такі як ARIMA, SARIMA, Prophet і VAR, застосовуються для прогнозування навантаження, використання ресурсів та пікових значень [14, с. 3942]. Вони дають змогу аналізувати тренди й деградацію продуктивності, що є важливими аспектами для управління ресурсами у хмарних інфраструктурах. Прогнозування на основі часових рядів забезпечує проактивний підхід до управління ризиками.

Виявлення аномалій – категорія, що охоплює методи Isolation Forest, LOF, One-Class SVM і Gaussian Mixture Models. Ці моделі допомагають виявляти відхилення від нормальної поведінки системи, які свідчать про наближення відмов. Завдяки цим технологіям можна оперативно реагувати на потенційні загрози, що знижує ймовірність виникнення серйозних проблем.

Підкріплювальне навчання, представлене такими методами, як Q-Learning, DQN і Policy Gradient Methods, застосовується для автоматичного адаптивного масштабування ресурсів та реагування на зміни навантаження. Цей підхід дає змогу мінімізувати ризик порушень шляхом самонавчання системи на основі отриманого досвіду [15, с. 747]. Відповідно, підкріплювальне навчання сприяє створенню більш стійких і адаптивних хмарних інфраструктур.

Аналіз логів і подій передбачає застосування таких моделей, як LogBERT, LogGPT і BERT for Logs. Ці технології дають змогу виявляти аномальні події в логах та прогнозувати послідовність лог-записів. Кореляція подій – важливий аспект цього процесу, оскільки вона допомагає встановити першопричини потенційних збоїв та оптимізувати моніторинг системи.

Графові нейронні мережі (GCN, GraphSAGE, GAT) пропонують нові можливості для прогнозування каскадних відмов та аналізу взаємозалежностей між мікросервісами. Побудова графів доступності допомагає візуалізувати та аналізувати складні зв'язки в системах, що може бути корисним для виявлення вузьких місць в архітектурі.

На ринку існує низка платформ та фреймворків для реалізації ШІ/МН у хмарному середовищі, таких як AWS AI/CloudWatch, Azure ML та Vertex AI. Ці інструменти надають готові рішення для моніторингу та прогнозування порушень, що спрощує процес інтеграції ШІ в бізнес-процеси [16, с. 104].

AIOps-технології (Moogsoft, BigPanda, IBM Watson AIOps) забезпечують автоматичну кореляцію подій та прогнозування інцидентів. Застосування цих технологій дає змогу аналізувати першопричини відмов та ефективно

їх усувати. Завдяки AIOps організації можуть значно підвищити свою оперативність і скоротити середній час до відновлення [17, с. 6].

Аналіз технологій ШІ для прогнозування відмов у хмарних інфраструктурах демонструє їхню важливу роль у забезпеченні надійності та ефективності систем. Кожна категорія технологій має специфічні методи та моделі, які взаємодоповнюють одна одну у створенні інтегрованого підходу до управління ризиками. Застосування цих технологій не лише знижує ймовірність відмов, а й оптимізує управління ресурсами в умовах постійно змінюваного навантаження.

У табл. 2 наведено переваги й обмеження застосування ШІ для прогнозування відмов у хмарних інфраструктурах.

Таблиця 2

Переваги та обмеження застосування ШІ для прогнозування відмов у хмарних інфраструктурах

Переваги	Обмеження
Висока точність прогнозування	Висока потреба у великих наборах даних
Можливість обробки великих обсягів телеметрії та логів	Висока обчислювальна вартість
Адаптивність до змін середовища	Необхідність спеціалістів із МН/ШІ
Раннє виявлення аномалій	Можливі хибні спрацювання
Автоматизація операцій	Складність інтеграції з наявною хмарною інфраструктурою
Підвищення надійності та доступності сервісів	Вразливість моделей до «drift»
Зменшення впливу людського чинника	Ризики кібербезпеки
Оптимізація витрат на інфраструктуру	Неоднорідність даних різних мікросервісів

Джерело: створено автором на основі [18, с. 104; 19, с. 503]

Застосування ШІ для прогнозування відмов у хмарних інфраструктурах має значні переваги, які роблять його привабливим для багатьох організацій. Насамперед моделі МН та ГН демонструють високу точність прогнозування, оскільки здатні виявляти приховані закономірності в даних, які часто залишаються непоміченими традиційними методами. Це дає змогу ефективніше реагувати на потенційні проблеми проактивно.

