DOI: https://doi.org/10.25140/2411-5363-2025-4(42)-245-252
УДК 004.8

Богдан Миколайович Пархоменко[1], Максим Валерійович Міщенко[2]

[1]аспірант кафедри інформаційних технологій та програмної інженерії,
Національний університет «Чернігівська політехніка» (Чернігів, Україна)
E-mail: bparkhomenko@stu.cn.ua. **ORCID:** https://orcid.org/0009-0005-1279-4981

[2] доктор філософії зі спеціальності «Комп'ютерні науки»,
викладач кафедри інформаційних технологій та програмної інженерії
Національний університет «Чернігівська політехніка» (Чернігів, Україна)
E-mail: max.mishchenko771@gmail.com. **ORCID:** https://orcid.org/0000-0001-9769-9759

ЗАСТОСУВАННЯ МОВНИХ МОДЕЛЕЙ ВЕЛИКОГО МАСШТАБУ ДЛЯ АНАЛІЗУ ЗАГОЛОВКІВ ТА ФОРМУВАННЯ РОЗШИРЕНИХ ОЗНАК У ПРОГНОЗУВАННІ ФІНАНСОВИХ РИНКІВ

Стаття є оглядово-інформаційним викладом матеріалу щодо використання сучасних мовних моделей великого масштабу для аналізу заголовків фінансових новин та формування індексу інформаційного настрою з метою покращення якості фінансових даних і підвищення точності подальшого прогнозування показників. В умовах зростання обсягів неструктурованої інформації та підвищеної динамічності фінансових ринків традиційні методи аналізу дедалі частіше виявляються недостатньо ефективними, що зумовлює необхідність застосування інтелектуальних інформаційних моделей. Поєднання глибинного семантичного аналізу текстових даних із методами машинного навчання забезпечує можливість розширення наборів ознак, збагачення вхідних даних та підвищення інформативності фінансових часових рядів.

На основі отриманих числових оцінок тональності будується коефіцієнт інформаційного настрою, який агрегує вплив інформаційного фону за певний часовий період та інтегрується у фінансові ряди як додаткова характеристика. Такий підхід дозволяє враховувати не лише історичну цінову динаміку, а і якісні аспекти інформаційного середовища, що впливають на поведінку учасників ринку та формування ринкових очікувань. Запропонований індекс може використовуватися як допоміжний індикатор для аналізу трендових змін і підсилення прогнозних моделей.

Використання запропонованих методів сприяє виявленню прихованих залежностей між інформаційним фоном та ринковими показниками, що є особливо важливим в умовах високої волатильності та невизначеності фінансових систем. Окрему увагу приділено питанням інтерпретованості отриманих результатів та стабільності побудованого індексу при зміні параметрів агрегації. На основі проведеного аналізу визначаються основні напрями подальшого удосконалення підходів до формування інформаційних індексів, а також можливості практичного застосування отриманих результатів у діяльності фінансових аналітиків та при побудові моделей прогнозування фінансових часових рядів.

***Ключові слова**: мовні моделі; машинне навчання; фінансова аналітика; інформаційні моделі; прийняття рішень.*
Табл.: 2. Рис.: 3. Бібл.: 7.

Актуальність теми дослідження. Використання сучасних мовних моделей великого масштабу для аналізу заголовків фінансових новин є актуальним в умовах зростання кількості доступної інформації, що впливає на поведінку учасників ринку та формування цінових трендів. Традиційні методи обробки текстів втрачають ефективність через складність інтерпретації фінансової інформації та швидкості її оновлення. Застосування підходів, заснованих на LLM, дозволяє точніше оцінювати інформаційний фон ринку та підтримувати процес прийняття рішень.

Дослідження має важливе значення для фінансової аналітики, оскільки формування індексу інформаційного настрою створює можливість розширення наборів ознак і покращення якості прогнозування. Використання таких інструментів підвищує адаптивність аналітичних моделей у періоди волатильності та турбулентності ринку. Подальший розвиток методів аналізу відкриває нові перспективи для підвищення ефективності управління інвестиціями.

Постановка проблеми. У зв'язку зі стрімким зростанням обсягів інформації та збільшенням впливу новинних потоків на фінансові ринки виникає необхідність розробки ефективних методів аналізу текстових даних, які можуть бути використані для покращення фінансових прогнозів. Традиційні алгоритми обробки текстів не здатні повною мірою врахувати контекст, динамічність та багатозначність заголовків фінансових новин, що обмежує їхню ефективність у задачах прогнозування.

