УДК 519.7: 004.8 **doi: 10.26906/SUNZ.2025.2.132-136**

І. В. Ільїна, Р. В. Артюх, М. М. Зимогляд

Харківський національний університет радіоелектроніки, Харків, Україна

АНАЛІЗ ДАНИХ ТА МАШИННЕ НАВЧАННЯ У ХМАРНИХ ТА ТУМАННИХ ПЛАТФОРМАХ ДЛЯ ЕФЕКТИВНОЇ ПЕРЕДАЧІ ДАНИХ

Анотація. У даній статті досліджується ефективність використання традиційних методів оптимізації передачі даних та запропонованої моделі на основі алгоритмів машинного навчання. В умовах стрімкого і постійного зростання обсягу інформації, розширення мережевої інфраструктури та зростаючих вимог до продуктивності систем з'являється необхідність адаптивного підходу до управління трафіком і зниженню енерговитрат. Особлива увага приділена інтеграції технологій машинного навчання з метою підвищення ефективності передачі даних у хмарних та туманних обчисленнях. Запропонована модель поєднує у собі методи прогнозування мережевого навантаження (LSTM, XGBoost), технології стиснення даних (VAE, K-Means) та програмно-визначені мережі (SDN) для балансування трафіку. У межах дослідження було проведено тестування комбінацій алгоритмів у розподілених середовищах із різними рівнями навантаження, що дало можливість оцінити ефективність кожного за критеріями продуктивності, надійності та енергоспоживання. Даний підхід забезпечує адаптивність до динамічних змін у трафіку і сприяє зменшенню затримок у передачі даних. Отримані результати досліджень підтверджують доцільність машинного навчання у застосуванні до оптимізації мережевої взаємодії і відкривають перспективи для подальших досліджень у сфері інтелектуального управління туманними обчисленнями. Запропоноване рішення може використовуватися у різних потребах, включаючи автономні транспортні системи, медичні інформаційні платформи, промисловий Інтернет речей та критично важливі інфраструктури.

Ключові слова: машинне навчання, туманні обчислення, хмарні обчислення, LSTM, XGBoost, VAE, K-Means, SDN, оптимізація навантаження, стиснення даних, затримка передачі, енергоефективність.

Вступ

Постановка проблеми. Сучасний етап розвитку інформаційних технологій характеризується стрімким зростанням обсягів даних, що генеруються різноманітними пристроями, системами та користувачами. Інтернет речей (IoT), розподілені обчислювальні платформи, автономні транспортні системи, медичні IoT-рішення та інші інноваційні технології вимагають миттєвої обробки та передачі інформації в умовах обмежених ресурсів. Це ставить нові виклики перед інфраструктурою хмарних і туманних обчислень, які покликані забезпечити ефективне функціонування розподілених систем. Однак традиційні підходи до управління мережевим трафіком, балансування навантаження та оптимізації енергоспоживання все частіше виявляються недостатніми для задоволення вимог сучасних додатків, де навіть мінімальні затримки або втрати даних можуть призвести до критичних наслідків.

Особливу актуальність набуває інтеграція методів машинного навчання (МН) та аналізу даних у хмарні й туманні середовища. Туманні обчислення, як проміжна ланка між периферійними пристроями та хмарою, дозволяють зменшити затримки та підвищити локалізовану обробку даних. Проте їхній потенціал залишається недостатньо розкритим у контексті динамічної адаптації до змін навантаження, оптимізації трафіку та енергоефективності. Існуючі рішення часто ґрунтуються на статичних алгоритмах, які не враховують нелінійні залежності та часові характеристики мережевих процесів. Це обмежує їхню здатність до масштабування в умовах експоненційного зростання кількості підключених пристроїв і різноманітності типів трафіку.

Метою даного дослідження є розробка адаптивної модели оптимізації передачі даних у хмарних і туманних платформах, яка поєднує передові методи машинного навчання, стиснення даних та програмно-визначених мереж (SDN).

Основна увага приділяється створенню гібридного підходу, що інтегрує прогнозування навантаження за допомогою LSTM (Long Short-Term Memory) та XGBoost (Extreme Gradient Boosting), стиснення даних на основі варіаційних автокодувальників (VAE) і кластеризації K-Means, а також динамічного балансування трафіку через SDN. Таке поєднання дозволяє не лише зменшити затримки та втрати пакетів, але й оптимізувати використання пропускної здатності мережі та енергоресурсів.

