УДК 004.8:519.23
JEL: A12; C14; C88
DOI: https://doi.org/10.32983/2222-4459-2025-9-106-113

ВЕЛИКІ МОВНІ МОДЕЛІ: НОВА ПАРАДИГМА ДЛЯ АНАЛІЗУ ДАНИХ

©2025 ТУМАНОВ О. О.

УДК 004.8:519.23
JEL: A12; C14; C88

Туманов О. О. Великі мовні моделі: нова парадигма для аналізу даних

Поява великих мовних моделей (ВММ) докорінно змінила парадигму аналізу даних, вплинувши на такі галузі, як аналітика соціальних медіа, охорона здоров'я та розробка програмного забезпечення. Ці моделі, прикладами яких є архітектури GPT і BERT, чудово справляються із завданнями, пов'язаними з неструктурованим текстом і контекстним розумінням, але їхнє широке впровадження вимагає критичної оцінки їхнього зв'язку та відмінностей від традиційних статистичних підходів. Ця стаття має на меті провести порівняльний аналіз між ВММ і традиційними статистичними методами, досліджуючи їхні сильні сторони, обмеження та сфери застосування. У дослідженні використано якісний метод порівняння за такими критеріями: тип даних (неструктуровані проти структурованих), інтерпретованість, призначення, вимоги до даних і відтворюваність. Аналіз показує, що, попри ефективність ВММ в обробці текстів природною мовою, їхня ймовірнісна природа зумовлює суттєві обмеження щодо надійності та етичності. Традиційні статистичні методи, натомість, забезпечують високу інтерпретованість і відтворюваність, що є критично важливим для отримання детермінованих висновків. Особлива увага приділена обмеженням обох підходів. Зокрема, ВММ схильні до «галюцинацій» та успадкування упереджень із навчальних даних, що може призвести до дискримінаційних результатів. Водночас традиційні статистичні моделі не застраховані від упереджень відбору. В результаті дослідження було виявлено, що ВММ і традиційні методи є не конкурентами, а взаємодоповнювальними інструментами, кожен з яких найкраще підходить для вирішення певних завдань. Головним висновком є те, що найбільш перспективним напрямком є інтеграція цих підходів у гібридні методології. ВММ можуть ефективно застосовуватися для попередньої обробки та перетворення неструктурованого тексту в структуровані кількісні дані, які потім можуть бути піддані класичному статистичному аналізу. Такий підхід дозволить подолати наявні обмеження та досягти більшої надійності, прозорості та точності в наукових дослідженнях.
Ключові слова: *великі мовні моделі, статистичні методи, аналіз даних, штучний інтелект, етичні міркування, гібридна модель.*
Рис.: *5.* ***Табл.:*** *2.* ***Бібл.:*** *12.*

Туманов Олексій Олександрович *– доктор філософії, старший викладач Харківського національного університету імені В. Н. Каразіна (майдан Свободи, 4, Харків, 61022, Україна)*
E-mail: *oleksii.tumanov@gmail.com*
ORCID: *https://orcid.org/0000-0003-0674-0037*

UDC 004.8:519.23
JEL: A12; C14; C88

Tumanov O. O. Large Language Models: A New Paradigm for Data Analysis

The emergence of large language models (LLMs) has fundamentally transformed the paradigm of data analysis, affecting fields such as social media analytics, healthcare, and software development. These models, examples of which include architectures like GPT and BERT, handle tasks related to unstructured text and contextual understanding very effectively, but their widespread adoption requires a critical assessment of their relationship to and differences from traditional statistical approaches. This article aims to provide a comparative analysis of LLMs and traditional statistical methods, examining their strengths, limitations, and areas of application. The study uses a qualitative comparison method based on the following criteria: data type (unstructured vs. structured), interpretability, purpose, data requirements, and reproducibility. The analysis shows that, despite the effectiveness of LLMs in natural language text processing, their probabilistic nature imposes significant limitations on reliability and ethics. Traditional statistical methods, in contrast, offer high interpretability and reproducibility, which are critically important for deriving deterministic conclusions. Special attention is paid to the limitations of both approaches. Specifically, LLMs are prone to «hallucinations» and inheriting biases from training data, which can lead to discriminatory outcomes. Meanwhile, traditional statistical models are not immune to selection biases. As a result, the study found that LLMs and traditional methods are not competitors but complementary tools, each best suited for addressing certain tasks. The main conclusion is that the most promising direction is the integration of these approaches into hybrid methodologies. Large language models (LLMs) can be effectively used for preprocessing and converting unstructured text into structured quantitative data, which can then be analyzed using classical statistical methods. This approach allows overcoming existing limitations and achieving greater reliability, transparency, and accuracy in scientific research.
Keywords: *large language models, statistical methods, data analysis, artificial intelligence, ethical considerations, hybrid model.*
Fig.: *5.* ***Tabl.:*** *2.* ***Bibl.:*** *12.*