Особливо важливою проблема стає для галузі фінансової аналітики та інвестування, де точність інтерпретації інформаційного фону ринку безпосередньо впливає на ухвалення рішень. Використання сучасних мовних моделей великого масштабу дозволяє значно підвищити якість аналізу текстових даних та створити нові інформаційні коефіцієнти, що розширюють можливості існуючих прогнозних моделей. Це набуває особливої актуальності в умовах ринкової невизначеності, де швидкий і точний аналіз настроїв може забезпечити суттєву аналітичну перевагу.

Аналіз останніх досліджень і публікацій. Упродовж останніх років зростає кількість робіт, присвячених застосуванню мовних моделей великого масштабу для обробки фінансових текстів та покращення аналітичних процесів. Сучасні LLM демонструють високу ефективність у визначенні семантичних характеристик та тональності фінансових повідомлень, що робить їх перспективними для використання в задачах прогнозування. Водночас дослідження вказують на низку проблем, серед яких – складність побудови універсальних моделей, здатних стабільно працювати на різнорідних фінансових даних, а також залежність результатів від якості текстових джерел та методів їх попередньої обробки [1].

Окремий напрям становлять роботи, спрямовані на створення спеціалізованих фінансових мовних моделей, які демонструють вищу точність у задачах сентимент-аналізу та інтерпретації ринкових новин. Такі моделі покращують результати аналізу, проте їх використання потребує ретельного налаштування та адаптації до певного типу текстових даних [2].

У новітніх дослідженнях набувають поширення ансамблеві підходи, де поєднання декількох мовних моделей дозволяє підвищити точність оцінювання настроїв ринку та знизити вплив окремих помилок класифікації. Такі методи демонструють перспективність, але не завжди враховують специфіку заголовків новин та особливості їх використання у формуванні інтегральних індексів інформаційного настрою [3].

Виділення недосліджених частин загальної проблеми. Попри значний прогрес у застосуванні мовних моделей для аналізу фінансових текстів, залишається низка аспектів, що вивчені недостатньо. Однією з ключових проблем є відсутність сталих методик формування індексів інформаційного настрою саме на основі заголовків новин, які є короткими, динамічними та часто містять концентровані ринкові сигнали.

Недостатньо досліджено також вплив запитів на отримані результати, оскільки якість запиту до мовної моделі безпосередньо визначає точність та стабільність оцінок тональності. У більшості робіт процес побудови промптів описано частково, що ускладнює відтворюваність результатів. Крім того, питання узгоджування оцінок, отриманих різними моделями, залишаються відкритими.

Таким чином, потребують додаткового вивчення методи аналізу коротких текстів - заголовків як швидкодоступного інформаційного потоку, а також побудова ефективних коефіцієнтів настрою, які можуть бути інтегровані у фінансові дані для подальшого покращення прогнозування.

Метою статті є дослідження використання сучасних мовних моделей великого масштабу для аналізу заголовків фінансових новин, розробка підходів до їхнього промптингу та формування коефіцієнта інформаційного настрою, що може бути інтегрований у фінансові дані з метою покращення їхньої інформативності й підвищення точності подальших прогнозів.

Виклад основного матеріалу. Сучасний розвиток штучного інтелекту та стрімке збільшення обсягів доступної інформації зумовлюють необхідність використання більш гнучких і точних методів для аналізу фінансових даних. Особливо це стосується текстових даних, які відіграють ключову роль у формуванні ринкових очікувань та поведінки інвесторів. Новини, аналітичні огляди та повідомлення про корпоративні події створюють складний інформаційний фон, який традиційні статистичні методи більше не в змозі

ефективно інтерпретувати [4]. Саме тому у фінансових дослідженнях усе ширше застосовуються мовні моделі великого масштабу (LLM), здатні аналізувати контекстуальні залежності та виявляти приховані семантичні зв'язки в текстах.