Крім того, ШІ здатен обробляти значні обсяги телеметрії та логів у режимі реального часу, що важливо в сучасному середовищі з мільйонами метрик і подій. Адаптивність моделей до змін у навантаженні або топології хмари також є перевагою, оскільки це дає змогу системам швидко реагувати на динамічні умови.

Одним зі значущих аспектів застосування ШІ є раннє виявлення аномалій, що може запобігти каскадним відмовам та серйозним інцидентам. Автоматизація операцій, зокрема через AIOps, дає змогу системам самостійно аналізувати інциденти, корелювати події та визначати першопричини проблем. Це підвищує надійність і доступність сервісів, адже прогностичні моделі забезпечують своєчасне масштабування та оптимізацію ресурсів [18, с. 104].

Проте, попри численні переваги, існують і значні обмеження, які потребують уваги. Зокрема, високі вимоги до великих наборів даних можуть стати перешкодою для ефективного навчання моделей. Це підкреслює необхідність якісних історичних даних про роботу хмарної системи. Також варто зазначити, що навчання складних моделей, таких як LSTM або Transformers, вимагає значних обчислювальних ресурсів, що може призводити до високих фінансових витрат [19, с. 503].

Важливим аспектом є потреба у кваліфікованих спеціалістах із МН та AIOps для інтеграції цих технологій. Без відповідних знань і досвіду реалізація може бути складною. Крім того, можливі хибнопозитивні спрацювання під час неправильного навчання або недостатньої кількості даних можуть призвести до негативних наслідків.

Складність інтеграції нових технологій із наявною хмарною інфраструктурою також може викликати труднощі, особливо якщо система нестандартизована. Моделі втрачають точність із часом через дрейф даних (drift), що вимагатиме регулярного перенавчання. Ризики кібербезпеки, пов'язані з маніпуляціями вхідними даними для обману моделей прогнозування, є ще одним серйозним викликом.

До того ж неоднорідність даних між мікросервісами може ускладнити створення універсальної моделі, що підкреслює важливість адаптації підходів до конкретних умов. Таким чином, хоча застосування ШІ для прогнозування відмов у хмарних інфраструктурах має значні переваги, необхідно враховувати потенційні недоліки та обмеження цієї технології.

Цей контекст відкриває нові можливості для впровадження інноваційних рішень, які можуть допомогти подолати ці виклики. Одним із таких є інтеграція цифрових двійників хмарної інфраструктури, що дає змогу створити керовану ШІ віртуальну копію середовища для тестування ризиків, сценаріїв відмов та оновлень. Такий підхід гарантує безпечне прогнозування проблем, не впливаючи на продуктивне середовище, і допомагає виявити вузькі місця системи.

Цифрові двійники можуть використовуватися для симуляції різних сценаріїв навантаження та тестування реакції системи на них. Це дає змогу не лише виявити потенційні точки відмови, але й оптимізувати конфігурацію

системи для підвищення її стійкості. Завдяки такій інтеграції, організації можуть швидше адаптуватися до змін у навантаженні та знизити ризики, пов'язані з оновленнями або змінами в налаштуваннях.

Крім того, необхідно розглянути практичні рекомендації для впровадження ШІ в моніторинг і управління ризиками хмарних систем. Наприклад, застосування аналітики даних у реальному часі може допомогти виявляти аномалії та реагувати на них ще до того, як вони спричинять серйозні проблеми. Поєднання цих технологій із традиційними методами моніторингу може значно підвищити загальну ефективність управління ризиками.