Tumanov Oleksii O. *– PhD, Senior Lecturer, V. N. Karazin Kharkiv National University (4 Svobody Square, Kharkiv, 61022, Ukraine)*
E-mail: *oleksii.tumanov@gmail.com*
ORCID: *https://orcid.org/0000-0003-0674-0037*

Швидка еволюція технологій штучного інтелекту обумовила появу нового класу моделей, які стали новою парадигмою для аналізу даних: великих мовних моделей (ВММ). Пройшовши навчання на величезних і різноманітних масивах даних, ВММ демонструють безпрецедентну здатність генерувати, узагальнювати та розуміти текст, що імітує людське спілкування. У контексті попередніх праць автора, присвячених статистичному аналізу соціальних медіа, поява ВММ означає значне відхилення від традиційних підходів [12]. Ці підходи історично зосеред-

жувалися на структурованих кількісних даних, що були в центрі уваги прогнозування та кластерного аналізу [4]. Сьогодні ВММ активно впроваджуються у системах, починаючи від охорони здоров'я, де вони сприяють біомедичним дослідженням, і до розробки програмного забезпечення, де їхнє використання підвищує продуктивність.

Незважаючи на їхні беззаперечні досягнення, імовірнісний та часто непрозорий характер, ВММ ставить під сумнів усталені статистичні принципи. У той час як дослідники зі статистичним досвідом високо цінують такі поняття, як p-значення, довірчі інтервали та інтерпретованість моделі, ВММ часто описують як «чорні скриньки», внутрішня робота яких є непрозорою. Це породжує ключові питання щодо їхньої надійності та належного використання в дослідженнях. Хоча ВММ пропонують значні переваги, порівняння з традиційними методами виявляє низку важливих співвідношень переваг та обмежень, зокрема щодо типів даних, інтерпретованості та відтворюваності.

З огляду на це, у цій статті ґрунтовно аналізуються ВММ крізь призму усталених статистичних практик. Досліджується їхнє сучасне застосування, підкреслюючи сильні сторони в обробці складних, неструктурованих даних. Також проводиться ретельне статистичне порівняння їхніх переваг і недоліків порівняно зі стандартним статистичним моделюванням. У роботі стверджується, що ВММ і традиційні статистичні методи не є взаємовиключними, а радше доповнюють одне одного, і кожен з яких має унікальні переваги та обмеження.

Експоненціальне зростання великих мовних моделей (ВММ) ініціювало хвилю наукових публікацій, що досліджують їхній потенціал та обмеження в порівнянні з усталеними методами аналізу даних. У вітчизняному науковому дискурсі особлива увага приділяється використанню штучного інтелекту (ШІ) та впливу ВММ на галузеві дослідження. Так, Мойсеєнко М. І., Кузишин М. М., Туровська Л. В. та інші досліджують стратегічні аспекти використання ВММ у сфері охорони здоров'я, підкреслюючи їхні можливості в аналізі даних [1]. Водночас Потятиник Б. досліджує можливості та обмеження використання ШІ в аналізі медіапотоків, а саме, застосування мовних моделей Claude та Google-PaLM для аналізу висвітлення саміту БРІКС у матеріалах Reuters, Укрінформу і ТАСС [2].

Ключовим аспектом, який послідовно підкреслюється в сучасних наукових публікаціях, є теза про взаємодоповнюваність, а не суперечливість між великими мовними моделями (ВММ) і традиційними статистичними підходами. Нещодавні дослідження, зокрема опубліковані на arXiv, акцентують, що, попри високу ефективність ВММ у роботі з неструктурованими текстовими даними, їх ймовірнісна природа зумовлює потребу інтеграції з класичними статистичними методами [9]. Такий синтез дозволяє підвищити надійність отриманих висновків, забезпечити прозорість аналітичних процедур і сприяти коректній інтерпретації результатів.

У цьому контексті особливу значущість має робота Kumar N. та ін., які у своїй публікації пропонують багатовимірну структуру для відповідальної оцінки та вибору ВММ [10]. Ця праця підкреслює, що критерії вибору не повинні обмежуватися лише ефективністю, а мають також включати етичні аспекти, як-от упередженість, конфіденційність та надійність. Це відповідає ідеї, що вибір ВММ має базуватись на широкому спектрі критеріїв, що виходять за рамки лише технічних характеристик.