У цьому дослідженні для класифікації тональності фінансових заголовків застосовано модель RoBERTa, адаптовану для коротких текстових повідомлень. Обрана модель демонструє високу ефективність у визначенні позитивної, нейтральної чи негативної тональності завдяки попередньому навчанню на великому корпусі коротких новинних та соціальних повідомлень, де лінгвістична структура подібна до фінансових заголовків [5]. Її перевагами є здатність працювати без додаткового донавчання, висока стійкість до шумних даних і доступність у вигляді відкритої моделі, що дозволяє проводити дослідження без значних обчислювальних витрат.

Використання LLM є ключовим елементом процесу побудови індексу інформаційного настрою S_t:

$$S_t = \frac{\sum_{i=1}^{N_t} \omega_i * sentiment(T_i)}{\sum_{i=1}^{N_t} \omega_i},$$

де T_i – окреме текстове повідомлення отримане в момент часу t;

$sentiment(T_i)$ – числова оцінка тональності тексту;

ω_i – ваговий коефіцієнт для кожного тексту, який може враховувати авторитетність джерела, кількість переглядів, поширень або іншу релевантність;

N_t – загальна кількість текстів за розглянутий проміжок часу;

S_t – узагальнений індекс інформаційного настрою, який характеризує тон інформаційного середовища в момент часу t.

Для формування кількісного індексу інформаційного настрою використовується структурований набір даних, який містить заголовки фінансових новин, відповідні дати їхніх публікації та тикери компаній, до яких ці новини належать [6]. Приклад даних наведено в табл. 1.

Таблиця 1 – Приклад вхідних даних із заголовками новин

Новина	Дата	Індекс компанії
B of A Securities Maintains Neutral on Agilent Technologies, Raises Price Target to $88"	2020-05-22	A
CFRA Maintains Hold on Agilent Technologies, Lowers Price Target to $85	2020-05-22	A

Джерело: розроблено автором на основі даних з kaggle.com.

Такий формат дозволяє пов'язати інформаційні повідомлення з конкретними активами та формувати часові ряди, що відображають зміни інформаційного фону для окремих акцій. Заголовки очищуються від технічних символів і повторів, після чого кожне повідомлення подається на вхід мовної моделі для визначення його емоційного забарвлення [7]. Модель повертає значення тональності в діапазоні від -1 до +1, де негативні значення відповідають негативному настрою, позитивні - позитивному, а нуль відображає нейтральний інформаційний контекст. У табл. 2 наведено приклад отриманого показника.

Таблиця 2 – Приклад вихідних даних мовної моделі

Новина	Дата	Індекс компанії	Сентимент
B of A Securities Maintains Neutral on Agilent Technologies, Raises Price Target to $88"	2020-05-22	A	0.106340
CFRA Maintains Hold on Agilent Technologies, Lowers Price Target to $85	2020-05-22	A	0.021060

Джерело: розроблено автором на основі даних з kaggle.com.

На основі отриманих значень формується добовий індекс інформаційного настрою S_t, який обчислюється як середнє значення тональності всіх заголовків за певний день. Оскільки на початковому етапі дослідження всі заголовки вважаються рівнозначними за важливістю, для кожного повідомлення встановлюється вага $\omega_i = 1$. Це дозволяє уникнути упередженості на ранній стадії обробки даних і забезпечує чисту оцінку базового впливу інформаційного фону без введення додаткових параметрів. Використання рівних ваг є також доцільним у випадку, коли заголовки мають подібну структуру і приблизно однакову інформаційну насиченість.

Після агрегації формується часовий ряд, який описує зміну загального тону новин щодо кожної компанії. Такий індекс можна подальше інтегрувати до фінансових даних як додатковий зовнішній фактор, що відображає інформаційний стан ринку.

На рис. 1 подано часову динаміку індексу інформаційного настрою S_t для акцій AAPL, сформованого на основі щоденних значень тональності новин. Синя лінія відображає фактичні значення індексу за кожний торговий день, що дозволяє оцінити короткострокові коливання інформаційного фону. Як видно з графіка, значення S_t демонструють значну варіативність протягом вибраного періоду, що зумовлено різною інтенсивністю новинного потоку та змінністю тону повідомлень, пов'язаних із діяльністю компанії.

Помаранчева пунктирна лінія представляє згладжене значення індексу, отримане за допомогою 7-денного ковзного середнього. Така фільтрація дозволяє виділити середньостроковий тренд інформаційного настрою, зменшивши вплив випадкових коливань і окремих аномальних значень. Порівняння обох кривих демонструє, що незважаючи на локальні коливання, загальний інформаційний фон має схильність до формування плавних фаз зростання або зниження, що може бути ознакою зміни ринкових очікувань та інвестиційної активності.