Важливим аспектом роботи є також аналіз взаємодії між різними компонентами моделі. Наприклад, комбінація LSTM і XGBoost дозволяє ефективно працювати як з часовими рядами, так і зі структурованими даними, тоді як інтеграція VAE та K-Means забезпечує гнучке стиснення без втрати критичної інформації. Використання SDN додає шаг динамічного управління, що дозволяє системі адаптуватися до змін у реальному часі. Актуальність роботи підкріплюється результатами експериментів, проведених у середовищах Google Cloud Platform (GCP), AWS Lambda та OpenStack.

Мета статті – представлення результатів дослідження, спрямованого на розробку та експериментального тестування адаптивної моделі оптимізації передачі даних у туманних обчисленнях, яка базується на методах машинного навчання, зокрема XGBoost, LSTM та Autoencoder.

Аналіз останніх досліджень та публікацій. Сучасні дослідження в галузі хмарних і туманних обчислень демонструють критичну роль інтеграції машинного навчання для оптимізації передачі даних.

Так, Bonomi та співавтори у статті [1] ввели концепцію туманних обчислень як проміжного шару між

хмарою та пристроями IoT. Це дозволяє суттєво зменшити затримки в розподілених системах. Ця робота стала фундаментом для подальших інновацій у сфері обробки даних у реальному часі.

Дослідження Shi та ін. у роботі [2] розширили цю концепцію, довівши ефективність туманних обчислень у поєднанні з edge computing для IoT-додатків.

У контексті машинного навчання ключову роль відіграють алгоритми прогнозування мережевого навантаження.

Наприклад, у статті [3] акцентує увагу на ефективності методів на кшталт XGBoost і LSTM, тоді як Chen та Guestrin у статті [4] емпірично довели переваги XGBoost у задачах класифікації та регресії завдяки унікальній архітектурі дерев рішень.

У роботі Hochreiter та Schmidhuber [5], які запропонували модель LSTM для аналізу часових рядів у розподілених системах, продемонстровано високу точність у прогнозуванні динамічних навантажень.

Важливим аспектом оптимізації трафіку є стиснення даних. Kingma та Welling у статті [6] запропонували теоретичні основи варіаційних автокодувальників (VAE), що відкрило можливості їх застосування в реальних мережах.

Для підвищення ефективності стиснення використовувався метод K-Means, ефективність підтверджується дослідженнями Jain [7] у контексті кластеризації даних.

Програмно-визначені мережі (SDN) були і залишаються ключовим інструментом для динамічного балансування навантаження. За дослідженням Feamster та ін. [8], архітектура SDN дозволяє нам адаптивно керувати трафіком шляхом централізованого контролю, що критично для систем із високими вимогами до часу відгуку.

Енергоефективність при впровадженні методів машинного навчання підтверджує робота [9], де інтелектуальний розподіл навантаження сприяє зменшенню енергоспоживання у хмарних середовищах.

Виклад основного матеріалу

Туманні та хмарні обчислення забезпечують ефективну передачу даних у розподілених середовищах, але вони створюють проблеми, що пов’язані із затримкою, балансуванням навантаження та оптимізацією пропускної здатності мережі. Наприклад, затримка передачі даних у медичних системах IoT може мати критичні наслідки для пацієнтів, а в автономному транспортуванні навіть невеликі затримки можуть спричинити надзвичайні ситуації.

Оскільки дедалі більше пристроїв генерують великі обсяги даних, потрібні передові методи керування мережею. Використовуючи машинне навчання, можливо передбачити навантаження на мережу та запобігати перевантаженням. Затримку передачі можливо зменшити за допомогою інтелектуального балансування трафіку, а використовування пропускної здатності можливо оптимізувати за допомогою стиснення даних і кластеризації.

В роботі запропонована модель на основі машинного навчання, що включає в себе три основні складові (рис. 1).

Рис. 1. Запропонована модель

Використовуючи синтетичні та реальні набори даних, експериментальні дослідження проводилися в середовищах Google Cloud Platform (GCP), AWS Lambda та OpenStack.