Незважаючи на значні досягнення, більшість досліджень зосереджена на демонстрації ефективності ВММ без суворого статистичного порівняння із загальноприйнятими кількісними методами. Таким чином, питання їх взаємодії, надійності та інтеграції у традиційний статистичний аналіз залишається недостатньо вивченим, що є невирішеною частиною загальної проблеми.

Метою цієї статті є проведення ретельного статистичного порівняння великих мовних моделей (ВММ) з традиційними методами аналізу даних і демонстрація того, як ВММ можуть служити надійним статистичним інструментом для трансформації неструктурованого тексту в кількісні емоційні показники. Це дозволяє подолати розрив між якісним аналізом і кількісними дослідженнями, обґрунтовуючи взаємодоповнювальний характер обох парадигм для досягнення більш глибоких і точних наукових висновків.

1. Великі мовні моделі: широке впровадження та застосування. Впровадження великих мовних моделей (ВММ) експоненціально зросло в усіх галузях. Глобальні тенденції показують, що значний відсоток організацій уже впровадив генеративний штучний інтелект, і, за прогнозами, ринок генеративного ШІ зросте до сотень мільярдів доларів.

З погляду статистики, мовна модель – це не просто генератор тексту, а складна ймовірнісна модель, що передбачає послідовність слів. Вона базується на принципах умовного розподілу ймовірностей, де ймовірність появи кожного наступного слова (токену) залежить від ймовірності появи попередніх (*рис. 1*). Таким чином, замість того, щоб бути лише лінгвістичним інструментом, ВММ є потужним статистичним апаратом для вивчення та моделювання взаємозв'язків між лінгвістичними одиницями. Це дозволяє використовувати їх для

отримання кількісних показників, що є центральною тезою нашого дослідження.

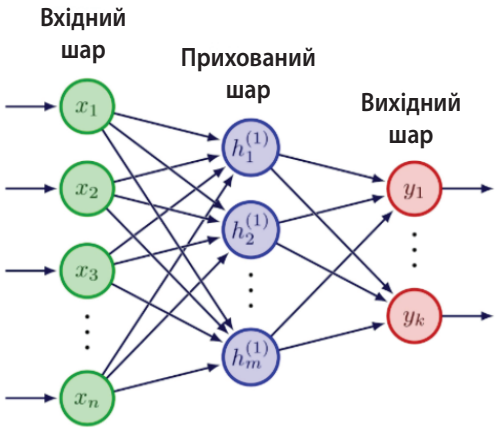

Рис. 1. Схематичне представлення архітектури нейронної мережі як основи для ВММ

Джерело: побудовано автором.

2. Сфери використання великих мовних моделей. ВММ відіграють ключову роль у трансформації різних сфер, і їхнє застосування особливо помітне в таких галузях, як:

- *Соціальні медіа та аналіз контенту.* ВММ використовуються для аналізу тональності даних соціальних мереж у масштабі та з нюансами, які раніше були недосяжними. На відміну від традиційних методів, що базуються на ключових словах або лексиконах, ВММ можуть зрозуміти сарказм, іронію та складний контекст, забезпечуючи більш точні оцінки настроїв [3]. Вони також сприяють створенню персоналізованого контенту та узагальненню великих обсягів контенту, створеного користувачами.
- *Охорона здоров'я.* У цій галузі ВММ інтегруються в клінічні та адміністративні робочі процеси. ВММ використовуються для узагальнення приміток пацієнтів, складання клінічної документації та відповідей на запитання пацієнтів, що значно знижує адміністративне навантаження на медичний персонал. Наприклад, у Великій Британії пілотний проєкт із використанням ШІ-інструменту допоміг скоротити кількість пропущених прийомів майже на третину, а в Іспанії система ШІ-сортування допомогла зменшити час очікування пацієнтів у відділеннях невідкладної допомоги [5].
- *Державний сектор.* За даними Управління з підзвітності уряду США, кількість випадків використання ШІ у федеральних установах майже подвоїлася з 2023 по 2024 роки, при цьому випадки використання генеративного ШІ зросли в дев'ять разів (*рис. 2*). Тоді як в Європейському Союзі, згідно з даними Євростату [6], понад 13% підприємств використовували технології ШІ у 2024 році, при цьому технології, які генерують текст, були одними з найпоширеніших (*рис. 3*). Це підтверджує, що ВММ стають невід'ємною частиною офіційної статистики та урядових ініціатив.
- *Розробка програмного забезпечення.* Роль великих мовних моделей у розробці програмного забезпечення постійно зростає. Ці моделі діють як інтелектуальні помічники, значно підвищуючи продуктивність розробників і оптимізуючи робочі процеси. ВММ допомагають на всіх етапах життєвого циклу розробки програмного забезпечення (SDLC) – від ініціації проєкту до підтримки. Зокрема, генерація коду є однією з найпомітніших функцій. Хоча перші інструменти автодоповнення коду, як-от Tabnine, з'явилися ще у 2018 році, вони використовували менші моделі [11]. Нова хвиля помічників, заснована на потужних ВММ, розпочалася із запуску GitHub Copilot компаніями GitHub та OpenAI у червні 2021 року, що стало значним проривом [8].