Таким чином, графік демонструє, що індекс інформаційного настрою є динамічним показником, який чутливо реагує на зміни в новинному просторі та відображає короткоі середньострокові коливання ринкових очікувань, що робить його корисним доповненням до традиційних фінансових індикаторів.

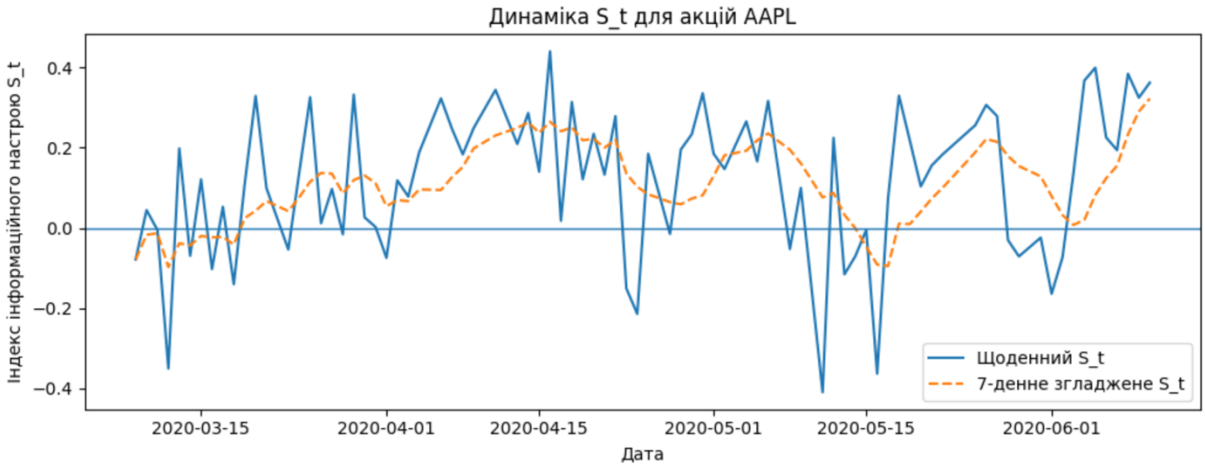


Рис. 1. Графік індексу інформаційного настрою для індексу компанії AAPL

На наступному рис. 2 подано графік порівняння індексу інформаційного настрою S_t із динамікою ціни закриття акцій AAPL за той самий період. Можна побачити взаємозв'язок між інформаційним фоном та фактичними ринковими змінами, оскільки ці дві величини відображають різні аспекти поведінки учасників ринку. Синя лінія демонструє значення індексу S_t, що відображають тональність новинних повідомлень, тоді як червона лінія репрезентує ціну закриття, яка є інтегральним показником ринкових очікувань та реакцій на зовнішні події.

Аналіз показує, що попри значні короткострокові коливання значень S_t, загальна структура кривої має певні елементи кореляції з динамікою ціни закриття. Періоди підвищення індексу позитивного настрою часто супроводжуються плавним зростанням ціни, тоді як зниження інформаційного фону нерідко передує невеликим локальним корекціям. Хоча ці зв'язки не завжди є синхронними, графік демонструє наявність тенденцій, які можуть свідчити про відкладений ефект впливу інформаційного середовища на ринкові рухи.