Наступним кроком є запровадження механізмів самовідновлення. Застосування моделей ШІ, які самостійно виявляють аномальну поведінку, дає змогу автоматично ініціювати відновлення, наприклад, міграцію віртуальних машин чи перезапуск сервісів. Це скорочує час простою та мінімізує необхідність ручного втручання.

Генеративні моделі, такі як GPT, також можуть бути корисними для аналізу логів. Вони генерують гіпотези щодо причин збоїв, описують сценарії ризику та пропонують варіанти їхнього усунення [20, с. 3]. Це прискорює процес аналізу кореневих причин інцидентів і забезпечує швидке реагування.

До того ж упровадження поведінкового профілю ресурсів дає змогу створити модель поведінки кожного сервісу на основі ШІ-аналітики, що сприяє виявленню навіть незначних відхилень від норми, які можуть передувати збоям. Використання графових нейронних мереж для моделювання каскадних відмов також є перспективним рішенням. Аналіз залежностей між сервісами у формі графа допомагає прогнозувати каскадні збої заздалегідь, що сприяє запобіганню масовим інцидентам і покращує планування резервування.

ШІ-орієнтоване тестування стійкості є ще одним важливим аспектом. Автоматичне генерування інтелектуальних порушень допомагає системам навчатися на реальних сценаріях, підвищуючи їхню стійкість і здатність до самовідновлення.

Створення системи колективного інтелекту сервісів через федеративне навчання дає змогу поєднувати локальні моделі без обміну сирими даними. Це підвищує точність прогнозування у великих розподілених хмарах та гарантує більшу безпеку даних.

Упровадження ШІ-прогнозування енергоефективності та навантаження є серйозним кроком у зменшенні ризику фізичних відмов обладнання. Це рішення передбачає поєднання прогнозу навантаження з оцінкою енергетичних ризиків, таких як перевантаження, перегрів та деградація обладнання. Завдяки такому інтегрованому підходу можна не лише оптимізувати витрати, а й стабілізувати роботу системи.

Реалізація превентивного масштабування на основі ШІ забезпечує можливість передбачати піки навантаження за кілька хвилин або годин до їхнього виникнення. Це дає змогу завчасно масштабувати ресурси, що значно знижує ймовірність падіння сервісів у критичні моменти.

Наступним етапом є автоматичне оцінювання ризику для кожного мікросервісу. Використовуючи історичні дані про інциденти, залежності та навантаження, модель може оцінювати рівень ризику збою в діапазоні від 0 до 100 %. Це забезпечує миттєву пріоритизацію компонентів системи, що є надзвичайно важливим для оперативного реагування на можливі загрози.

ШІ також може створювати теплові карти аномалій, які візуалізують проблемні зони в хмарній інфраструктурі. Це дає змогу швидко діагностувати проблемні компоненти та визначати зони ризику, що підвищує загальну ефективність моніторингу системи.

Важливою також є інтеграція моделей прогнозування в DevOps-цикли через MLOps-based CI/CD. Під час кожного релізу автоматично тестується вплив змін на ризик відмов, і моделі оновлюються [21, с. 456]. Це істотно зменшує кількість інцидентів після оновлень і забезпечує стабільні релізи.

Завершальним етапом є застосування ШІ для симуляції критичного навантаження. ШІ генерує навантаження, яке імітує реальні умови, що дає змогу перевірити стійкість системи до непередбачуваних піків і зовнішніх чинників. Таким чином, інтеграція цих технологій створює надійнішу та ефективнішу хмарну інфраструктуру, готову до викликів сучасності.

Отже, застосування технологій ШІ для прогнозування відмов у хмарних інфраструктурах відкриває нові можливості для підвищення надійності та продуктивності систем. Кожен із підходів має свої особливості, переваги й недоліки, тому вибір конкретного методу залежить від специфіки завдання та доступних даних.

Висновки

У результаті дослідження детально проаналізовано різні моделі та методи ШІ, які застосовуються для прогнозування відмов у хмарних інфраструктурах. Зокрема, розглянуто алгоритми, такі як Random Forest, XGBoost, LSTM та графові нейронні мережі, які продемонстрували високу ефективність в аналізі великих обсягів даних та виявленні аномалій. Це дає змогу організаціям підвищити точність прогнозування та знизити операційні витрати, що важливо в сучасному бізнес-середовищі.

