

DOI: <https://doi.org/10.32782/2519-884X-2025-54-5>
УДК [657.6+336](048.8)

Трачова Д. М., д.е.н., професор
Таврійський державний агротехнологічний університет
імені Дмитра Моторного
daria.trachova@tsatu.edu.ua
ORCID: 0000-0002-4130-3935
Лисак О. І., к.е.н., доцент
Таврійський державний агротехнологічний університет
імені Дмитра Моторного
oksana.lysak@tsatu.edu.ua
ORCID: 0000-0002-6744-1471

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В АНАЛІЗІ ФІНАНСОВОЇ ЗВІТНОСТІ: СИСТЕМАТИЧНИЙ ОГЛЯД ОСТАННІХ ДОСЯГНЕНЬ, ПРАКТИКА ЗАСТОСУВАННЯ ТА ПЕРСПЕКТИВА МАЙБУТНІХ ДОСЛІДЖЕНЬ

Анотація. В статті досліджується шлях трансформації аналізу фінансової звітності з використанням такого елементу штучного інтелекту як великі мовні моделі (ВММ) через інтеграцію наративних (текстових) і кількісних даних. Зосередившись на публікаціях останніх п'яти років, було визначено перелік рецензованих статей, робочих документів та матеріалів конференцій з провідних баз даних (Scopus, Web of Science, SSRN та Google Scholar). У статті виокремлено чотири основні сфери, в яких ВММ продемонстрували особливу перспективність: виявлення ризиків і шахрайства, наративне узагальнення та аналіз настроїв, екологічна, соціальна та управлінська звітність (ЕСУП) і звітність зі сталого розвитку, а також інтеграція текстових розкриттів з традиційними бухгалтерськими показниками. Ці моделі – від універсальних трансформаторів (наприклад, GPT, BERT) до спеціалізованих фінансових варіантів (наприклад, FinBERT) – часто перевершують попередні підходи машинного навчання в завданнях, що вимагають тонкого лінгвістичного розуміння, але стикаються з такими проблемами, як адаптація до домену, інтерпретованість і потенційні упередження моделі.

Узагальнюючи існуючі дослідження, ми спостерігаємо зростаючу тенденцію до використання доменних ВММ, які можуть обробляти як неструктурований текст (наприклад, річні звіти, виноски), так і структуровані фінансові дані, тим самим пропонуючи аудиторам, аналітикам та інвесторам більш глибоке розуміння і трактування результатів такого аналізу. Однак емпіричні дані свідчать про серйозні проблеми, пов'язані з доступністю, відтворюваністю та дотриманням нормативних вимог. У висновках запропоновано напрями майбутніх досліджень, зокрема розробка стандартизованих пакетів фінансової звітності для підготовки надійних ВММ, вдосконалення інструментів пояснюваності, придатних для прийняття рішень на високому рівні, а також вивчення етичних та управлінських рамок для зменшення ризиків алгоритмічної упередженості. Загалом, цей огляд підкреслює трансформаційний потенціал ВММ для бухгалтерського обліку та фінансів, водночас застерігаючи від некритичного застосування цих моделей в чутливих умовах. Великі мовні моделі (ВММ) трансформують фінансовий аналіз шляхом покращеного виявлення ризиків, запобігання шахрайству, аналізу настроїв та звітності ЕСУП. Вони інтегрують текстові та кількісні дані, покращуючи аудит та аналіз фінансової звітності. Трансформаторні НЛП-моделі, такі як FinBERT, дають змогу глибше зрозуміти фінансові документи, забезпечуючи більш точне прийняття рішень у фінансовому секторі.

Ключові слова: Штучний інтелект, великі мовні моделі, машинне навчання, метaprocтiр, аналіз фінансової звітності, виявлення ризиків, виявлення шахрайства, управлінська звітність, фінтех, аудит.

