

Формування когнітивних упереджень на фінансовому ринку в умовах розвитку штучного інтелекту

У статті досліджено трансформацію поведінкових патернів учасників фінансового ринку під впливом стрімкого впровадження штучного інтелекту (ШІ). Зазначено, що традиційні моделі фінансової поведінки, які ґрунтувалися на обмеженій раціональності, зазнають суттєвих змін. ШІ не лише реагує на ринкові стимули, а й самостійно формує їх, створюючи нові цифрові патерни поведінки. Особлива увага приділена когнітивним упередженням, що виникають як у користувачів, так і в алгоритмах ШІ, враховуючи автоматизаційне упередження, надмірну впевненість, упередження доступності та алгоритмічні викривлення. Розглянуто механізми взаємодії в системі «людина – ШІ», де визначено ролі, канали та механізми передачі упереджень від людини до алгоритмів. Запропоновано класифікацію учасників системи «людина – ШІ» (користувач, розробник, анотатор, експерт, регулятор тощо) та з'ясовано, як на кожному етапі може виникати або посилюватись упередженість. Описано «ефект снігової кулі», коли навіть незначні когнітивні похибки користувачів з часом масштабуються завдяки особливостям роботи ШІ. У дослідженні представлено інструментарій для контролю когнітивних упереджень на алгоритмічному та особистісному рівнях. Підкреслюється, що недооцінка психологічного впливу ШІ на користувачів може призводити до ризикованих рішень і зниження фінансової стійкості. Авторка наголошує на необхідності інтеграції соціотехнічного підходу до вивчення впливу ШІ на фінансову поведінку, що враховує як технічні, так і когнітивні та соціальні чинники. Обґрунтовано потребу в розробці комплексних стратегій моніторингу, контролю та етичного регулювання ШІ у фінансовій сфері. Перспективи подальших досліджень полягають у глибшому аналізі психологічних механізмів формування «цифрових» поведінкових патернів та створенні нових методик зниження впливу когнітивних викривлень.

Ключові слова: поведінка; когнітивні упередження; штучний інтелект; фінансові рішення; автоматизація; алгоритмічна упередженість; інвестиційна поведінка

Актуальність теми. Сучасний фінансовий ринок в умовах розвитку та активного використання ШІ переживає докорінні зміни у механізмах функціонування, взаємодії суб'єктів та, що найважливіше, у формуванні поведінкових патернів його учасників.

Аналітики спостерігають експоненційне зростання обсягів використання ШІ на фінансовому ринку. За даними Deloitte, у 2024 році понад 85 % фінансових установ світу вже інтегрували або планують інтегрувати рішення на базі ШІ у свою діяльність, що на 25 % більше, ніж у 2022 році [1]. PwC прогнозує [2], що до 2030 року ШІ може додати до світового ВВП \$15,7 трильйона, значна частина яких припадатиме на фінансовий сектор. Це свідчить про глибоке проникнення ШІ у всі сфери фінансових операцій – від автоматизованої торгівлі та управління портфелями до оцінки ризиків та взаємодії з клієнтами. Впровадження алгоритмів машинного навчання, нейронних мереж та BigData дозволяє здійснювати операції з безпрецедентною швидкістю та точністю, що, безумовно, впливає на динаміку ринку та, як наслідок, на поведінку його учасників. Крім того, зміна структури ринкових взаємодій свідчить про те, що традиційні моделі поведінки, що ґрунтувалися на ірраціональності або обмеженій раціональності людського фактора, вимагають переосмислення. Зростаюча частка алгоритмічної торгівлі, яка, за деякими оцінками, становить до 70–80 % обсягу торгів на певних ринках, зокрема на валютних та високочастотних, нівелює багато традиційних психологічних аспектів. Однак це не означає зникнення поведінкових патернів, а їх трансформацію. Тепер ШІ не лише реагує на ринкові стимули, але й сам стає потужним джерелом формування цих стимулів, створюючи нові «цифрові» поведінкові патерни, що можуть відрізнятися від людських аналогів і мати непередбачувані наслідки. Наприклад, ефект Flash Crash, що спостерігався на фондових ринках, часто пов'язують з каскадними діями алгоритмічних систем, що взаємодіють між собою без людського втручання. Це свідчить про необхідність вивчення не лише поведінки людини на ринку, але й поведінки самих алгоритмів, їх взаємодії та впливу на загальну динаміку.

Слід враховувати, що ШІ надає унікальні інструменти для збору, обробки та аналізу величезних обсягів фінансових даних, що відкриває безпрецедентні можливості для виявлення прихованих кореляцій, аномалій та нових поведінкових патернів, які раніше були недоступними для людського аналізу. Дослідження поведінкових патернів з використанням ШІ може виявити закономірності, які дозволять не лише краще зрозуміти поточні процеси, а й розробити більш точні моделі прогнозування, підвищити ефективність інвестиційних стратегій та управління ризиками. Однак, як зазначає Міжнародний валютний

фонд [3], не слід недооцінювати можливі наслідки штучного інтелекту для фінансової стабільності, серед яких установа виділяє такі, як ризик концентрації ринку, надмірна швидкість коригування ринку, алгоритмічні маніпуляції та упередженість, неефективність. При цьому залежність від зовнішніх постачальників моделей штучного інтелекту / машинного навчання та вхідних даних, що збільшує ризик операційних збоїв.