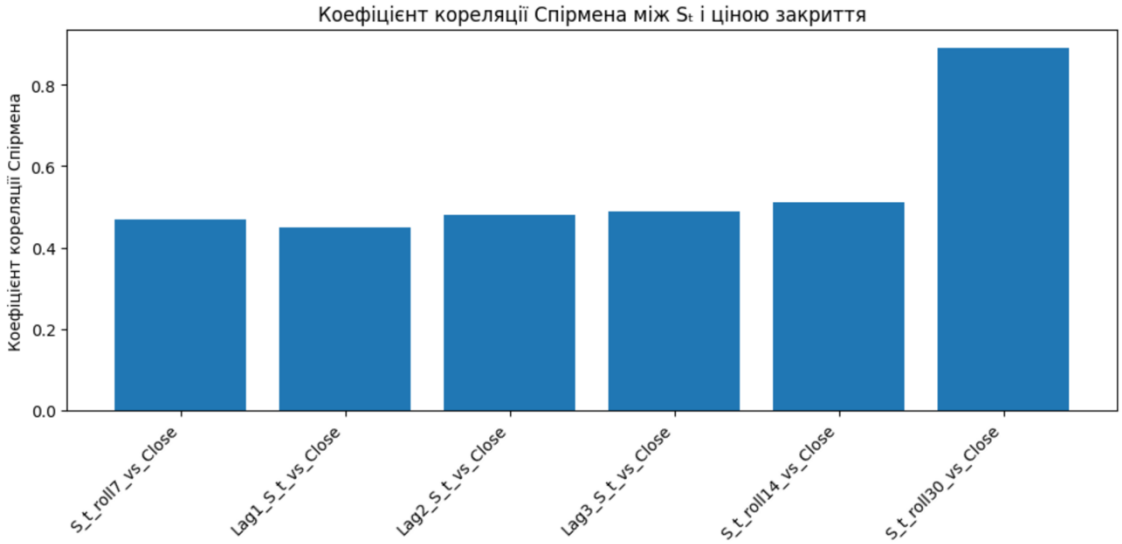
Рис. 2. Графік індексу інформаційного настрою для індексу компанії AAPL та ціни акцій

Особливої уваги заслуговує той факт, що значні коливання S_t, пов'язані з появою окремих новин, що здатні формувати середньострокові очікування на ринку. Такі інформаційні імпульси не завжди одразу відображаються в ціні, однак можуть створювати передумови для подальших трендових змін. Це узгоджується з попередніми спостереженнями щодо необхідності використання ковзних середніх та лагових значень індексу, оскільки ринок часто реагує на інформацію не миттєво, а з певною затримкою. Завдяки поданій візуалізації стає очевидно, що індекс інформаційного настрою може виступати додатковим індикатором для розуміння ринкових процесів.

На наступному рисунку 3 подано порівняльний аналіз коефіцієнтів кореляції Спірмена між різними модифікаціями індексу інформаційного настрою та ціною закриття акцій.

Відповідно до проведених обчислень, короткострокове 7-денне згладжування індексу S_{t_roll7} демонструє помірну кореляцію на рівні близько 0.47. Подібні значення спостерігаються і для лагових індексів, зміщених на один, два та три дні. Це підтверджує наявність стабільного, хоча й не миттєвого зв'язку між інформаційним настроєм та ринковою динамікою. Важливо зазначити, що відносна стабільність кореляцій у межах кількох лагів узгоджується з природою реакції ринку на інформацію: новини не завжди одразу відображаються у ціні, і для формування ринкової відповіді може знадобитися певний час.

Більш виражені залежності спостерігаються при застосуванні довших інтервалів згладжування. Для 14-денного середнього кореляція зростає до значення 0.51, що свідчить про узгодженість між середньостроковим інформаційним тлом та загальним трендом ціни акцій. Найбільш показовою є кореляція між ціною закриття та 30-денним згладженим індексом S_{t_roll30}, яка досягає величини 0.89. Таке значення фактично вказує на дуже високий рівень лінійного зв'язку, що є нетипово сильним для фінансових часових рядів. Це дозволяє стверджувати, що тривалий кумулятивний інформаційний фон має значний вплив на поведінку ринку та може виступати важливим предиктором у моделей прогнозування цінових трендів.




```
=== Коефіцієнт кореляції Спірмена значення ===
S_t_roll7_vs_Close: 0.47
Lag1_S_t_vs_Close: 0.45
Lag2_S_t_vs_Close: 0.48
Lag3_S_t_vs_Close: 0.49
S_t_roll14_vs_Close: 0.51
S_t_roll30_vs_Close: 0.89
```

Рис.3. Розрахунок коефіцієнтів кореляції Спірмена між модифікаціями індексу інформаційного настрою та ціною закриття акцій

Таким чином, аналіз представленого графіка однозначно свідчить про значну інформативність індексу інформаційного настрою у довгостроковому вимірі. Його сильний зв'язок з ціною закриття при збільшенні вікна згладжування вказує на те, що інформаційний фон здатен акумулювати ефект новин і формувати стабільні ринкові тренди. Це робить індекс перспективним інструментом для інтеграції у моделі машинного навчання, де він може використовуватись як предиктор майбутніх цінових рухів, а також як додатковий фактор при дослідженні поведінки інвесторів та динаміки ринкових очікувань.