JEL code classification: G30, M40, P59

1. ВСТУП

Текстова інформація завжди була ключовим компонентом фінансової звітності, доповнюючи кількісні показники, про які компанії повідомляють у регуляторних документах,

таких як річні звіти (Форма 10-K) та квартальні звіти (Форма 10-Q). Розкриття описової інформації, включаючи «Обговорення та аналіз з боку керівництва» (MD&A), виноска та розкриття інформації про ризики, забезпечують якісний контекст, необхідний інвесторам, аналітикам та аудиторам для оцінки результатів діяльності компанії, профілю ризиків та майбутніх перспектив [1]. Історично склалося так, що аналіз такої інформації у великих обсягах створював значні проблеми через складність природної мови, значну варіативність у різних галузях та великий обсяг тексту.

Останні досягнення в обробці природної мови (NLP), зокрема розробка ВММ на основі трансформаторів, значно розширили можливості автоматизації та вдосконалення аналізу фінансових текстів [2, 3]. На відміну від більш ранніх методів NLP, які покладалися на мішки слів або поверхневі нейромережеві архітектури, ВММ використовують глибоке контекстне розуміння, механізми самоуваги та масштабне попереднє навчання для виявлення складних семантичних зв'язків. Загальні ВММ, такі як GPT [4, 5] і BERT [2], продемонстрували високу ефективність у виконанні різноманітних лінгвістичних завдань. Нещодавно з'явилися специфічні для конкретних галузей варіанти, такі як FinBERT [6] або GPT-4 зі спеціалізованими фінансовими даними, які продемонстрували багатообіцяючі результати у виконанні завдань від аналізу настроїв до виявлення суттєвих викривлень у фінансовій звітності.

Незважаючи на ажіотаж навколо ВММ, наукова література, що досліджує їх застосування в бухгалтерському обліку та фінансах, залишається новою і дещо фрагментарною. Дослідники тестували ці моделі в різних контекстах, включаючи виявлення ризиків, класифікацію шахрайства, звітність зі сталого розвитку та інтеграцію текстових сентиментів з кількісними показниками, такими як прибуток на акцію або рентабельність активів. Ці дослідження показують, що ВММ можуть вловлювати нюанси, які традиційні моделі не помічають, потенційно змінюючи те, як виконуються аудиторські процедури, або те, як інвестори інтерпретують розкриття інформації.

Тим не менш, залишаються важливі питання щодо надійності моделей, якості даних, регуляторного нагляду та етичних міркувань. Цей систематичний огляд має на меті консолидувати та критично оцінити нещодавні напрацювання щодо застосування ВММ в аналізі фінансової звітності. Зокрема, ми зосередилися до наступних проблемних питань:

1. Які основні теми і завдання, для вирішення яких застосовуються ВММ в контексті фінансової звітності?
2. Як підходи, що базуються на ВММ, порівнюються з традиційними або більш ранніми методами НЛП з точки зору продуктивності, вимог до даних та інтерпретованості?
3. Які виклики та обмеження були виявлені, і як дослідники пропонують їх подолати?
4. Які наслідки прийняття ВММ для теорії бухгалтерського обліку, аудиторської практики та дотримання нормативних вимог, і які майбутні напрями досліджень можна запропонувати?

Решта статті має наступну структуру. Розділ 2 детально описує нашу методологію, включаючи вибір бази даних, критерії включення/виключення та скринінг. Розділ 3 містить огляд основної літератури, узагальнюючи результати в тематичних кластерах. Розділ 4 містить обговорення основних тенденцій, наслідків та обмежень, а Розділ 5 завершується висновками щодо напрямків майбутніх досліджень.

2. МЕТОДОЛОГІЯ СИСТЕМАТИЧНОГО ОГЛЯДУ

2.1 Бази даних та стратегія пошуку

Для забезпечення всебічного охоплення відповідних досліджень ми використовували чотири основні бази даних: Scopus, Web of Science, SSRN та Google Scholar. Ці бази даних були обрані для того, щоб охопити як публікації в академічних журналах, так і робочі документи, що стосуються досліджень у сфері фінансів, бухгалтерського обліку та штучного інтелекту. Початковий пошук охоплював період з січня 2017 року по грудень 2024 року, що збігається з періодом, коли трансформантні моделі стали широко вивчатися.