Зважаючи на масштаби впровадження та використання ШІ на фінансовому ринку, варто враховувати, що застосування ШІ в персональному фінансовому плануванні, робо-едвайзингу та автоматизованих інвестиційних платформах не просто впливає на результати прийняття фінансових рішень учасниками ринку, а саме змінює способи прийняття рішень мільйонами роздрібних інвесторів. Згідно з даними Allied Market Research [4], глобальний ринок робо-консультаційних послуг у 2022 році оцінювався в 7,9 млрд доларів, і, за прогнозами, до 2032 року досягне 129,5 млрд доларів, із середньорічним темпом зростання 32,5 % з 2023 по 2032 рік. Відповідно все більше інвесторів будуть покладатися на рекомендації, сформовані алгоритмами, що може призвести до уніфікації інвестиційних стратегій та нових форм «стадної» поведінки, опосередкованої технологіями. Для корпорацій ШІ стає незамінним інструментом для оцінки ризиків, оптимізації капіталу та розробки конкурентних стратегій. Розуміння того, як ці зміни впливають на формування поведінкових патернів, є ключовим для всіх учасників ринку.

Таким чином, постає необхідність глибшого розуміння складних динамічних процесів сучасного фінансового ринку та розробки науково обґрунтованих рекомендацій для всіх його учасників – від індивідуальних інвесторів та корпорацій до регуляторних органів для забезпечення стабільності, ефективності та справедливості світової фінансової системи.

Аналіз останніх досліджень та публікацій, на які спирається автор. Питання впливу ШІ на різні сфери людського життя, у т. ч. економічної активності, знаходяться під пильною увагою науковців, дослідників та практиків. Окремою сферою наукових інтересів дослідників є визначення характеру, сили та наслідків впливу застосування технологій ШІ в сфері прийняття фінансових рішень. У роботах таких вчених, як Д.Канеман, А.Тверські, Р.Талер, Р.Шиллер, Г.Саймон та ін. розглядаються базові аспекти ірраціональної поведінки учасників економічних процесів, зокрема фінансового ринку. Однак питання впливу ШІ на поведінкові аспекти прийняття рішень набуло актуальності серед науковців саме в останні роки. Так Dimple Patil [5] в своєму дослідженні зазначає, що в процесі оцінки фінансових ризиків задля виявлення потенційно шахрайських операцій системи на базі ШІ можуть значно підвищити ефективність та швидкість виявлення потенційно небезпечних операцій, використовуючи алгоритми машинного навчання та аналітику BigData та виявляючи певні поведінкові закономірності та аномалії у поведінці клієнта або груп клієнтів, що вказують на відповідні ризики або шахрайство. Під відповідний аналіз підпадають поведінка витрат клієнта, історія його транзакцій, соціальні дані тощо, що у сукупності дозволяє сформувати відповідний профіль клієнта з точки зору типової та нетипової поведінки, ознак ризикованості операцій, потенційного шахрайства. Більш того, висока адаптивність ШІ та здатність до швидкого навчання дозволяє також швидко реагувати на нові схеми шахрайства, знижуючи ризикованість фінансових операцій як для клієнтів, так і для фінансових установ. Однак, як зазначає автор [5], не слід ідеалізувати використання ШІ, оскільки багато ризиків виникає саме внаслідок використання алгоритмів ШІ, а саме: загроза конфіденційності даних клієнта, упередженість результатів навчання, якщо воно відбувається на основі нерепрезентативних або спотворених даних.

Аналогічні занепокоєння щодо достовірності та етичності використання ШІ висловлюються в публікації [6], в якій авторами обґрунтовується думка щодо можливих ризиків використання ШІ під час прийняття фінансових рішень. Авторами вирізняється три основні проблеми використання ШІ в фінансах, такі як, етичні проблеми щодо можливості потенційного впливу на поведінку інвесторів, можливості порушення конфіденційності даних; труднощі для регуляторних органів, які мають швидко адаптувати механізми та інструменти контролю використання ШІ у фінансах; системні ризики, у т. ч. інформаційні каскади, які можуть збільшувати волатильність ринку та призводити до непередбачуваних наслідків використання алгоритмів ШІ у фінансах.

Фокус уваги дослідників [7] спрямовано на два «критичні» ризики, пов'язані з інтенсифікацією використання ШІ у фінансовому секторі, зокрема:

- зростаюча залежність від технологій, яка може призводити до значних фінансових втрат внаслідок таких подій, як системні збої, витоки даних, шахрайські технологічні дії тощо;
- потенціал для маніпулювання ринком внаслідок свідомих та несвідомих дій як людей, так і самих алгоритмів ШІ, які можуть базуватися на хибній, спотвореній, неповній або неточній інформації, що потенційно підриває цілісність та стабільність ринку.