Ці моделі можуть автоматично генерувати фрагменти коду, а також цілі функції та класи на основі короткого опису задачі природною мовою. Це прискорює процес написання коду, особливо для рутинних або стандартних завдань. Крім того, вони відіграють важливу роль у налагодженні та оптимізації коду, допомагаючи знаходити помилки та пропонуючи ефективніші алгоритми. Таким чином, ВММ не просто автоматизують окремі завдання, а інтегруються в екосистему розробки, перетворюючи її на більш ефективний і продуктивний процес.

3. Переваги та обмеження: порівняння великих мовних моделей і традиційних статистичних підходів. Незважаючи на те, що ВММ пропонують значні переваги, їхнє пряме порівняння з усталеними статистичними методами виявляє низку важливих співвідношень переваг та обмежень. Основні порівняльні характеристики ВММ і традиційних статистичних методів наведено в *табл. 1*.

3.1. Переваги великих мовних моделей. Аналізуючи наведену табл. 1, можна виділити такі сильні сторони ВММ.

- *Обробка неструктурованих даних.* Великі мовні моделі відмінно справляються із завданнями, що включають неструктуровані дані, зокрема текст природною мовою. На відміну від статистичних моделей, які ви-

Примітка: інформація агентств та аналіз GAO-звітів про випадки використання штучного інтелекту (ШІ) в агентствах, поданих до Управління з питань менеджменту та бюджету | GAO-25-107653

Рис. 2. Використання ШІ у федеральних установах США за 2023–2024 рр.

Джерело: побудовано за [7].

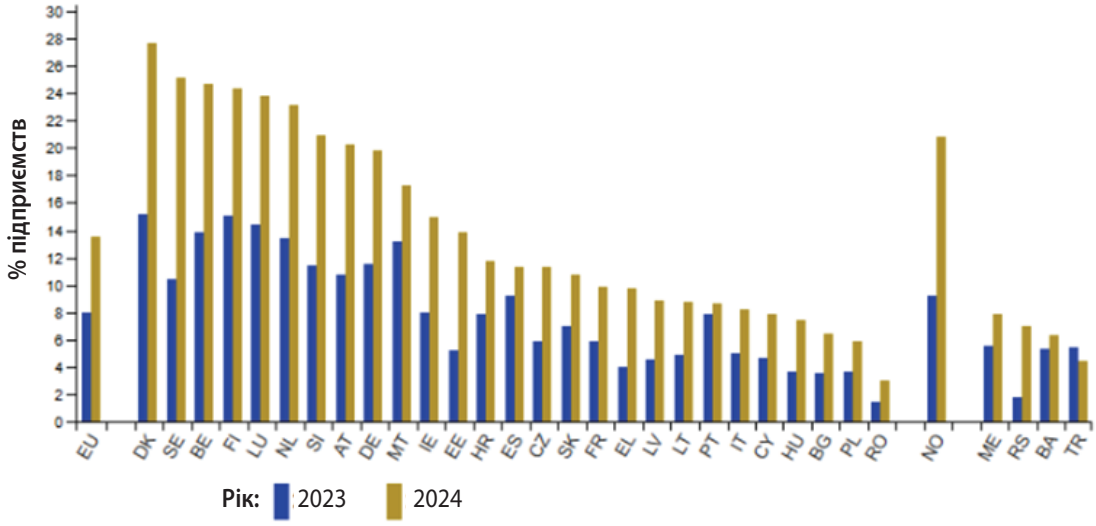

Рис. 3. Відсоток підприємств у розрізі країн Європейського Союзу, що використовували ШІ у 2023–2024 рр.

Джерело: Євростат [6].

магають структурованих числових вхідних даних, ВММ можуть обробляти та отримувати інформацію з даних у довільній формі, таких як публікації в соціальних мережах, історії пацієнтів та наукові статті.