Прогнозування. Прогнозування навантаження виконувалось за допомогою LSTM і XGBoost для аналізу мережевого трафіку (рис. 2, 3).

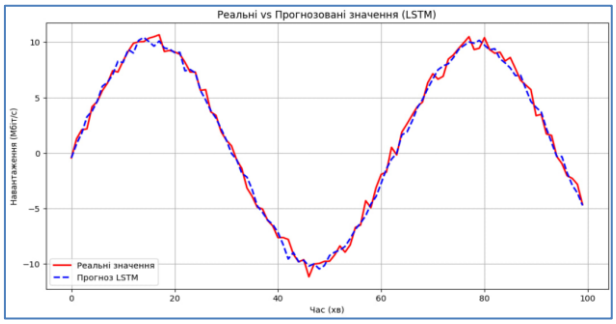

Рис. 2. Графік реальних та прогнозованих значень для LSTM

Рис. 3. Графік реальних та прогнозованих значень для XGBoost

ARIMA була обрана як базова модель для порівняння через її широке використання в аналізі часових рядів.

Модель ARIMA (AutoRegressive Integrated Moving Average) є класичним статистичним методом для аналізу часових рядів і широко використовується в задачах прогнозування навантаження. Її було включено до порівняння, що дозволило оцінити переваги сучасних методів машинного навчання (LSTM, XGBoost) над традиційними підходами. Наприклад, ARIMA не враховує нелінійні залежності та потребує ручного налаштування параметрів, що обмежує її ефективність у динамічних середовищах.

LSTM була обрана через здатність аналізувати довгострокові залежності в часових рядах, що є критичним для прогнозування мережевого трафіку.

XGBoost було використано через його швидкість, стійкість до шуму та можливість роботи зі структурованими даними.

Обидва методи дозволяють автоматично адаптуватися до змін у навантаженні, на відміну від ARIMA. Використання LSTM, XGBoost пояснюється їх високою ефективністю в задачах аналізу часових рядів і прогнозування динамічних змін мережі. Під час реалізації моделі використовувалися оптимізовані гіперпараметри:

– для LSTM – двошаровий підхід із розміром комірки 50 та використанням функції активації ReLU;

– для XGBoost – кількість дерев рішень 100, глибина 5, швидкість навчання 0,1.

Основні результати підтвердили ефективність інтеграції машинного навчання точності прогнозування навантаження мережі для підвищення продуктивності туманних обчислень (рис. 4).

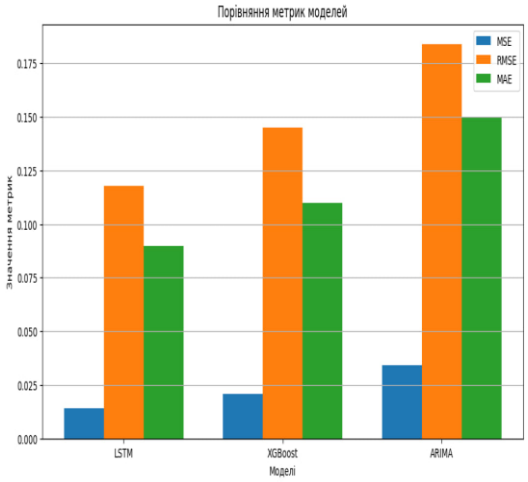

Рис. 4. Порівняння метрик (MSE, RMSE, MAE) між методами при прогнозуванні навантаження

Метрики, які були використані для оцінки методів прогнозування:

1. Середньоквадратична похибка (MSE) [10] для прогнозування навантаження розрахована:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 , \quad (1)$$

2. Корінь середньоквадратичної похибки (RMSE) розраховувався за формулою:

$$RMSE = \sqrt{MSE} , \quad (2)$$

3. Середня абсолютна похибка (MAE) розраховувалась таким чином:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}| y_i - \hat{y}_i | , \quad (3)$$

де y_i - реальне значення, $\hat{y}_i$ – прогнозоване, N- кількість спостережень.

Дослідження показали, що в порівнянні зі ARIMA методи LSTM та XGBoost мають менші значення метрик, що доводить їх здатність точніше прогнозувати навантаження мережі. Наприклад, MSE для LSTM (0.014) значно нижчий, ніж у ARIMA (0.034), що підтверджує вищу точність сучасних методів. А LSTM показала найвищу точність завдяки здатності аналізувати довгострокові залежності (табл. 1).