У цьому дослідженні [7] пропонується ціла низка стратегій щодо зменшення рівня ризиків, у т. ч. розробка відповідних систем управління ШІ з активним та збалансованим залученням людського нагляду, забезпеченням прозорості та кіберзахисту в системах ШІ тощо.

Також у дослідженнях доволі часто розкриваються питання ідентифікації традиційних когнітивних та емоційних упереджень за допомогою інструментів ШІ, зокрема, надмірна впевненість, небажання втрат,

ментальний облік та стадна поведінка [8]. При цьому дослідники науково доводять, що, наявність зазначених когнітивних упереджень призводить як до негативних наслідків, що проявляється у зниженні прибутковості від 2 до 25 %, так і до позитивних, але у короткостроковому періоді. Так, наприклад, стадна поведінка призводила до зростання прибутковості на 15 % у 65 % випадків спостереження. Таким чином, дослідники продемонстрували суттєвий вплив когнітивних упереджень на ефективність фінансових рішень під час використання алгоритмів ШІ.

В окремих дослідженнях зазначається, що когнітивні та емоційні упередження, характерні для людини, можуть виникати також при використанні систем ШІ [9]. В процесі дослідження петлі зворотного зв'язку між людиною та ШІ було встановлено, що технології ШІ змінюють людське сприйняття, емоційні та соціальні судження, що згодом посилює когнітивні та емоційні упередження у людей. При цьому це посилення значно інтенсивніше, ніж в системі комунікацій між людьми. Все це відбувається через неусвідомлення ступеня впливу ШІ на прийняття рішень. Автори презентують так званий «ефект снігової кулі», коли незначні на перший погляд помилки людей у судженнях переростають у значно більші.

У [10] зауважується, що існує нагальна потреба розглядати ШІ не просто як технологічний інструмент досягнення ефективності фінансових рішень, а як соціотехнічну систему, невіддільну від умов, в яких ШІ розроблений, впроваджений та використаний. Як підкреслює автор, значний вплив людських та суспільних факторів на упередження ШІ ігнорується, що призводить до певних викривлень та дисбалансів у використанні ШІ. Відповідно, виникає необхідність розгляду когнітивних упереджень та людської евристики як основи забезпечення справедливості рішень ШІ, а також дослідження того, як шкідливі дії людини впливають на загальний життєвий цикл ШІ. Для цього автором пропонується застосовувати порівняльний аналіз людської евристики з відображенням упереджень ШІ.

Таким чином, узагальнюючи викладене вище, можна констатувати, що більшість дослідників спрямовують свою увагу на те, який саме вплив спричиняє поведінка людини на алгоритми ШІ, визнається наявність зворотного зв'язку, що призводить до зміни у поведінці самих користувачів алгоритмів, однак поки що ці дослідження мають точковий характер і потребують подальшого опрацювання та розвитку.

Метою статті є дослідження механізмів формування поведінкових патернів суб'єктів фінансового ринку під час застосування фінансових технологій, заснованих на використанні алгоритмів ШІ різного типу та виявлення можливих ризиків таких моделей поведінки економічних суб'єктів.

Викладення основного матеріалу. Досліджуючи увагу науковців до питань врахування психологічних аспектів у діяльності фінансового ринку, автори сходяться у думках щодо наявності такого впливу та необхідності його оцінки. Автори [11–13] акцентують увагу на таких когнітивних упередженнях, як небажання втрат, надмірна впевненість, стадний ефект, закріплення активів та ін., що є основними причинами нерегульованих ринків, та роль, яку відіграють когнітивні упередження у дестабілізації ринку. При цьому наголошується на необхідності виявлення та розуміння упереджень інвесторів, що має вирішальне значення для розробки ефективних стратегій управління ризиками [12, 13]. Однак, як було вказано вище, з активним впровадженням ШІ почали формуватися «зворотні петлі» впливу [9]. Як зазначають у своєму дослідженні M.Glickman та T.Sharot [9], взаємодія людини зі ШІ призводить до того, що навіть незначні упередження посилюються і передаються наступним користувачам відповідного алгоритму ШІ, що практично не спостерігається при взаємодії системи «людина – людина». При цьому під час взаємодії людини з ШІ під виглядом людини перші були менш схильні до посилення упереджень, ніж при взаємодії зі ШІ, розуміючи, що це ШІ. Однак, з іншого боку, автори зазначають, що взаємодія з точним алгоритмом ШІ підвищує точність, а, відповідно, й ефективність прийнятих рішень.