- *Узагальнення та навчання з мінімальним прикладом.* Пройшовши підготовку на масивних корпусах, ВММ володіють обширними «світовими знаннями». Це дозволяє їм ефективно виконувати нові завдання з мінімальними або без спеціальних навчальних даних. Це становить різкий контраст із традиційними статистичними моделями, які часто вимагають великих, позначених наборів даних для ефективного навчання.
- *Творчість і генерація.* Ключовою перевагою ВММ є їхня генеративна здатність. Вони можуть створювати новий, послідовний і кон-

Таблиця 1

Порівняння ВММ і традиційних статистичних методів

Критерій порівняння	Великі мовні моделі (LLM)	Традиційні статистичні методи
Тип даних	Неструктуровані (текст, зображення, аудіо)	Переважно структуровані (числові, категоріальні)
Інтерпретованість	Низька («чорний ящик»)	Висока (зрозумілі коефіцієнти, p-значення)
Призначення	Генерація, узагальнення, розуміння контексту	Прогнозування, причинно-наслідковий зв'язок, виведення
Вимоги до даних	Екстремально великі корпуси (терабайти)	Часто менші, але репрезентативні вибірки
Відтворюваність	Низька (імовірнісна природа, «галюцинації»)	Висока (детерміновані, математично обґрунтовані)
Масштабованість	Дуже висока (зростає з обсягом даних)	Обмежена (залежить від обчислювальної складності)

Джерело: побудовано автором.

текстуально релевантний контент, наприклад маркетингові тексти для соціальних медіа або початкові чернетки резюме досліджень. Ця можливість виходить за рамки описових чи прогностичних функцій більшості стандартних статистичних моделей.

3.2. Переваги стандартних статистичних підходів. Своєю чергою, стандартні статистичні підходи мають переваги в такому.

- *Інтерпретованість і пояснюваність.* Традиційні статистичні моделі (наприклад, лінійна та логістична регресія, ANOVA) мають високу інтерпретованість. Коефіцієнти в цих моделях мають чіткі, апріорні значення, що дозволяє дослідникові зрозуміти, чому було зроблено конкретне передбачення. Це має вирішальне значення в таких галузях, як охорона здоров'я, де обґрунтування рішення має бути прозорим і підлягати перевірці. І навпаки, ВММ часто називають «чорними скриньками», оскільки їхня внутрішня робота є непрозорою.
- *Статистична точність і висновок.* Стандартні статистичні методи побудовані на основі математичної та статистичної теорії. Вони розроблені для причинно-наслідкових висновків, перевірки гіпотез і кількісного визначення невизначеності. Вони забезпечують засоби для вимірювання мінливості (наприклад, стандартної помилки) та для створення надійних, відтворюваних висновків про генеральну сукупність на основі вибірки. Ці принципи є важливими для того, щоб зробити обґрунтовані висновки на основі даних. Результати ВММ, будучи імовірнісними, можуть бути недетермінованими та схильними до «галюцинацій» (фабрикації інформації).
- *Обчислювальна ефективність і вимоги до даних.* Для аналізу структурованих даних традиційні статистичні моделі часто є більш обчислювально ефективними та вимагають значно менше даних для навчання та розгортання, ніж великі ВММ. Більше того, хоча ВММ можуть здійснювати короткочасне навчання, вони все ще схильні до перенавчання, коли навчальні дані обмежені, що є проблемою, щодо якої статистичні методи часто більш надійні.

З огляду на зазначені сильні сторони ВММ і традиційних підходів, стає очевидно, що вибір методу залежить від конкретної задачі та характеру даних. Щоб проілюструвати цю взаємодію, у *табл. 2* наведено порівняння рішень на основі ВММ і статистичних аналогів у різних сферах застосування.

4. Обмеження та етичні міркування. Обидва підходи мають притаманні їм обмеження, особливо щодо упередженості та етичного використання. Зокрема, великі мовні моделі (ВММ) відомі тим, що вони успадковують і підсилюють упередження, присутні в їхніх навчальних даних, що може призвести до дискримінаційних або несправедливих результатів. Наприклад, ВММ, орієнтована на сферу охорони здоров'я, може закріпити існуючі нерівності, якщо її навчальні дані недостатньо репрезентують певні демографічні групи. Подібним чином традиційні статистичні моделі не захищені від упередженості.

Такі центральні для статистичної теорії концепції, як упередженість відбору та зміщення ви-

Таблиця 2

Приклади застосування ВММ і аналогічних традиційних статистичних методів

Сфера застосування	Завдання	Рішення на основі ВММ	Рішення на основі статистичних методів
Соціальні медіа	Аналіз настроїв	Розуміння сарказму, іронії, контексту	Класифікація тексту на основі словника, частотний аналіз
Охорона здоров'я	Аналіз клінічних записів	Вилучення й узагальнення ключової інформації з вільного тексту	Прогноз результатів лікування на основі структурованих параметрів (тиск, пульс)
Бізнес-аналітика	Прогнозування трендів	Аналіз настроїв у соцмережах для прогнозу попиту на продукт	ARIMA-моделі, експоненційне згладжування, регресійний аналіз

Джерело: побудовано автором.