Таблиця 1 – **Результати порівняльного аналізу методів за метриками MSE, RMSE, MAE**

Метод	MSE	RMSE	MAE
LSTM	0.014	0.118	0.090
XGBoost	0.021	0.145	0.110
ARIMA	0.034	0.184	0.150

Стиснення. Стиснення даних виконувалось за допомогою варіаційного автокодувальника (VAE), що зменшує обсяг переданих даних. Архітектура VAE включала прихований рівень із 64 елементами та простором затримки розміром 32.

Застосування методу K-Means для стиснення дозволяє скоротити надмірний обсяг даних шляхом залишення кілька найбільш характерних представників від кожного кластера.

Поєднанням стиснення на основі VAE та K-Means, вдалося зменшити переданий трафік значно знизивши навантаження на мережеву інфраструктуру, що свідчить про синергію між кластеризацією та глибоким навчанням (табл. 2).

Оптимізація пропускної здатності за рахунок зменшення розміру даних розрахована:

$$R = \left(\frac{Original\ Size - Compressed\ Size}{Original\ Size} \right) \times 100\% , \quad (4)$$

де *Original Size* - обсяг даних до стиснення; *Compressed Size* - обсяг даних після стиснення.

Таблиця 2 – **Результати порівняння обсягу даних, переданих у різних сценаріях**

Метод стиснення	Обсяг переданих даних (%)	Скорочення трафіку (%)
Без стиснення	100	0
K-Means	72	28
VAE	55	45
K-Means+VAE	43	57

Балансування. Балансування навантаження виконувалось за допомогою програмно-визначеної мережі (SDN), яка дозволяє динамічно перерозподіляти потоки даних залежно від поточного навантаження на вузли мережі. Балансування маршрутизації в SDN базується на алгоритмі Shortest Path First з додатковим зважуванням для затримок у вузлі.

Було виявлено, що балансування навантаження через SDN зможе оптимізувати пересилання запитів

у реальному часі, і це вирішальне значення для роботи інтелектуальних систем IoT.

Результати тестування підтвердили, що застосування запропонованих підходів дозволяє ефективно керувати навантаженням у туманних і хмарних обчисленнях, зменшує споживання ресурсів і покращує загальну продуктивність системи.

У порівнянні з іншими підходами, такими як традиційні методи балансування за фіксованими правилами, дослідження показало, що запропонований підхід може покращувати розподіл трафіку та зменшувати середню затримку.

Середня затримка в мережі розраховувалась за параметрами [11]:

$$Latensy\,avg = \frac{1}{N}\sum_{i=1}^{N}(T_{Response} - T_{Requst}), \qquad (5)$$

де $T_{Response}$ - час отримання відповіді; T_{Requst} - час відправлення запиту; N - кількість спостережень.

Для комплексної оцінки ефективності запропонованої моделі було проведено порівняльний аналіз зі звичайними методами оптимізації. Традиційні методи включають ARIMA для прогнозування, статичне балансування та базове стиснення даних. Енергоспоживання розраховано на основі експериментів у середовищі OpenStack з використанням стандартних інструментів моніторингу.

Результати порівняння представлені в табл. 3.

Таблиця 3 – **Результати порівняльного аналізу звичайних методів оптимізації трафіку за допомогою машинного навчання**

Метрика	Традиційні методи	Запропонована модель
Точність прогнозування навантаження (%)	85,2	95,6
Скорочення обсягу переданих даних (%)	22,1	57
Зменшення середньої затримки (%)	18,4	48,6
Енергоспоживання (%)	100	60

Точність прогнозування навантаження досягла 95,6% за допомогою методу LSTM.

Запропонована модель включає в себе кластеризацію даних за допомогою K-Means і використання варіаційного автокодувальника (VAE) для стиснення трафіку. Ефективність цих підходів оцінювалась шляхом порівняння обсягу даних, переданих у різних сценаріях. За результатами проведених експериментів, комбінований підхід K-Means і VAE зменшив обсяг переданих даних на 57%, значно підвищивши ефективність використання пропускної здатності мережі.

Виявлено, що середня затримка була зменшена на 48,6% за допомогою адаптивного балансування.