Така дуалістична природа впливу ШІ на точність прийняття рішень, яка також визнається авторами [8], свідчить про те, що потребує вивчення питання побудови ефективної стратегії взаємодії в системі «людина – ШІ», яка буде посилювати «умовно-позитивні» впливи та зменшувати або навіть запобігати появі «умовно-негативних» впливів. У цьому контексті доцільно визначитися з самою моделлю взаємодії в системі «людина – ШІ», яка дозволяє зрозуміти та спроектувати ефективну взаємодію, виявити потенційні проблеми, оптимізувати продуктивність та забезпечити відповідальне використання ШІ. Як вже було зазначено [10], ШІ – це не просто технічний процес, а складна соціотехнічна система, що вимагає врахування як технічних аспектів ШІ, так і психологічних, когнітивних та соціальних аспектів людини. Відповідно, для розуміння саме когнітивних аспектів взаємного впливу доцільно визначитися з елементами та каналами взаємодії, а також ролями людини та ШІ в цій системі.

Якщо розглядати елемент «людина» в процесі взаємодії з ШІ, то слід чітко розуміти, що роль пересічного користувача та роль, наприклад, тренера даних для систем ШІ є принципово різною за впливом на наслідки роботи ШІ. Усвідомлюючи той факт, що людина є апріорі носієм когнітивних упереджень, які можуть підпадати під групову динаміку, а власне групова динаміка може підсилювати або послаблювати особисті когнітивні упередження, можемо стверджувати, що залежно від певної ролі людини у створенні та експлуатації алгоритмів ШІ змінюється і ступінь впливу особистих та групових когнітивних упереджень на алгоритм ШІ (табл. 1).

Вплив людини на формування когнітивних упереджень в алгоритмах ІІІ

Роль	Мета взаємодії	Когнітивні упередження	
		канали передачі	механізм передачі
Користувач	Отримання користі від функціоналу ІІІ	Форма запиту (голос, текст, жести, дані з датчиків тощо); стереотипи та суспільні норми, вбудовані в мову та культуру	Емоційність реакції; вияв довіри до результатів; явні та неявні реакції; зворотний зв'язок
Розробник / Інженер ІІІ	Створення, навчання, тестування, оптимізація алгоритмів ІІІ	Коди, архітектура моделей, характер роботи з даними, функції втрат, критерії оптимізації	Критерії вибору алгоритмів та платформ кодування, архітектури моделей (об'єктивні та суб'єктивні)
Тренер / Анотатор даних	Створення ефективних моделей на основі великих обсягів даних	Метадані та атрибути даних; розмітка, класифікація даних; транскрибовані аудіодані	Оцінка якості даних, класифікаційні групи та ознаки даних; критерії надійності та достовірності даних; ідентифікація емоцій у текстах, аудіо- та відеоданих
Експерт доменної області	Вирішення конкретних завдань у певній галузі, зокрема фінансові ринки	Експертна передача знань про фінансові ринки, інструменти, системи та механізми	Вибір алгоритмів прийняття фінансових рішень; знання про фінансові ринки, фінансові інструменти; регуляторні фінансові обмеження тощо; критерії перевірки коректності та обмежень рішень ІІІ
Тестувальник / Верифікатор	Оцінка коректності та етичності роботи алгоритмів ІІІ	Функціональне тестування	Виявлені помилки, упередження або несподівана поведінка; критерії оцінки етичності роботи
Регулятор / Етик	Регулювання використання; розробка етичних норм	Рамки, стандарти, керівні принципи для відповідального створення, використання	Критерії конфіденційності, упередженості та прозорості
Оператор / Монітор	Контроль роботи ІІІ у режимі реального часу	Механізм втручання під час збоїв або нештатних ситуацій	Порядок, критерії та наслідки втручання; прийняття остаточних рішень

Джерело: розроблено автором на базі [14–16]

Відповідно, на кожному етапі створення, використання та вдосконалення алгоритмів ІІІ можливий вплив когнітивних упереджень учасників, що може призводити як до накопичувального ефекту когнітивних упереджень, так і до певного балансування ступеня упередженості внаслідок впливів регуляторів, тестувальників, етиків, операторів алгоритмів ІІІ. При цьому вже самі алгоритми ІІІ стають джерелом формування когнітивних упереджень у людини, формуючи «петлю зворотного зв'язку» в системі «людина – ІІІ». Хоча ІІІ пропонує значний потенціал для пом'якшення традиційних людських упереджень, його інтеграція у процес прийняття фінансових рішень не позбавлена власних складнощів, зокрема, упередженість автоматизації, надмірна впевненість, упередженість доступності, алгоритмічна упередженість, алгоритмічне стадна поведінка, проблема прозорості (ефект чорної скриньки).

Упередженість автоматизації, яку також називають «схильністю до автоматизації», трапляється, коли люди сліпо довіряють рішенням, прийнятим алгоритмами ІІІ, не оцінюючи критично їхню точність чи справедливість, навіть коли ці рішення суперечать їхнім власним судженням [17]. Внаслідок виникнення

такого когнітивного упередження виникає надмірна залежність від рекомендацій та рішень ШІ, тобто людина занадто довірливо починає ставитися до рішень та рекомендацій, що надаються ШІ, коли навіть очевидні помилки залишаються без належної уваги просто тому, що вони виникли в системі на основі ШІ. Саме в цих умовах для фінансових інституцій є критичним створення відповідних систем моніторингу та контролю в межах ризик-менеджменту, який би дозволяв використовувати відповідні «запобіжники», створюючи «контрольний буфер», який би дозволив за певних визначених параметрів угоди формувати процедури на кшталт «stop-loss», що б убезпечувало клієнтів, які підпадають під упередження автоматизації, від прийняття помилкових рішень. При цьому також важливим етапом зниження рівня ризику інвестора є встановлення параметрів такої «помилковості» рішення, які б з одного боку убезпечували клієнта від дійсно хибних рішень, які призводять до значних як поточних, так і стратегічних втрат, а з іншого – не ставали б обмеженням або навіть перешкодою на шляху виконання можливо високоризикованих, але прийнятних для клієнта операцій.