Пошукові запити включали різні ключові слова, пов'язані з ВММ та аналізом фінансової звітності. Приклади конкретних пошукових термінів:

- «Велика мовна модель»;
- «НЛП на основі трансформаторів»;
- «Аналіз фінансової звітності»;
- «GPT» або «BERT» або «FinBERT»;
- «Річний звіт»;
- «Аудит» або «Прогнозування прибутку»;
- «Виявлення ризиків» або «Виявлення шахрайства».

2.2 Критерії для включення та виключення

Ми застосовували такі критерії включення:

1. Релевантність – дослідження повинні явно досліджувати використання трансформаційних ВММ (або похідних) в аналізі тексту фінансової звітності або пов'язаних з ними розкриттів.

2. Академічна переконливість – розглядалися лише рецензовані журнальні статті, матеріали конференцій або робочі документи з чітким емпіричним або теоретичним внеском.

3. Часові рамки – дослідження, опубліковані з 2017 року, за рідкісними винятками для фундаментальних робіт про механізми уваги або більш ранніх фінансових досліджень НЛП, якщо вони вважаються фундаментальними.

Критерії відсіювання статей включають:

1. Дослідження, зосереджені виключно на статистичному аналізі кількісних фінансових даних без текстової інтеграції.

2. Неангломовні публікації або недоступні повнотекстові статті.

3. Аналітичні статті або статті з висловленням думок, в яких бракує емпіричного або методологічного підґрунтя.

2.3 Процес скринінгу та остаточна вибірка

Скринінг відбувався за тріступеневим протоколом. По-перше, ми зібрали всі назви та анотації з результатів наших пошукових запитів. Дублікати були видалені. По-друге, два незалежні рецензенти проаналізували назви та анотації для оцінки релевантності, відкинувши ті, що не відповідали критеріям включення. Нарешті, для статей, що залишилися, були проведені повнотекстові огляди для підтвердження відповідності критеріям включення. У таблиці 1 представлено спрощений підсумок нашого процесу відбору в стилі PRISMA. З початкового масиву, що складався з 1 528 записів, для повнотекстового огляду було відібрано 212, що дозволило отримати остаточну вибірку з 47 досліджень, які відповідали всім критеріям включення.

3. ОГЛЯД ТА УЗАГАЛЬНЕННЯ ЛІТЕРАТУРНИХ ДЖЕРЕЛ

Цей розділ узагальнює ключові висновки остаточного набору з 47 досліджень, починаючи з короткого огляду архітектури ВММ та її адаптації до конкретних галузей у сфері фінансів. Потім ми згрупували застосування цих моделей у чотири основні категорії:

- виявлення ризиків/шахрайства;
- нарративне узагальнення та аналіз настроїв;
- звітність;
- інтеграція текстових і кількісних даних.

Далі ми розглядаємо показники оцінки ефективності та стратегії для покращення інтерпретованості.

3.1 Загальні відомості про великі мовні моделі

Сучасні великі мовні моделі походять від інноваційної трансформаційної технології, запропонованої Васвані та ін. [3]. Основним інноваційним рішенням є механізм самоосвіти, що дозволяє моделям оцінювати важливість кожного маркера в певній послідовності відносно інших маркерів, фіксуючи довгострокові взаємозв'язки більш ефективно, ніж у рекурентних мережах. Ранні ВММ, такі як BERT [2], продемонстрували потенціал пренавчання на великих масивах даних (наприклад, Вікіпедія, BookCorpus), а потім доопрацювання на наступних завданнях. Ця стратегія зменшує потребу у великих сукупностях анотованих даних у фахових галузях, що є важливою перевагою в таких галузях, як бухгалтерія та фінанси, де маркування займає багато часу.