Аналіз технічних збоїв, відхилень у продуктивності, кіберзагроз, проблем обслуговування, інтеграційних порушень та помилок, пов'язаних із людським фактором, засвідчив наявність різних типів відмов у хмарних інфраструктурах.Зокрема, технічні збої охоплюють апаратні й програмні порушення, що можуть спричиняти втрату даних або зупинку системи. Відхилення продуктивності відображають нетипові зміни в роботі інфраструктури

та сигналізують про потенційні порушення. Кіберзагрози стосуються атак на систему, таких як DDoS-атаки та фішинг. Проблеми обслуговування виникають у разі недостатньої технічної підтримки або труднощів із наданням послуг. Інтеграційні помилки виникають під час впровадження нових технологій у вже наявну інфраструктуру, а помилки людського походження виникають через неправильні дії чи недогляд персоналу.

Разом із перевагами впровадження технологій ШІ дослідження також окреслило низку викликів. До них належать необхідність у великих наборах даних, висока обчислювальна вартість та ризики кібербезпеки. Ці чинники можуть стати перешкодою для організацій, які прагнуть інтегрувати такі технології. Важливо враховувати ці обмеження під час планування впровадження ШІ у хмарні системи.

Наукова новизна дослідження полягає в систематизації та порівняльному аналізі наявних методів і моделей ШІ для прогнозування відмов у хмарних інфраструктурах та в розробленні практичних рекомендацій щодо їхньої інтеграції.

Перспективи подальших досліджень можуть охоплювати поглиблене вивчення впливу нових технологій, зокрема федеративного навчання та AIOps, на підвищення надійності хмарних систем. Важливим напрямом є також аналіз можливостей адаптації моделей до змін середовища та зміцнення їхньої стійкості до дрейфу даних (drift), що сприятиме підвищенню ефективності прогнозування відмов. Крім того, перспективним напрямом є вивчення питань інтеграції технологій ШІ з методами автоматизованого управління ресурсами, що відкриває можливість формування високодинамічних та адаптивних систем моніторингу й керування хмарними інфраструктурами.

Список використаної літератури

1. Voutsas F., Violos J., Leivadeas A. Mitigating alert fatigue in cloud monitoring systems: A machine learning perspective. *Computer Networks*. 2024. Vol. 250. Article 110543. DOI: https://doi.org/10.1016/j.comnet.2024.110543

2. Agboola O. A., Ogeawuchi J. C., Gbenle T. P., Abayomi A. A., Uzoka A. C. Advances in risk assessment and mitigation for complex cloud-based project environments. *Journal of Frontiers in Multidisciplinary Research*. 2023. Vol. 06, № 01. P. 309–320.DOI: https://doi.org/10.54660/.jfmr.2023.4.1.309-320

3. Drissi S., Chergui M., Khatar Z. A systematic literature review on risk assessment in cloud computing: Recent research advancements. *IEEE Access*. 2025. № 13. DOI: https://doi.org/10.1109/access.2025.3561123

4. Козак Ю. Б. Аналіз даних та машинне навчання на хмарних та туманних платформах як основа ефективної передачі даних. *Вчені записки ТНУ імені В. І. Вернадського. Серія: Технічні науки*. 2021. Т. 32 (71), № 5. С. 100–107. DOI: https://doi.org/10.32838/2663-5941/2021.5/16

5. Кушнерьов О. С., Позовна І. В., Сокол В. Вплив нейронних мереж на розвиток кібербезпеки в умовах регуляторних змін. *Безпека інформації*. 2024. Т. 30, № 2. С. 261–269.DOI: https://doi.org/10.18372/2225-5036.30.19238