Аналогічний механізм моніторингу та контролю є необхідним при виникненні упередження надмірної впевненості у тих порадах та рекомендаціях, які генеруються ШІ, оскільки задекларована розробниками та, можливо, навіть доведена певним досвідом використання висока точність і прогностична сила моделей ШІ можуть призвести до того, що інвестори починають занадто довірливо ставитися до рекомендацій і, як наслідок, переоцінюють власну здатність приймати ефективні фінансові рішення. Людська інтелектуальна властивість піддавати сумніву зовнішню інформацію в умовах використання ШІ знижується за наявності суттєвого позитивного досвіду попереднього використання алгоритмів ШІ, ігноруючи можливість галюцинації ШІ [18].

Ще одним аспектом, що підлягає усвідомленню, є баланс між короткостроковими операціями та довгостроковими стратегіями інвесторів. У своїй роботі Морган Гаусел [19] зазначає, що інвестори, наслідуючи дії когось поруч, практично ніколи не можуть знати, в яку саме гру грає інший інвестор, чи його гра «в коротку», спрямована на спекулятивний дохід, чи «в довгу» – спрямована на зростання вартості капіталу. Тому просте покладання на приклади дій інших інвесторів, а іноді й залучення до когнітивного упередження стадного ефекту може призводити до хибних рішень. Водночас ШІ, маючи можливість обробляти величезні масиви найактуальніших даних не тільки фінансового ринку, а взагалі матеріальної інформації (фінансові звіти, новини, аналітичні дописи та звіти, наукові статті, пресрелізи компаній тощо), може сприяти упередженості доступності, яка є формою упередженості щодо високого ступеня новизни згенерованої інформації, роблячи нещодавню інформацію ще більш помітною та легкозасвоюваною, та, відповідно, спонукати інвесторів надмірно зосереджуватися на короткострокових змінах ринку, ненавмисно випускаючи з уваги важливі довгострокові стратегічні фундаментальні показники ринку, потенційно спотворюючи інвестиційні перспективи. Таким чином, з'являється необхідність формування ще одного «запобіжника» щодо переоцінки новизни інформації та дотримання інвестором обраної стратегії. І знов таки при функціонуванні такого «запобіжника» постає питання балансу та адекватності умов його роботи. Іноді інвесторам дійсно необхідно змінити принципово свою стратегію дій, а іноді і повний розворот концепції ведення бізнесу, вдавшись до PIVOT, замість впертого ігнорування актуальних сигналів ринку, що зі свого боку також є когнітивним упередженням.

Значним поведінковим ризиком на фінансових ринках, керованих ШІ, є явище алгоритмічного упередження, яке формується внаслідок схильності алгоритмів ШІ до успадкування та посилення упереджень, наявних в історичних наборах даних, які часто є ненавмисними, виникають внаслідок людських рішень, але при цьому стають основою навчання алгоритмів ШІ [6]. При цьому слід враховувати, що якщо недоліком навчального набору даних є низький рівень різноманітності, певних дискримінацій (гендерних, вікових, територіальних, економічних, фінансових, суспільних, стереотипних тощо), то і результати навчання та моделювання можуть відображати дискримінаційні рекомендації, надаючи перевагу певним «знайомим» для моделі об'єктам, що не завжди буде мати раціональний характер та може призводити до дискримінаційних результатів у критичних функціях [20], таких як прийняття позитивних або негативних рішень щодо надання кредитів, оцінки рівня кредитного скорингу, у страхуванні під час визначення страхових платежів, інвестиційних стратегій тощо.

Більш того, зважаючи на згаданий вище ефект зворотної петлі в системі «людина – ШІ» [9], такі алгоритмічні упередження, ймовірно, можуть призводити й до розвитку відповідних дискримінаційних упереджень у користувачів таких рішень. Якщо кредитний менеджер з певною регулярністю отримує типові рекомендації для певних категорій аплікантів на кредит, то з плином часу в нього може сформуватись або посилитись упередження щодо відповідної категорії клієнтів. Особливо якщо це формується за «нефінансовими» ознаками, такими як гендер, вік, професія, територіальне розташування тощо. Тобто, сама присутність ШІ змінює психологію інвестора, іноді таким чином, що це породжує нові форми ірраціональності або посилює існуючі, що підкреслює важливість розуміння того, що простого розгортання ШІ без усвідомлення його психологічного впливу на користувачів критично недостатньо.