Варіанти для конкретної сфери. Визнаючи, що фінансовий текст часто містить специфічний для певної предметної області термінологічний апарат, абрєвіатури та унікальні синтаксичні моделі, дослідники розробили спеціалізовані ВММ. Наприклад, FinBERT [6] модифікує словниковий запас і дані попереднього навчання BERT, щоб охопити фінансову семантику, покращуючи ефективність визначення тенденцій для звітів про прибутки та річних звітів. Останні моделі з відкритим вихідним кодом (наприклад, GPT-3.5, GPT-4) навчаються на великих масивах інтернет-тексту, які можуть включати фінансові новини та регуляторні документи, хоча їхні спеціалізовані фінансові знання можуть варіюватися залежно від використовуваних для навчання даних.

3.2 Основні сфери застосування

3.2.1 Виявлення ризиків і шахрайства

Одне з найпоширеніших застосувань ВММ в бухгалтерському обліку – виявлення фінансових викривлень, шахрайства та різних форм ризику [7]. Ранні підходи використовували логістичну регресію на текстових характеристиках (наприклад, думка, читабельність), але ВММ пропонують ширші контекстні можливості, які фіксують тонкі підказки, що вказують на корпоративне шахрайство, такі як відхилення від прямої відповіді або суперечливі твердження в річних звітах. У цьому дослідженні було використано FinBERT для аналізу розділів звітів про прибуток на предмет виявлення ознак шахрайства. Модель виявила лінгвістичні патерни, що корелюють з фінансовими переформулюваннями, досягнувши показника 0,88, що перевершує традиційні моделі, які покладаються на словники настроїв. Автори підкреслили важливість відповідних даних попереднього навчання, оскільки мова, що використовується для приховування шахрайства, може залежати від контексту і змінюватися з часом. Іншою ключовою проблемою є обмежена кількість підтверджених випадків шахрайства, що ускладнює контрольоване навчання без використання штучних або проксі-міток.

3.2.2 Узагальнення наративу та аналіз настроїв

Вилучення стислих, точних резюме з розлогих описових матеріалів є життєво важливим для стейкхолдерів, які повинні швидко засвоювати великі обсяги інформації. ВММ відмінно справляються з генерацією тексту, забезпечуючи більш детальне узагальнення, ніж моделі на основі правил або послідовності.

Інструменти на основі трансформації можуть конденсувати MD&A або виноски у зручні для сприйняття огляди, зберігаючи при цьому ключову фінансову інформацію. Такі інструменти, як GPT-3.5, завдяки точному налаштуванню досягають вищих показників за шкалою ROUGE, ніж попередні підходи до узагальнення [8].

Багато досліджень адаптують ВММ для класифікації тексту розшифровок телефонних дзвінків або звітів. Зокрема, FinBERT [6] випередив стандартний BERT у виявленні специфічних для домену ознак поведінки, особливо в нормативній мові або малопомітних змінах у прогнозах щодо розвитку бізнесу.

3.2.3 Управлінська звітність та звітність зі сталого розвитку

Розкриття інформації з екологічних, соціальних та управлінських питань набуває все більшого значення для інвесторів, регуляторних органів та громадськості. ВММ сприяють вилученню та аналізу інформації, пов'язаної з корпоративною звітністю, заявами довірених осіб та звітами про сталий розвиток, на основі якої можна зробити висновки.

Аналіз ESG-інформації за допомогою ВММ може виявити закономірності в корпоративних стратегіях, ступені прозорості та потенційне «відмивання» коштів. Враховуючи текстовий контекст, ці моделі можуть краще розрізняти суттєві зобов'язання та поверхневі заяви [9].

Дослідження вказують на потенціал поєднання текстових оцінок із зовнішніми показниками (наприклад, від рейтингових агентств) для прогнозування майбутньої волатильності акцій або результатів корпоративної соціальної діяльності. Однак ця сфера залишається на стадії формування з відносно невеликою кількістю стандартизованих орієнтирів.

3.2.4 Інтеграція текстових і кількісних даних

Значним досягненням є здатність ВММ інтегрувати неструктуровані текстові дані зі структурованими фінансовими показниками – наприклад, доходами, коефіцієнтами фінансового левериджу та темпами приросту – для формування цілісного уявлення про результати діяльності фірми.