6. Duda O., Shakleina I., Luchkevych M. Increasing the efficiency of DevOps through the use of artificial intelligence and machine learning. *Herald of Khmelnytskyi National University. Technical sciences*. 2025. Vol. 351, № 3(1). P. 143–149. DOI: https://doi.org/10.31891/2307-5732-2025-351-17

7. Adewusi A. O., Okoli U. I., Olorunsogo T., Adaga E., Daraojimba D. O., Obi O. C. Artificial intelligence in cybersecurity: Protecting national infrastructure: A USA review. *World Journal of Advanced Research and Reviews*. 2024. Vol. 21, № 1. P. 2263–2275.DOI: https://doi.org/10.30574/wjarr.2024.21.1.0313.

8. Burov Y., Zhovnir Y., Zakharya O. The vision and implementation of intelligent security system. *Herald of Khmelnytskyi National University. Technical sciences*. 2024. Т. 341, № 5. P. 497–509. DOI: https://doi.org/10.31891/2307-5732-2024-341-5-72

9. Sarker I. H., Furhad M. H., Nowrozy R. AI-driven cybersecurity: An overview, security intelligence modeling and research directions. *SN Computer Science*. 2021. Vol. 2. Article 173. DOI: https://doi.org/10.1007/s42979-021-00557-0

10. Davydov V., Hrebeniuk D. Development the resources load variation forecasting method within cloud computing systems. *Advanced Information Systems*. 2020. Vol. 4, № 4. P. 128–135. DOI: https://doi.org/10.20998/2522-9052.2020.4.18

11. Ahmed S. A., Khalifa E. H., Nawaz M., Abdalla F. A., Mahmoud A. F. A. Enhancing cloud data center security through deep learning: A comparative analysis of RNN, CNN, and LSTM models for anomaly and intrusion detection. *Engineering, Technology & Applied Science Research*. 2025. Vol. 15, № 1. P. 20071–20076. DOI: https://doi.org/10.48084/etasr.9445

12. Tengku Asmawi T. N., Ismail A., Shen J. Cloud failure prediction based on traditional machine learning and deep learning. *Journal of cloud computing*. 2022. Vol. 11. Article 47. DOI: https://doi.org/10.1186/s13677-022-00327-0

13. Noor A. Cloud-based deep learning for real-time URL anomaly detection: LSTM/GRU and CNN/LSTM models. *Computer systems science and engineering*. 2025. № 49. P. 259–286. DOI: https://doi.org/10.32604/csse.2025.060387

14. Saha S., Sarkar J., Dhavala S., Mota P., Sarkar S. Quantile-long short term memory: A robust, time series anomaly detection method. *IEEE transactions on artificial intelligence*. 2024. Vol. 5, № 8. P. 3939–3950. DOI: https://doi.org/10.1109/tai.2024.3353163

15. Zhao Z., Xu C., Li B. A LSTM-based anomaly detection model for log analysis. *Journal of signal processing systems*. 2021. Vol. 93. P. 745–751. DOI: https://doi.org/10.1007/s11265-021-01644-4

16. Shaikh R., Muntean C. H., Gupta S. Prediction of resource utilisation in cloud computing using machine learning. *Proceedings of the 14th International Conference on Cloud Computing and Services Science CLOSER*. 2024. Vol. 1. P. 103–114. DOI: https://doi.org/10.5220/0012742200003711

17. Al-Ghuwairi A. R., Sharrab Y., Al-Fraihat D., AlElaimat M., Alsarhan A., Algarni A. Intrusion detection in cloud computing based on time series anomalies utilizing machine learning. *Journal of cloud computing*. 2023. Vol. 12. Article 127. DOI: https://doi.org/10.1186/s13677-023-00491-x

18. Dmytriv Y., Orlov M. Use of artificial intelligence methods and tools in the construction of cloud IT infrastructures. *Вісник Національного університету «Львівська політехніка». Серія Інформаційні системи та мережі*. 2025. Вип. 17. С. 101–113. DOI: https://doi.org/10.23939/sisn2025.17.101