Проблему алгоритмічного упередження також ускладнює проблема «чорної скриньки», де ключові алгоритмічні процеси прийняття рішень моделями ШІ є непрозорими або важкими для інтерпретації та

пояснення людьми, що своєю чергою ускладнює процес оцінки об'єктивності рішень, виявлення та усунення упереджень і помилок, які внесені системою ІІІ як базовий набір даних, що створює ризик увіковічення таких дискримінаційних тенденцій. Відповідно, етичний імператив ІІІ у фінансах поширюється не лише на результати окремих інвесторів, а й на суспільство в цілому, зменшуючи рівень фінансової інклюзії.

Узагальнюючи розглянуті упередження, спричинені використанням ІІІ, а також проаналізований вплив на фінансові рішення, доцільно запропонувати відповідний інструментарій контролю впливу. При цьому важливим є як контроль на рівні алгоритмів, так і контроль на рівні особистого сприйняття певних явищ особистістю (табл. 2).

Таблиця 2

Когнітивні упередження, спричинені використанням ІІІ під час прийняття фінансових рішень, та інструменти їх контролю

Когнітивні упередження	Результат впливу	Інструменти контролю впливу	
		алгоритмічного	когнітивного
Упередження автоматизації	Неоскаржені помилки; слабкий людський нагляд; поодинокі точки відмови	Об'єктивний моніторинг алгоритму та результатів його використання; людський нагляд за системою «stop-loss»; дослідження причин і наслідків збоїв та галюцинацій ІІІ	Критичне мислення; раціональний сумнів; встановлення особистих фінансових обмежень
Надмірна впевненість	Надмірний ризик; потенційні фінансові втрати; надмірна залежність від ІІІ	Адекватні параметричні характеристики для інвестиційних задач; портфель повноважень ІІІ у фінансових рішеннях	Критичне мислення; раціональний сумнів; встановлення особистих меж використання ІІІ
Упередження доступності	Нехтування довгостроковими показниками ринку; спотворення перспективи ринку	Адекватні параметричні характеристики для інвестиційних задач; часові налаштування промптів з визначенням пріоритетів	Критичне мислення Пріоритетність цілей Особистий моніторинг матеріальної інформації
Алгоритмічне упередження	Дискримінаційні результати	Контроль об'єктивності та різноманітності даних для навчання; забезпечення прозорості «чорної скриньки»	Виявлення та контроль власних стереотипів, атрибутивності груп, соціальної ідентичності тощо

Джерело: розроблено автором

Окремими питаннями, що потребують вивчення, є аспекти системного ризику та волатильності фінансового ринку внаслідок застосування, які виникають в умовах застосування алгоритмів ІІІ зокрема для високочастотної торгівлі (High Frequency Trading), коли масовість та швидкість обробки виставлених ордерів може спричиняти підвищення волатильності ринку, а іноді й спричиняти ринкові колапси, підкреслюючи та посилюючи ринкові тенденції замість їх згладжування та нівелювання, що викликає появу алгоритмічного стадного упередження. Крім того, спираючись на історичні дані та не володіючи людською властивістю інтуїтивного відчуття ринку та адаптивної здатності, моделі ІІІ можуть ставати обмежено придатними для прийняття фінансових рішень в умовах появи так званих «чорних лебедів».

З алгоритмічної точки зору існують спроби пояснити та обчислити певні параметри, що будуть свідчити про наближення «чорного лебедя» [21], однак, як зазначають дослідники теорії екстремальних значень (EVT) на фінансових ринках, якщо можна точно наближено оцінити величину такої рідкісної події, то не можна передбачити, коли подія відбудеться, з будь-яким значним ступенем впевненості. Відповідно, можемо зробити висновок, що ті інвестори, які вже мають упередження автоматизації, доступності та новизни, занадто покладаючись на рекомендації та алгоритми ІІІ, та внаслідок власного ймовірного недостатнього рівня фінансової грамотності й досвіду, можуть зазнавати значних збитків через ігнорування таких критичних подій або відсутності запасної стратегії на випадок появи «чорного лебедя».

Таким чином, ІІІ – це не нейтральний інструмент генерації відповідей на поставлені запитання або алгоритмізації виконання поставлених завдань. ІІІ стає активним учасником, який може змінювати поведінку людини, особливо в тих випадках, коли людина схильна до довірливості, володіє недостатньою базою знань, навичок, вмінь та досвідом у відповідній галузі, легко підпадає під зовнішній вплив.

Тобто, відповідальна інтеграція ШІ повинна включати стратегії управління новими, виникаючими упередженнями взаємодії людини та ШІ, що вимагає глибшого розуміння когнітивної психології в контексті передових технологій.