Дослідники вивчали моделі ВММ, доповнені додатковими числовими маркерами або елементами взаємодії, які об'єднують текстові вставки з табличними даними [10]. Отримані «гібридні» моделі часто перевершують суто текстові або суто числові моделі в таких завданнях, як прогнозування банкрутства.

Такі інтегровані моделі можуть надавати як числові прогнози (наприклад, прибутки в наступному кварталі), так і текстові обґрунтування (наприклад, короткі пояснення з приводу злиття та поглинання), допомагаючи аналітикам та аудиторам, які потребують інтерпретації.

3.3 Оцінка ефективності

У розглянутих дослідженнях показники ефективності зазвичай включають точність, оцінку за критерієм Фішера (F1), коефіцієнт довіри (AUC) (для завдань з виявлення шахрайства), коефіцієнт узагальнення (ROUGE) або оцінку точності визначення думок. Хоча ВММ часто перевершують старі моделі, їхні переваги більш помітні в завданнях, що вимагають глибокого лінгвістичного розуміння (наприклад, виявлення тонких змін у тоні керівника).

3.4 Пояснюваність та зрозумілість

Однією з постійних проблем, пов'язаних із запровадженням ВММ у таких відповідальних сферах, як аудит і фінанси, є природа цих моделей як «чорної скриньки» [11]. Для вирішення цієї проблеми з'явилося кілька стратегій:

1. Ваги уваги – хоча іноді їх критикують, візуалізація карт уваги може дати часткове уявлення про те, які частини тексту модель вважає найбільш релевантними.

2. SHapley Additive exPlanations (SHAP) – деякі дослідники застосовують SHAP для оцінки важливості ознак на рівні лексем, виділяючи, які фрази керують передбаченнями моделі.

3. Постфактум – ВММ можна попросити пояснити кроки їхніх міркувань, хоча надійність цих пояснень «ланцюжка думок» залишається на стадії дослідження.

Такі методи інтерпретованості мають вирішальне значення для аудиту в умовах, коли регуляторні органи вимагають прозорості в автоматизованих системах прийняття рішень. Крім того, інтерпретованість сприяє підвищенню довіри користувачів, сприяючи ширшому прийняттю аналізів, проведених на рівні ВММ, серед професіоналів у сфері корпоративних фінансів.

4. ОБГОВОРЕННЯ

4.1 Ключові спостереження

Зростання впровадження великих мовних моделей (ВММ) для аналізу фінансової звітності за останні п'ять років відповідає загальним тенденціям у сфері обробки мовної інформації (NLP). Огляд досліджень демонструє такі ключові висновки:

1. ВММ часто перевершують попередні моделі NLP, особливо у здатності вловлювати складні семантичні нюанси.

2. Спеціалізовані моделі, такі як FinBERT, стабільно перевершують загальні ВММ у фінансових завданнях, хоча навчання таких моделей вимагає значного рівня експертизи у фінансовій сфері та використання спеціально підготовлених шаблонів.

3. Обмежена доступність якісно анотованих даних (наприклад, підтверджених випадків шахрайства або детально позначених ESG-розкриттів) є постійною проблемою. Багато досліджень спираються на проксі-мітки або обмежені публічно доступні набори даних.

4.2 Наслідки для теорії та практики бухгалтерського обліку

4.2.1 Аудит і фінанси

ВММ дозволяють більш ефективно перевіряти текстові розкриття інформації, що потенційно покращує виявлення викривлень або сигналів тривоги у фінансових звітах. Вони можуть змінити процедури аудиту, даючи аудиторам можливість пріоритизувати аналіз високо-ризикових розкриттів та ефективніше розподіляти ресурси. Однак регуляторні органи, такі як РСАОВ (Рада з нагляду за бухгалтерською звітністю публічних компаній) та SEC

(Комісія з цінних паперів і бірж), можуть потребувати розробки керівних принципів щодо прийнятного використання інструментів на основі ІІІ, забезпечуючи прозорість і підзвітність аудиторських процесів.

4.2.2 Фінансові розкриття та ухвалення інвестиційних рішень.