бірки, можуть спотворити результати моделі, якщо дані не є репрезентативними. Проте структурований характер статистичних методів дозволяє використовувати більш прозорий і систематичний підхід до виявлення та пом'якшення цих проблем.

Досліднику легше простежити джерело упередження в регресійній моделі, ніж у мільярдах параметрів ВММ, що схематично зображено на *рис. 4*.

Ще одним критичним обмеженням ВММ є їхня схильність до «галюцинацій» – генерації фактично неточних або сфабрикованих даних. Ця проблема виникає через імовірнісну природу моделей, які генерують найбільш імовірну послідовність слів, не перевіряючи її на відповідність реальним фактам.

У таких чутливих сферах, як медицина чи юриспруденція, подібні «галюцинації» можуть мати серйозні наслідки, тоді як результати традиційних статистичних моделей, побудованих на математичних принципах, є детермінованими та підлягають верифікації. Таким чином, для забезпечення надійності та точності результатів вкрай важливо впроваджувати суворі протоколи валідації.

5. Поєднання підходів. Найефективніший шлях розвитку аналітичних методів полягає не в перевазі одного підходу над іншим, а в їхній інтеграції. Дослідник, який володіє досвідом як у статистиці, так і в розробці програмного забезпечення, має унікальні можливості для керування цією інтеграцією. На *рис. 5* проілюстровано потенційну схему поєднання ВММ і традиційних статистичних методів для комплексного аналізу даних.

Як показано на рис. 5, ВММ можуть бути використані для попередньої обробки та узагальнення неструктурованих даних соціальних медіа, які потім можуть бути введені в традиційну статистичну модель для надійного прогнозування або

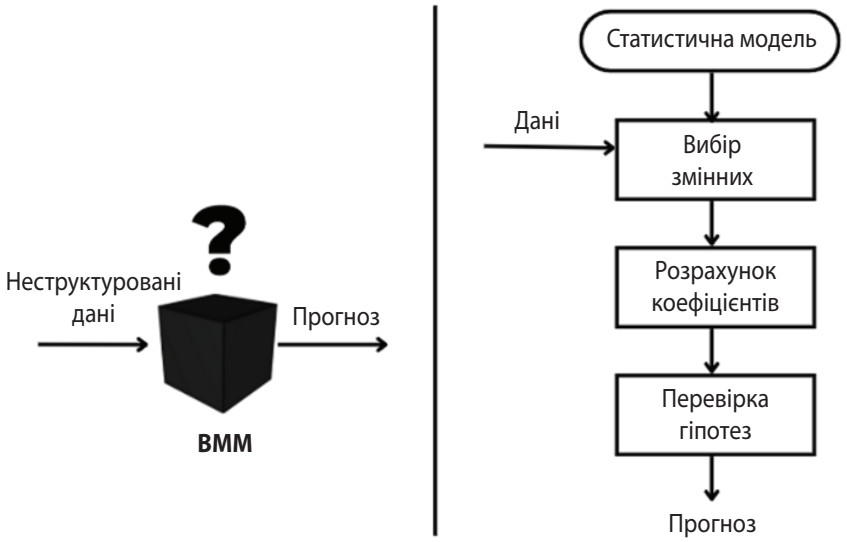

Рис. 4. Порівняння прозорості ВММ і статистичної моделі

Джерело: побудовано автором.

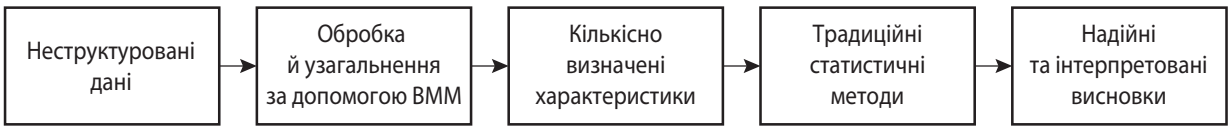

Рис. 5. Спрощена схема використання ВММ і традиційних статистичних методів в аналізі даних

Джерело: побудовано автором.

кластерного аналізу. Аналогічно, у сфері охорони здоров'я ВММ можуть допомогти проаналізувати та витягти ключову інформацію з клінічних записів, яку стандартна статистична модель могла б потім використовувати для прогнозування результатів пацієнтів.