Висновки та перспективи подальших досліджень. Сучасний фінансовий ринок зазнає суттєвих трансформацій під впливом стрімкого розвитку та впровадження ШІ. При цьому зміни не обмежуються технічними аспектами функціонування ринку, а глибоко торкаються механізмів формування поведінкових патернів його учасників. ШІ трансформує традиційні моделі поведінки, що ґрунтувалися на ірраціональності або обмеженій раціональності людського фактора, однак, не дивлячись на високий рівень алгоритмізації фінансових операцій, що певним чином нівелює багато традиційних психологічних аспектів, це призводить не до зникнення поведінкових патернів, а до їхньої кардинальної трансформації. ШІ не лише реагує на ринкові стимули, але й сам стає потужним джерелом їх формування, створюючи нові «цифрові» поведінкові патерни, які можуть мати непередбачувані наслідки. Це підкреслює необхідність дослідження не тільки поведінки людини на ринку, але й поведінки самих алгоритмів, їхньої взаємодії, впливу на загальну динаміку, а особливо на поведінку користувачів ШІ.

При цьому, як засвідчили дослідження, когнітивні та емоційні упередження, характерні для людини, можуть виникати також при використанні систем ШІ. Взаємодія людини з ШІ може посилювати існуючі когнітивні та емоційні упередження у людей, причому це посилення значно інтенсивніше, ніж у системі комунікацій між людьми, що відбувається через недооцінку ступеня впливу ШІ на прийняття рішень та може призводити до так званого «ефекту снігової кулі», коли незначні помилки у судженнях переростають у значно більші, утворюючи петлю зворотного зв'язку між людиною та ШІ.

У дослідженні було проаналізовано вплив людини на формування когнітивних упереджень в алгоритмах ШІ залежно від його ролі в системі «людина – ШІ» та визначено канали й механізми передачі когнітивних упереджень від людини в алгоритми ШІ. Це дозволяє зрозуміти, на яких етапах використання алгоритмів ШІ можуть формуватися, інтегруватися та посилюватися відповідні людські когнітивні упередження, створюючи основу для специфічних когнітивних упереджень внаслідок використання ШІ.

Серед специфічних когнітивних упереджень, що виникають під впливом ШІ, було проаналізовано упередженість автоматизації, надмірну впевненість, упередженість доступності, алгоритмічне упередження, які у сукупності суттєво впливають на поведінку людей, їх фінансові рішення, а відповідно і рівень життя. В дослідженні запропоновано відповідний інструментарій контролю впливу когнітивних упереджень як на алгоритмічному, так і на особистісному рівні.

Враховуючи виявлені виклики, подальші дослідження планується зосередити на глибшому дослідженні психологічних механізмів формування нових «цифрових» поведінкових патернів, зокрема на розумінні когнітивних та емоційних процесів, що лежать в основі виникнення та поширення нових поведінкових патернів, опосередкованих ШІ, а також розробці комплексних інструментів контролю та моніторингу когнітивних упереджень у системі «людина – ШІ». При цьому особливий акцент має бути зроблений на «запобіжниках», які убезпечують від хибних рішень, спричинені дією відповідних упереджень.

References:

1. «AI can help banks unleash a new era of software engineering productivity», [Online], available at: <https://www.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2025/ai-and-bank-software-development.html>
2. PWC (2017), «Sizing the prize. What's the real value of AI for your business and how can you capitalise?», [Online], available at: <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>
3. IMF (2025), «The Real Impact of Artificial Intelligence on Finance and Regulation», [Online], available at: <https://galapagoscapital.com/en/imf-2025-the-real-impact-of-artificial-intelligence-on-finance-and-regulation-2/>
4. Allied Market Research (2023), *Robo Advisory Market Size, Share, Competitive Landscape and Trend Analysis Report, by Business Model, by Provider, by Service Type, by End User: Global Opportunity Analysis and Industry Forecast, 2022 – 2032*, [Online], available at: <https://www.alliedmarketresearch.com/robo-advisory-market>
5. Patil, D. (2025), *Artificial Intelligence In Financial Risk Assessment And Fraud Detection: Opportunities And Ethical Concerns*, doi: 10.2139/ssrn.5057434.
6. Kumar, R. and Jha, N. (2025), «Artificial intelligence in behavioural finance: opportunities and challenges», *Deleted Journal*, Vol. 07, No. 04 (I), pp. 194–200, doi: 10.62823/ijarcms/7.4(i).7074.
7. Petronijević, J., Radić, N. and Gavrilović, M. (2024), «Dependence on technology and market manipulation as potential risks of using artificial intelligence in finance», pp. 916–929, doi: 10.5937/eee24084p.
8. Podille, V.R., Varun M.A., Dhanurmika, S. et al. (2024), «Artificial Intelligence Based Behavioural Finance in Shaping Investment Strategies to Analysis of Key Biases and Heuristics», *Journal of Computer Allied Intelligence (JCAI)*, No. 2 (6), pp. 1–18, doi: 10.69996/jcai.2024026.
9. Glickman, M. and Sharot, T. (2022), «How human-AI feedback loops alter human perceptual, emotional and social judgements», *Biased AI systems produce biased humans*, doi: 10.31219/osf.io/c4e7r.
10. Vakali, A. and Tantalaki, N. (2024), «Rolling in the deep of cognitive and AI biases», *arXiv.Org*, doi: 10.48550/arxiv.2407.21202.