Висновок. Проведене дослідження демонструє, що застосування сучасних мовних моделей великого масштабу для аналізу фінансових заголовків є ефективним підходом до формування кількісних індикаторів інформаційного середовища. Використання LLM дало можливість отримати числові оцінки тональності новинних повідомлень, які згодом були агреговані у добовий індекс інформаційного настрою S_t. Побудований індекс відображає динаміку інформаційного фону та виявляє чутливість до різких змін у новинному потоці, зберігаючи при цьому чіткі середньострокові тенденції.

Порівняння S_t з ціною закриття показало наявність низки важливих закономірностей. Короткострокові коливання індексу суттєво відрізняються за амплітудою та часто містять високочастотні коливання, спричинені локальними новинними подіями. Проте згладжені значення S_t демонструють стійкі тренди, які співпадають з розвитком ринкової динаміки. Це відображено на графіках, де індекс настрою та ціна закриття змінюються у схожих часових фазах, що свідчить про наявність статистичного зв'язку між інформаційним фоном і поведінкою акцій.

Кореляційний аналіз підтвердив ці спостереження. Коефіцієнти Спірмена для коротких вікон і лагових значень знаходяться на рівні 0.44–0.49, що можна розглядати як помірний прямий зв'язок. Значно сильніша кореляція спостерігається при збільшенні періоду згладжування. Для 14-денного вікна кореляція підвищується до 0.51, а використання

30-денного індикатора демонструє надзвичайно високе значення — 0.89. Така узгодженість свідчить, що довгостроковий інформаційний фон є вагомим носієм ринкових очікувань і може застосовуватися як трендовий індикатор. Натомість кореляції між S_t і короткостроковою дохідністю залишаються низькими, що узгоджується з природою фінансових ринків, де доходність формується під впливом багатьох випадкових факторів.

Отримані результати підтверджують перспективність інтеграції індексу інформаційного настрою до моделей машинного навчання для прогнозування фінансових часових рядів. Завдяки поєднанню семантичних можливостей LLM та методів статистичної обробки даних індекс здатен підсилювати прогнозні моделі, особливо у контексті трендових оцінок. У подальших дослідженнях доцільно розглянути впровадження вагових коефіцієнтів для заголовків новин, використання більш глибоких моделей трансформерної архітектури та інтеграцію альтернативних джерел інформації, що може забезпечити ще точніше моделювання інформаційного впливу на ринок.

Заява про використання генеративного ШІ та технологій на основі ШІ в процесі написання текстів.

Під час написання цього матеріалу автори використовували ChatGPT 5.1 – для виправлення помилок в тексті та пришвидшення написання розділів з поясненням графічних матеріалів.

Після використання цього інструменту автори переглянули та відредагували зміст за потреби і взяли на себе повну відповідальність за зміст публікації.

Список використаних джерел

1. Ouyang, M., Thomas, J. J., Tianzhou, Y., & Fiore, U. (2025). *LLM-guided semantic feature selection for interpretable financial market forecasting in low-resource financial markets*. Discover Computing.

2. Iacovides, G., Konstantinidis, T., Xu, M., & Mandic, D. (2024). *FinLlama: LLM-based financial sentiment analysis for algorithmic trading*. In *Proceedings of the 5th ACM International Conference on AI in Finance (ICAIF'24)*.

3. FinSentLLM: Multi-LLM and structured semantic signals for enhanced financial sentiment forecasting. (2025). *arXiv*. https://arxiv.org/html/2509.12638v1

4. Kim, A., Muhn, M., & Nikolaev, V. (2024). *Financial statement analysis with large language models*. arXiv. https://arxiv.org/html/2407.17866v1.

5. Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). *TweetEval: Unified benchmark and comparative evaluation for tweet classification*. arXiv. https://arxiv.org/abs/2010.12421

6. Aenlle, M. (2023). *Daily Financial News for 6000+ Stocks* [Dataset]. Kaggle. https://www.kaggle.com/datasets/miguelaenlle/massive-stock-news-analysis-db-for-nlpbacktests/data.