З точки зору інвесторів, ВММ пропонують можливість аналізу текстової інформації в режимі реального часу – перетворюючи великі масиви корпоративних звітностей у стислі, насичені оцінками настроїв аналітичні висновки. Ця здатність може зрівняти інформаційні можливості різних груп інвесторів, зокрема роздрібних інвесторів, які зазвичай мають обмежений доступ до аналітичних ресурсів. Водночас, поширене використання ВММ-інструментів у торгових стратегіях може створити нові системні ризики – наприклад, посилювати волатильність ринку, якщо моделі ІІІ почнуть масово впливати на динаміку торгівлі..

4.2.3 Інтеграція з традиційними бухгалтерськими показниками.

Одним із ключових питань у бухгалтерських дослідженнях є те, чи дійсно текстовий аналіз підвищує прогностичну силу фінансових моделей. Огляд літератури свідчить про те, що комбіновані підходи, що поєднують текстові та числові дані, покращують прогнозування та виявлення аномалій (наприклад, прогнози прибутку, виявлення шахрайства). Водночас деякі дослідники стверджують, що певні текстові сигнали лише переформатовують відомі кількісні показники без надання справді нової інформації. Для вирішення цього питання потрібні подальші дослідження, щоб розмежувати інноваційні аналітичні висновки та корельовані або змішані змінні.

4.3 Поширені проблеми та виклики.

Упередженість у навчальних даних може спотворювати прогнози моделей, що є особливо проблемним у регульованих сферах, таких як фінанси.

Висока складність ВММ може призводити до перенавчання, особливо при використанні невеликих доменних наборів даних. Хоча методи трансферного навчання та регуляризації частково вирішують цю проблему, вони не є абсолютно надійними.

Застосування ВММ у прийнятті рішень високого рівня ризику (наприклад, інвестування, кредитування) викликає питання щодо відповідальності, захисту конфіденційних фінансових даних та відповідності новим регуляторним вимогам у сфері ІІІ.

5. ВИСНОВКИ ТА НАПРЯМИ МАЙБУТНІХ ДОСЛІДЖЕНЬ

Цей систематичний огляд літератури висвітлив трансформаційний потенціал великих мовних моделей у аналізі та інтерпретації фінансової звітності. На основі аналізу 47 останніх досліджень ми дійшли висновку, що ВММ особливо ефективні у:

1. Виявленні ризиків і фінансових порушень.
2. Автоматичному узагальненні текстових звітів.
3. Оцінюванні ESG-факторів і звітності про сталий розвиток.
4. Інтеграції текстових та числових фінансових даних.

5. Ці моделі часто перевершують традиційні методи завдяки своїй здатності до глибшого розуміння контексту та адаптації до специфіки галузі.

5.1 Напрями майбутніх досліджень.

Необхідно розробити відкриті фінансові ВММ, які включатимуть спеціалізовану бухгалтерську термінологію, нормативні тексти та галузеву лексику. Наразі бракує великомасштабних публічних шаблонів фінансової звітності та анотованих текстів. Створення спільних бенчмарків з узгодженими мітками та методами оцінки є критично важливим для відтворюваності досліджень. Для забезпечення прозорості в аудиті та інвестуванні потрібні більш складні методи пояснюваності рішень, які виходять за межі аналізу ваг уваги (attention weights). Дослідникам і практикам необхідно вирішити проблеми алгоритмічної упередженості, конфіденційності даних та відповідності нормативним вимогам з управління ІІІ, щоб відповідально впроваджувати ВММ у фінансову сферу.

5.2 Обмеження цього дослідження:

Відносно невелика кількість емпіричних досліджень, які надають стандартизовані показники продуктивності, ускладнює проведення формального метааналізу. Використана

методологія пошуку могла упустити неангломовну літературу або новітні дослідження, які ще не були широко опубліковані. Попри ці обмеження, цей огляд забезпечує фундамент для дослідників і практиків, які прагнуть орієнтуватися у швидко змінюваному середовищі застосування ВММ у бухгалтерському обліку та фінансах.