11. Shivhare, S. (2025), «A Comprehensive Review of Human Psychology's Role in Shaping Stock Market Investments: Behavioral Finance Insights», *International Journal of Scientific Research in Engineering and Management*, Vol. 09, pp. 1–9, [Online], available at: <https://ijsrem.com/download/a-comprehensive-review-of-human-psychologys-role-in-shaping-stock-market-investments-behavioral-finance-insights/>
12. Kanapickienė, R., Vasiliauskaitė, D., Keliuotytė-Staniulėnienė, G. et al. (2024), «A comprehensive review of behavioral biases in financial decision-making: from classical finance to behavioral finance perspectives», *Journal of Business Economics and Management*, No. 25 (5), pp. 1006–1029, doi: 10.3846/jbem.2024.22314.
13. Othman, N.A. (2024), «Mind Games in the Market: Unraveling the Impact of Psychological Biases on Your Stock Portfolio», *Social Science Research Network*, doi: 10.2139/ssrn.4844961.
14. Ramya, S. and Ajay, Ch. (2021), «Biases in AI systems. Commun», *ACM*, No. 64, Vol. 8, pp. 44–49, doi: 1145/3464903.
15. Shakhovska, N.B., Kaminskyi, R.M. and Vovk, O.B. (2018), *Systemy shtuchnoho intelektu, navchalnyi posibnyk*, Vydavnytstvo Lvivskoi politekhniki, Lviv, 392 p.
16. Steimers, A. and Schneider, M. (2022), «Sources of Risk of AI Systems», *International Journal of Environmental Research and Public Health*, No. 19 (6), doi: 10.3390/ijerph19063641.
17. Bayakhmetova, A., Rudenko, L., Krylova, L. et al. (2025), «Artificial Intelligence in Financial Behavior: Bibliometric Ideas and New Opportunities», *Journal of Risk and Financial Management*, No. 18 (3), doi: 10.3390/jrfm18030159.
18. Ziwei, J., Nayeon, L., Frieske R. et al. (2023), «Survey of Hallucination in Natural Language Generation», *ACM Comput. Surv.*, No. 55 (12), doi: 10.1145/3571730.
19. Housel, M. (2020), *The Psychology of Money. Timeless Lessons on Wealth, Greed and Happiness*, Harriman House Publishing, 288 p.
20. Kavitha, D.M., Hanumanthu, D.K.D., Sai, O.N. et al. (2025), «The Role of Artificial Intelligence in Financial Decision-Making», *Journal of Marketing & Social Research*, No. 2 (2), pp. 189–198, [Online], available at: <https://jmsr-online.com/article/the-role-of-artificial-intelligence-in-financial-decision-making-65/>
21. Manhire, J. (2018), «Measuring Black Swans in Financial Markets», *Journal of Mathematical Finance*, No. 8, pp. 227–239, doi: 10.4236/jmf.2018.81016.

Єрмошкіна Олена Вячеславівна – доктор економічних наук, професор кафедри фінансів Національного університету «Львівська політехніка».

<https://orcid.org/0000-0002-4013-3575>.

Наукові інтереси:

– поведінкові фінанси.

E-mail: olena.v.yermoshkina@lpnu.ua.

Yermoshkina O.V.

Formation of cognitive biases in the financial market under artificial intelligence development

The article examines the transformation of behavioral patterns of the participants on the financial market under the influence of the rapid implementation of artificial intelligence (AI). It was noted that traditional models of financial behavior, based on bounded rationality, are undergoing significant changes. AI not only responds to market stimulus, but also independently shapes them, creating new digital behavioral patterns. Particular attention was paid to cognitive biases that arise in both users and AI algorithms, including automation bias, overconfidence, accessibility bias, and algorithmic bias. The mechanisms of interaction in the «Human – AI» system were considered, the roles, channels, and mechanisms for transmitting biases from humans to algorithms were identified. The research offers the classification of participants in the «Human-AI» system (user, developer, annotator, expert, regulator, etc.) and reveals how bias can arise or intensify at each stage. The «snowball effect» was described, when even minor cognitive errors of users scale over time due to the peculiarities of AI. The study presents the tools for controlling cognitive biases at the algorithmic and personal levels. It was emphasized that underestimating the psychological impact of AI on users can lead to risky decisions and reduce financial stability. The author emphasizes the need of integration of a sociotechnical approach to investigating the impact of AI on financial behavior, based on technical, cognitive and social factors. The necessity for developing comprehensive strategies for monitoring, controlling and ethical regulation of AI in the financial sector was substantiated. Prospects for further research targeted on a deeper analysis of the psychological mechanisms of forming «digital» behavioral patterns and creating new methods for reducing the impact of cognitive distortions were discussed.

Keywords: behavior; cognitive biases; artificial intelligence; financial decisions; automation; algorithmic bias; investment behavior.

Стаття надійшла до редакції 05.08.2025.