7. CardiffNLP. (2024). *twitter-roberta-base-sentiment-latest* [Model]. Hugging Face. https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest.

References

1. Ouyang, M., Thomas, J. J., Tianzhou, Y., & Fiore, U. (2025). *LLM-guided semantic feature selection for interpretable financial market forecasting in low-resource financial markets*. Discover Computing.

2. Iacovides, G., Konstantinidis, T., Xu, M., & Mandic, D. (2024). *FinLlama: LLM-based financial sentiment analysis for algorithmic trading*. In *Proceedings of the 5th ACM International Conference on AI in Finance (ICAIF'24)*.

3. FinSentLLM: Multi-LLM and structured semantic signals for enhanced financial sentiment forecasting. (2025). *arXiv*. https://arxiv.org/html/2509.12638v1

4. Kim, A., Muhn, M., & Nikolaev, V. (2024). *Financial statement analysis with large language models*. arXiv. https://arxiv.org/html/2407.17866v1.

5. Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). *TweetEval: Unified benchmark and comparative evaluation for tweet classification*. arXiv. https://arxiv.org/abs/2010.12421

6. Aenlle, M. (2023). *Daily Financial News for 6000+ Stocks* [Dataset]. Kaggle. https://www.kaggle.com/datasets/miguelaenlle/massive-stock-news-analysis-db-for-nlpbacktests/data.
7. CardiffNLP. (2024). *twitter-roberta-base-sentiment-latest* [Model]. Hugging Face. https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest.

Дата першого надходження статті до видання: 18.11.2025
Дата прийняття статті до друку після рецензування: 01.12.2025

UDC 004.8

Bohdan Parkhomenko[1], Maksym Mishchenko[2]

[1]PhD student of the Department of Information Technology and Software Engineering
Chernihiv Polytechnic National University (Chernihiv, Ukraine)
E-mail: bparkhomenko@stu.cn.ua. **ORCID:** https://orcid.org/0009-0005-1279-4981

[2]PhD Philosophy in Computer Science, Lecturer at the Department of Information Technologies and Software Engineering
Chernihiv Polytechnic National University (Chernihiv, Ukraine)
E-mail: max.mishchenko771@gmail.com. **ORCID:** https://orcid.org/0000-0001-9769-9759

APPLICATION OF LARGE-SCALE LANGUAGE MODELS FOR HEADLINE ANALYSIS AND FORMATION OF ADVANCED CHARACTERS IN FINANCIAL RISKS FORECASTING

The article provides an overview of the information presented based on using of current large-scale models to analyze the headlines of financial news and the formation of an index of information sentiment using the method of reducing the brilliance financial data and improving the accuracy of further forecasting indicators. In the face of growing concerns about unstructured information and the increased dynamics of financial markets, traditional methods of data analysis are often found to be ineffective, which is understandable the need to establish intelligent information models. The addition of deep semantic analysis of text data using machine learning methods will ensure the possibility of expanding sets of characters, enriching input data and increasing the information content of financial time series.

Based on the abstraction of numerical assessments of sentiment, there will be a coefficient of information mood, which aggregates the influx of information background over the previous hourly period and is integrated into the financial series as an additional characteristic. This approach allows us to take into account not only the historical price dynamics, but also the clear aspects of the information environment that influence the behavior of market participants and the formation of market formations. The propionation index can be used as an additional indicator for analyzing trend changes and strengthening forecast models.

The use of proven methods reveals the presence of deposits between the information background and market indicators, which is especially important in the minds of high volatility and insignificance of financial systems. We pay special attention to the ease of interpretation of the results obtained and the stability of the resulting index when changing aggregation parameters. Based on the analysis carried out, the main directions for further improving approaches to the formation of information indices, as well as the possibility of practical stagnation of the withdrawal of results from the activities of financial analysts, are identified and with case-by-case models forecasting financial time series.

Keywords: *language models; machine learning; financial analytics; information models; decision-making.*
Table: 2. Fig.: 3. Bibliography: 7.

Пархоменко Б. М., Міщенко М. В. Застосування мовних моделей великого масштабу для аналізу заголовків та формування розширених ознак у прогнозуванні фінансових ринків. *Технічні науки та технології,* 2025. № 4(42). С. 245-252. DOI: https://doi.org/10.25140/2411-5363-2025-4(42)-245-252.