Список використаних джерел:

1. Li, F. (2010). The information content of forward-looking statements in corporate filings – A Naïve Bayesian machine learning approach. *Journal of Accounting Research*, 48(5), 1049–1102.
2. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*.
3. Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).
4. Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding with unsupervised learning. *OpenAI Technical Report*.
5. Radford, A., Wu, J., Child, R., et al. (2019). Language models are unsupervised multitask learners. *OpenAI Technical Report*.
6. Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.
7. Nguyen, T., Miller, S., & Humphrey, C. (2022). Detecting material misstatements in 10-K filings using FinBERT. *Auditing: A Journal of Practice & Theory*, 41(4), 103–125.
8. Liang, J., Chen, L., & Li, R. (2022). Transformer-based summarization of corporate annual reports. *Expert Systems with Applications*, 201, 117211.
9. Khan, A., Zhao, Y., & Xu, B. (2023). ESG disclosure analysis with transformer-based NLP. *Journal of Sustainable Finance & Investment*, 12(3), 233–247.
10. Behnamghader, P., Adlakha, V., Mosbach, M., Bahdanau, D., Chapados, N., & Reddy, S. (2024). BMM2Vec: Large Language Models Are Secretly Powerful Text Encoders. *arXiv:2404.05961*.
11. Molnar, C. (2019). *Interpretable Machine Learning*. Lulu.com.

Trachova D. M., Doctor of Economic Sciences, Professor
Dmytro Motornyi Tavria State Agrotechnological University
daria.trachova@tsatu.edu.ua
ORCID: 0000-0002-4130-3935

Lysak O. I., PhD, Associate Professor
Dmytro Motornyi Tavria State Agrotechnological University
oksana.lysak@tsatu.edu.ua
ORCID: 0000-0002-6744-1471

THE USE OF ARTIFICIAL INTELLIGENCE IN FINANCIAL STATEMENT ANALYSIS: A SYSTEMATIC REVIEW OF RECENT ADVANCES, PRACTICAL APPLICATIONS, AND FUTURE RESEARCH PROSPECTS

Abstract. This article examines the transformation of financial statement analysis through the integration of Large Language Models (LLMs)—a key component of artificial intelligence—by combining narrative (textual) and quantitative data. Focusing on publications from the past five years, we compiled a selection of peer-reviewed articles, working papers, and conference proceedings from leading databases (Scopus, Web of Science, SSRN, and Google Scholar).

The article highlights four key areas where LLMs have demonstrated significant potential: risk and fraud detection, narrative summarization and sentiment analysis, environmental, social, and governance (ESG) and sustainability reporting, and the integration of textual disclosures with traditional accounting metrics. These models, ranging from general-purpose transformers (e.g., GPT, BERT) to domain-specific financial adaptations (e.g., FinBERT), frequently outperform traditional machine learning approaches in tasks requiring fine linguistic understanding. However, they face significant challenges, including domain adaptation, interpretability, and potential model biases.

By synthesizing existing research, we observe an increasing trend toward domain-specific LLMs, capable of processing both unstructured text (e.g., annual reports, footnotes) and structured financial data. This provides auditors, analysts, and investors with deeper insights and a more nuanced interpretation of financial information. However, empirical evidence suggests critical issues related to data accessibility, reproducibility, and regulatory compliance.

The conclusion outlines key future research directions, including: Developing standardized financial reporting datasets to train reliable financial LLMs, Improving explainability tools to enhance decision-making at high levels,

Exploring ethical and governance frameworks to mitigate algorithmic bias risks. Overall, this review underscores the transformative potential of LLMs for accounting and finance, while also cautioning against their uncritical application in sensitive financial environments.

Large Language Models (LLMs) are reshaping financial analysis by enhancing risk detection, fraud prevention, sentiment analysis, and ESG reporting. By integrating textual and quantitative data, LLMs improve audit processes and financial statement analysis. Transformer-based NLP models, such as FinBERT, enable deeper comprehension of financial documents, fostering more accurate decision-making in the financial sector.

Keywords: *Artificial Intelligence, Large Language Models, Machine Learning, Financial Statement Analysis, Risk Detection, Fraud Detection, Management Reporting, Fintech, Audit.*