19. Іванченко Ю., Аверичев І., Рижаков М. Узагальнена модель прогнозування та виявлення аномалій кібербезпеки на основі штучного інтелекту. *Кібербезпека: освіта, наука, техніка*. 2025. № 4(28). С. 493–510. DOI: https://doi.org/10.28925/2663-4023.2025.28.823

20. Трапаідзе С., Швецова К. Генеративний штучний інтелект у створенні маркетингового контенту для українських компаній. *Економіка та суспільство*. 2025. № 72. DOI: https://doi.org/10.32782/2524-0072/2025-72-161

21. Pravorska N. Method of applying machine learning to enhance the efficiency of DevOps processes. *Herald of Khmelnytskyi National University, Technical sciences*. 2024. Vol. 343, № 6(1). P. 454–463. DOI: https://doi.org/10.31891/2307-5732-2024-343-6-68

References

1. Voutsas, F., Violos, J., & Leivadeas, A. (2024). Mitigating alert fatigue in cloud monitoring systems: A machine learning perspective. *Computer Networks*, 250, Article 110543. https://doi.org/10.1016/j.comnet.2024.110543

2. Agboola, O. A., Ogeawuchi, J. C., Gbenle, T. P., Abayomi, A. A., & Uzoka, A. C. (2023). Advances in risk assessment and mitigation for complex cloud-based project environments. *Journal of Frontiers in Multidisciplinary Research*, 06(01), 309–320. https://doi.org/10.54660/.jfmr.2023.4.1.309-320

3. Drissi, S., Chergui, M., & Khatar, Z. (2025). A systematic literature review on risk assessment in cloud computing: Recent research advancements. *IEEE Access*, 13. https://doi.org/10.1109/access.2025.3561123

4. Kozak, Yu. B. (2021). Analiz danykh ta mashynne navchannia na khmarnykh ta tumannykh platformakh yak osnova efektyvnoi peredachi danykh [Data analysis and machine learning on cloud and fog platforms as a basis for efficient data transmission]. *Vcheni zapysky TNU imeni V.I. Vernadskoho. Seriia: Tekhnichni nauky* [Scientific notes of TNU named after V.I. Vernadsky. Series: Technical Sciences], 32(71), no. 5, 100–107. https://doi.org/10.32838/2663-5941/2021.5/16 (in Ukrainian).

5. Kushnerov, O. S., Pozovna, I. V., & Sokol, V. (2024). Vplyv neironnykh merezh na rozvytok kiberbezpeky v umovakh rehuliatornykh zmin [Impact of neural networks on cybersecurity development in the context of regulatory changes]. *Bezpeka informatsii* [Information security], 30(2), 261–269. https://doi.org/10.18372/2225-5036.30.19238 (in Ukrainian).

6. Duda, O., Shakleina, I., & Luchkevych, M. (2025). Increasing the efficiency of DevOps through the use of artificial intelligence and machine learning. *Herald of Khmelnytskyi National University. Technical sciences*, 351, no. 3(1), 143–149. https://doi.org/10.31891/2307-5732-2025-351-17

7. Adewusi, A. O., Okoli, U. I., Olorunsogo, T., Adaga, E., Daraojimba, D. O., & Obi, O. C. (2024). Artificial intelligence in cybersecurity: Protecting national infrastructure: A USA review. *World Journal of Advanced Research and Reviews*, 21(1), 2263–2275. https://doi.org/10.30574/wjarr.2024.21.1.0313

8. Burov, Y., Zhovnir, Y., & Zakharya, O. (2024). The vision and implementation of intelligent security system. *Herald of Khmelnytskyi National University. Technical sciences*, 341, no. 5, 497–509. https://doi.org/10.31891/2307-5732-2024-341-5-72

9. Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). AI-driven cybersecurity: An overview, security intelligence modeling and research directions. *SN Computer Science*, 2, Article 173. https://doi.org/10.1007/s42979-021-00557-0

10. Davydov, V., & Hrebeniuk, D. (2020). Development the resources load variation forecasting method within cloud computing systems. *Advanced Information Systems*, 4(4), 128–135. https://doi.org/10.20998/2522-9052.2020.4.18