Отримані результати підтверджують, і дають зрозуміти, що інтеграція методів машинного навчання може значно підвищити продуктивність туманних і хмарних платформ. Зокрема, основними перевагами запропонованого підходу є скорочення затримки на 48,6% і зниження споживання енергії на 40%.

Це відкриває нові можливості для впровадження в системах реального часу, таких як медичні моніторингові платформи, автономний транспорт та промислові IoT-мережі, де ефективність передачі даних безпосередньо впливає на безпеку та продуктивність.

Висновки

Сучасні технології, такі як туманні та хмарні обчислення, відкривають нові можливості для ефективної передачі даних, але водночас ставлять складні виклики, пов’язані з затримками, енергоефективністю та обробкою великих обсягів інформації.

Це дослідження показало, що поєднання машинного навчання з інноваційними методами оптимізації може суттєво покращити роботу таких систем. Наприклад, модель LSTM, яка аналізує часові ряди, дозволила передбачати навантаження з рекордною точністю, а XGBoost прискорила обробку даних, навіть коли частина інформації була відсутня.

Окремо варто відзначити ефективність стиснення даних: завдяки поєднанню кластеризації K-Means та глибокого навчання (VAE) вдалося значно скоротити обсяг переданої інформації. Це не лише зменшило навантаження на мережу, але й заощадило енергію, що критично важливо для пристроїв IoT з обмеженими ресурсами. Крім того, програмно-визначені мережі (SDN) допомогли динамічно розподіляти трафік, знизивши середню затримку майже на половину. Такі результати особливо актуальні для медичних систем, де кожна мілісекунда може мати вирішальне значення, або для автономних автомобілів, де стабільність передачі даних безпеко-критична.

Важливим аспектом роботи стала також енергоефективність. Запропоновані рішення дозволили знизити споживання енергії на 40%, що робить їх привабливими не лише з технічної, а й з екологічної точки зору.

Дослідження показало, що даний підхід не лише теоретично обґрунтований, але може бути впроваджений у реальні системи передачі даних, де швидкість, точність та енергоощадність стають запорукою технологічного прориву.

У статті підкреслюється важливість поєднання нових рішень із існуючими, щоб зменшити навантаження на канали зв’язку та адаптувати їх до зростаючих вимог.

Запропонована модель адаптивної оптимізації для передачі даних покращує точність прогнозування навантаження, зменшує затримку передачі даних, оптимізує пропускну здатність та зменшує енергоспоживання у поєднанні традиційних методів з використанням інтелектуального розподілу навантаження між обчислювальними ресурсами. Також, можна зробити висновок, що цей підхід може бути застосований у розподілених системах у режимі реального часу, що дозволить покращити ефективність їх використання.

Подальші дослідження можуть включати впровадження квантових алгоритмів для прискорення обчислень, гібридних архітектур нейронних мереж та

розширення методів стиснення для різних типів даних. Умови, створені стрімким розвитком технологій, вимагають постійного вдосконалення таких систем, і проведені дослідження це підтверджують.