Зрештою, потужність ВММ полягає в їхній здатності опрацьовувати «неупорядкований», неструктурований світ людської мови, тоді як сила статистичних методів полягає в їхній здатності забезпечити сувору, прозору та кількісно визначену структуру для її розуміння. Використовуючи сильні сторони обох підходів, ми можемо відкрити нові шляхи для досліджень і створити більш надійні гібридні системи штучного інтелекту.

ВИСНОВКИ

Проведене дослідження підтверджує, що поява великих мовних моделей (ВММ) відкрила нову парадигму для аналізу даних, доповнюючи традиційні статистичні підходи. ВММ демонструють безпрецедентні можливості в обробці складних неструктурованих даних, зокрема природної мови, що дозволяє вирішувати задачі, які раніше були недосяжними для класичних методів. Однак, як було показано, їхній імовірнісний і непрозорий характер створює суттєві обмеження щодо відтворюваності, інтерпретованості та упередженості.

Водночас традиційні статистичні методи залишаються незамінними. Їхня сила полягає в детермінованості, прозорості та математичній обґрунтованості, що є критично важливим для отримання надійних і валідованих висновків. Порівняльний аналіз показав, що вибір між ВММ і традиційними методами залежить від конкретної задачі, типу даних і вимог до інтерпретованості результатів.

Головним висновком статті є те, що ВММ і традиційна статистика є не конкурентами, а взаємодоповнювальними інструментами. Найбільший потенціал для майбутніх досліджень лежить у розробці гібридних методологій, які використовують переваги обох підходів. Наприклад, ВММ можуть ефективно застосовуватися для попередньої обробки та перетворення неструктурованого тексту в структуровані кількісні дані, які потім можуть бути піддані суворому статистичному аналізу.

Перспективи подальших досліджень. Майбутні дослідження в цьому напрямку можуть бути зосереджені на таких аспектах:

- *Розробка стандартів для оцінки надійності ВММ*. Необхідно створити уніфіковані метрики для кількісного вимірювання «галюцинацій» та упередженості ВММ, що дозволить використовувати їх у чутливих сферах, таких як медицина та юриспруденція.
- *Створення гібридних моделей*. Розробка інтегрованих платформ, які автоматично поєднують сильні сторони ВММ і статистичні методи, щоб підвищити точність і надійність аналізу даних.
- *Дослідження етичних аспектів*. Подальше вивчення етичних викликів, що виникають при використанні ВММ, особливо в контексті конфіденційності даних і соціальної справедливості.
- *Адаптація методологій до українського контексту*. Розробка ВММ, навчених на україномовних корпусах даних, та їх інтеграція у вітчизняні наукові дослідження та державні установи.

Таким чином, інтеграція ВММ у сферу аналізу даних не лише змінює інструментарій дослідника, але й вимагає переосмислення фундаментальних принципів, що відкриває широке поле для подальших наукових розвідок. ■

БІБЛІОГРАФІЯ

1. Мойсеєнко М. І., Кузишин М. М., Туровська Л. В. та ін. Великі мовні моделі штучного інтелекту в медицині. *Сучасні інформаційні технології та інноваційні методики навчання в підготовці фахівців: методологія, теорія, досвід, проблеми*. 2024. Вип. 72. С. 73–88. DOI: https://doi.org/10.31652/2412-1142-2024-72-73-88
2. Потятиник Б. Використання ШІ в аналізі медіапотоків: можливості та обмеження. *Вісник Львівського університету. Серія «Журналістика»*. 2025. Вип. 57. С. 62–74. DOI: https://doi.org/10.30970/vjo.2025.57.13290
3. Туманов О. О. Розробка системи статистичних показників для моніторингу емоційного стану у соціальних медіа. *Матеріали III Міжнародної науко-*