11. Ahmed, S. A., Khalifa, E. H., Nawaz, M., Abdalla, F. A., & Mahmoud, A. F. A. (2025). Enhancing cloud data center security through deep learning: A comparative analysis of RNN, CNN, and LSTM models for anomaly and intrusion detection. *Engineering, Technology & Applied Science Research*, 15(1), 20071–20076. https://doi.org/10.48084/etasr.9445

12. Tengku Asmawi, T. N., Ismail, A., & Shen, J. (2022). Cloud failure prediction based on traditional machine learning and deep learning. *Journal of Cloud Computing*, 11, Article 47. https://doi.org/10.1186/s13677-022-00327-0

13. Noor, A. (2025). Cloud-based deep learning for real-time URL anomaly detection: LSTM/GRU and CNN/LSTM models. *Computer Systems Science and Engineering*, 49, 259–286. https://doi.org/10.32604/csse.2025.060387

14. Saha, S., Sarkar, J., Dhavala, S., Mota, P., & Sarkar, S. (2024). Quantile-long short term memory: A robust, time series anomaly detection method. *IEEE Transactions on Artificial Intelligence*, 5(8), 3939–3950. https://doi.org/10.1109/tai.2024.3353163

15. Zhao, Z., Xu, C., & Li, B. (2021). A LSTM-based anomaly detection model for log analysis. *Journal of Signal Processing Systems*, 93, 745–751. https://doi.org/10.1007/s11265-021-01644-4

16. Shaikh, R., Muntean, C. H., & Gupta, S. (2024). Prediction of resource utilisation in cloud computing using machine learning. *Proceedings of the 14th International Conference on Cloud Computing and Services Science CLOSER*, 1, 103–114. https://doi.org/10.5220/0012742200003711

17. Al-Ghuwairi, A. R., Sharrab, Y., Al-Fraihat, D., AlElaimat, M., Alsarhan, A., & Algarni, A. (2023). Intrusion detection in cloud computing based on time series anomalies utilizing machine learning. *Journal of Cloud Computing*, 12, Article 127. https://doi.org/10.1186/s13677-023-00491-x

18. Dmytriv, Y., & Orlov, M. (2025). Use of artificial intelligence methods and tools in the construction of cloud IT infrastructures. *Visnyk Natsionalnoho universytetu "Lvivska politekhnika". Seriia Informatsiini systemy ta merezhi* [Bulletin of Lviv Polytechnic National University. Series Information Systems and Networks], 17, 101–113. https://doi.org/10.23939/sisn2025.17.101

19. Ivanchenko, Yu., Averychev, I., & Ryzhakov, M. (2025). Uzahalnena model prohnozuvannia ta vyiavlennia anomalii kiberbezpeky na osnovi shtuchnoho intelektu [Generalized model for forecasting and detecting cybersecurity anomalies based on artificial intelligence]. *Kiberbezpeka: osvita, nauka, tekhnika* [Cybersecurity: education, science, technology], no. 4(28), 493–510. https://doi.org/10.28925/2663-4023.2025.28.823 (in Ukrainian).

20. Trapaidze, S., & Shvetsova, K. (2025). Heneratyvnyi shtuchnyi intelekt u stvorenni marketynhovoho kontentu dlia ukrainskykh kompanii [Generative artificial intelligence in creating marketing content for Ukrainian companies]. *Ekonomika ta suspilstvo* [Economy and society], no. 72. https://doi.org/10.32782/2524-0072/2025-72-161 (in Ukrainian).

21. Pravorska, N. (2024). Method of applying machine learning to enhance the efficiency of DevOps processes. *Herald of Khmelnytskyi National University, Technical sciences*, 343, no. 6(1), 454–463. https://doi.org/10.31891/2307-5732-2024-343-6-68

Дата першого надходження статті до видання: 17.01.2026
Дата прийняття статті до друку після рецензування: 20.02.2026
Дата публікації (оприлюднення) статті: 30.04.2026