СПИСОК ЛІТЕРАТУРИ

1. Bonomi, F. Fog computing and its role in the internet of things [Text] / F. Bonomi, R. Milito, J. Zhu, S. Addepalli // Proceedings of the MCC Workshop on Mobile Cloud Computing. – 2012. – P. 13–16. DOI:10.1145/2342509.2342513.
2. Shi, W. Edge computing: Vision and challenges [Text] / W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu // IEEE Internet of Things Journal. – 2016. – Vol. 3, No. 5. – P. 637–646. DOI:10.1109/JIOT.2016.2579198.
3. Stock-Price Forecasting Based on XGBoost and LSTM / P. Hoang Vuong et al. Computer Systems Science and Engineering. 2022. Vol. 40, no. 1. P. 237–246. DOI.org/10.32604/csse.2022.017685
4. Chen, T. XGBoost: A scalable tree boosting system [Text] / T. Chen, C. Guestrin // Proceedings of the 22nd ACM SIGKDD International Conference. – 2016. – P. 785–794. DOI:10.1145/2939672.2939785.
5. Hochreiter, S. Long short-term memory [Text] / S. Hochreiter, J. Schmidhuber // Neural Computation. – 1997. – Vol. 9, No. 8. – P. 1735–1780. DOI:10.1162/neco.1997.9.8.1735.
6. Kingma, D. P. Auto-encoding variational Bayes [Text] / D. P. Kingma, M. Welling // arXiv preprint. – 2013. – arXiv:1312.6114.
7. Jain, A. K. Data clustering: 50 years beyond K-means [Text] / A. K. Jain // Pattern Recognition Letters. – 2010. – Vol. 31, No. 8. – P. 651–666. DOI:10.1016/j.patrec.2009.09.011.
8. Feamster, N. The Road to SDN: An Intellectual History of Programmable Networks [Text] / N. Feamster, J. Rexford, E. Zegura // ACM SIGCOMM Computer Communication Review. – 2014. – Vol. 44, No. 2. – P. 87–98. DOI:10.1145/2602204.2602219.
9. Volk M.O. Models optimization of resources in cloud computing: a hybrid approach to automation of operations and energy saving / M. O. Volk et al. Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences. 2024. Vol. 1, no. 5. P. 91–96. URL: https://doi.org/10.32782/2663-5941/2024.5.1/15.
10. Belas A. O., Bidyuk P. I. CHOOSING A QUALITY CRITERION FOR EVALUATING THE FORECAST OF NONLINEAR NON-STATIONARY PROCESSES. KPI Science News. 2021. No. 2. URL: https://doi.org/10.20535/kpisn.2021.2.236936
11. Kreutz, D. Software-Defined Networking: A Comprehensive Survey [Text] / D. Kreutz, F. M. V. Ramos, P. E. Veríssimo // Proceedings of the IEEE. – 2015. – Vol. 103, No. 1. – P. 14–76. DOI:10.1109/JPROC.2014.2371999.

Received (Надійшла) 25.03.2025
Accepted for publication (Прийнята до друку) 14.05.2025

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Ільїна Ірина Віталіївна – кандидат технічних наук, доцент, доцент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Iryna Ilina – candidate of technical sciences, Associate Professor, Associate Professor of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: iryna.ilina@nure.ua; ORCID Author ID: https://orcid.org/0000-0003-3132-7949;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57202223262.

Артюх Роман Володимирович – кандидат технічних наук, доцент кафедри комп'ютерно-інтегрованих технологій, автоматизації та робототехніки, Харківський національний університет радіоелектроніки, Харків, Україна;
Roman Artiukh – candidate of technical sciences, Associate Professor of the Department of Computer-Integrated Technologies, Automation and Robotics, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: roman.artyuh77@gmail.com; ORCID Author ID: http://orcid.org/0000-0002-5129-2221;
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57208344824.

Зимогляд Микола Миколайович– студент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;
Mykola Zymohliad – student, Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;
e-mail: nickzymoglyad@gmail.com; ORCID Author ID: https://orcid.org/0009-0001-7136-1968;

Data analysis and machine learning in cloud and fog platforms for efficient data transmission

Iryna Ilina, Roman Artiukh, Mykola Zymohliad

Abstract. This article investigates the effectiveness of traditional methods for optimizing data transmission and the proposed model based on machine learning algorithms. In the context of rapid and constant growth in the volume of information, expansion of network infrastructure and increasing requirements for system performance, there is a need for an adaptive approach to traffic management and reducing energy consumption. Special attention is paid to the integration of machine learning technologies in order to increase the efficiency of data transmission in cloud and fog computing. The proposed model combines network load prediction methods (LSTM, XGBoost), data compression technologies (VAE, K-Means) and software-defined networks (SDN) for traffic balancing. As part of the study, combinations of algorithms were tested in distributed environments with different load levels, which made it possible to evaluate the effectiveness of each according to the criteria of performance, reliability and energy consumption. This approach provides adaptability to dynamic changes in traffic and helps reduce data transmission delays. The obtained research results confirm the feasibility of machine learning in the application to the optimization of network interaction and open up prospects for further research in the field of intelligent control of fog computing. The proposed solution can be used for various needs, including autonomous transportation systems, medical information platforms, industrial Internet of Things and critical infrastructures.

Keywords: machine learning, fog computing, cloud computing, LSTM, Extreme Gradient Boosting) Variational Autoencoder, K-Means, SDN, load optimization, data compression, transmission latency, energy efficiency.