во-практичної конференції «Східноєвропейський центр наукових досліджень» (м. Одеса, 16 серпня 2025 р.). Одеса : Research Europe, 2025. С. 114–118. DOI: https://doi.org/10.64076/eecsr250816
4. Туманов О. О. Статистичне оцінювання розвитку соціальних медіа : автореф. … канд. екон. наук : 08.00.10. Київ, 2021. URL: http://nasoa.edu.ua/wp-content/uploads/zah/tumanov_avt.pdf
5. EU report finds promise and hurdles in using AI for healthcare. *CADE*. 14.08.2025. URL: https://cadeproject.org/updates/eu-report-finds-promise-and-hurdles-in-using-ai-for-healthcare/
6. Use of artificial intelligence in enterprises. *Eurostat*. 2024. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php/Use_of_artificial_intelligence_in_enterprises
7. Artificial Intelligence: Generative AI Use and Management at Federal Agencies. *GAO*. 29.07.2025. URL: https://www.gao.gov/products/gao-25-107653
8. Introducing GitHub Copilot: your AI pair programmer. *GitHub Blog*. 2021. URL: https://github.blog/news-insights/product-news/introducing-github-copilot-ai-pair-programmer/
9. Ji W., Yuan W., Getzen E. et al. An Overview of Large Language Models for Statisticians. arXiv preprint. *arXiv:2502.17814*. 25.02.2025. URL: https://arxiv.org/html/2502.17814v1
10. Kumar N., Ramavath Sh. K., Venkatesh V. Bridging Context, Statistics, and Practice: A Multi-Dimensional Framework for Responsible LLM Evaluation and Selection. *TechRxiv*. 2023. DOI: https://doi.org/10.36227/techrxiv.175691248.85284358/v1
11. Tabnine. About | Tabnine: The AI code assistant that you control. 2025. URL: https://www.tabnine.com/about/
12. Туманов О. О. Статистичні методи аналізу даних соціальних медіа. *Бізнес Інформ*. 2020. № 2. С. 266–272. DOI: https://doi.org/10.32983/2222-4459-2020-2-266-272

REFERENCES

CADE. (2025, August 14). *EU report finds promise and hurdles in using AI for healthcare*. https://cadeproject.org/updates/eu-report-finds-promise-and-hurdles-in-using-ai-for-healthcare/

Eurostat. (2024). *Use of artificial intelligence in enterprises*. https://ec.europa.eu/eurostat/statistics-explained/index.php/Use_of_artificial_intelligence_in_enterprises

GAO. (2025, July 29). *Artificial Intelligence: Generative AI Use and Management at Federal Agencies (GAO-25-107653)*. https://www.gao.gov/products/gao-25-107653

GitHub Blog. (2021). *Introducing GitHub Copilot: your AI pair programmer*. https://github.blog/news-insights/product-news/introducing-github-copilot-ai-pair-programmer/

Ji, W., Yuan, W., Getzen, E., et al. (2025, February 25). *An Overview of Large Language Models for Statisticians*. https://arxiv.org/html/2502.17814v1

Kumar, N., Ramavath, Sh. K., & Venkatesh, V. (2023). *Bridging Context, Statistics, and Practice: A Multi-Dimensional Framework for Responsible LLM Evaluation and Selection*. TechRxiv. https://doi.org/10.36227/techrxiv.175691248.85284358/v1

Moiseienko, M. I., Kuzyshyn, M. M., Turovska, L. V., & in. (2024). Velyki movni modeli shtuchnoho intelektu v medytsyni [Large language models of artificial intelligence in medicine]. *Suchasni informatsiini tekhnolohii ta innovatsiini metodyky navchannia v pidhotovtsi fakhivtsiv: metodolohiia, teoriia, dosvid, problemy, 72*, 73–88. https://doi.org/10.31652/2412-1142-2024-72-73-88

Potiatynyk, B. (2025). Vykorystannia ShI v analizi mediapotokiv: mozhlyvosti ta obmezhennia [The use of AI in the analysis of media flows: opportunities and limitations]. *Visnyk Lvivskoho universytetu. Seriia «Zhurnalistyka», 57*, 62–74. https://doi.org/10.30970/vjo.2025.57.13290

Tabnine. (2025). About | *Tabnine: The AI code assistant that you control*. https://www.tabnine.com/about/

Tumanov, O. O. (2020). Statystychni metody analizu danykh sotsialnykh media [Statistical methods of social media data analysis]. *Biznes Inform, (2)*, 266–272. https://doi.org/10.32983/2222-4459-2020-2-266-272

Tumanov, O. O. (2021). *Statystychne otsiniuvannia rozvytku sotsialnykh media* [Statistical assessment of social media development] [Avtoreferat dysertatsii kandydata ekonomichnykh nauk, Natsionalna akademiia statystyky, obliku ta audytu]. http://nasoa.edu.ua/wp-content/uploads/zah/tumanov_avt.pdf

Tumanov, O. O. (2025). Rozrobka systemy statystychnykh pokaznykiv dlia monitorynhu emotsiinoho stanu u sotsialnykh media [Development of a system of statistical indicators for monitoring the emotional state in social media]. *Materialy III Mizhnarodnoi naukovo-praktychnoi konferentsii «Skhidnoievropeiskyi tsentr naukovykh doslidzhen»*, 114–118. Research Europe.