

ЕКОНОМІЧНА ВАЛІДАЦІЯ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ В АЛГОРИТМІЧНОМУ ТРЕЙДІНГУ: КОМПЛЕКСНИЙ ПІДХІД ДО ОЦІНЮВАННЯ ПРОГНОЗНОЇ ЯКОСТІ, ВИКОНУВАНOSTІ СИГНАЛІВ І РИЗИКУ

ECONOMIC VALIDATION OF MACHINE-LEARNING MODELS IN ALGORITHMIC TRADING: AN INTEGRATED APPROACH TO ASSESSING PREDICTIVE QUALITY, SIGNAL EXECUTABILITY AND RISK

У статті обґрунтовано необхідність комплексної економічної валідації моделей машинного навчання в алгоритмічному трейдингу. Актуальність дослідження зумовлена тим, що висока прогнозна точність не гарантує практичної придатності моделі після врахування умов виконання угод, транзакційних витрат, ризику та зміни ринкових режимів. Метою є розроблення послідовної рамки оцінювання ML-моделей у цілісному торговому процесі. Методологічну основу становлять аналіз наукових праць, систематизація критеріїв прогновної, торгові та ризик-скоригованої результативності. Запропоновано п'ятирівневу рамку перевірки прогновної переваги, виконуваності сигналу, нетто-результату після ринкових фрикцій, ризик-скоригованої результативності та стабільності за різних режимів. Практична цінність полягає у застосуванні рамки для відсіву статистично привабливих, але економічно непридатних моделей.

Ключові слова: алгоритмічний трейдинг, машинне навчання, економічна валідація, прогнозна якість, ринкова мікροструктура, торговий ризик.

The article examines the economic validation of machine-learning models in algorithmic trading. Its relevance stems from the growing use of ML-based forecasting in digital financial markets and the limits of conventional predictive metrics for assessing practical applicability. High values of accuracy, AUC, RMSE, MAE, or out-of-sample R^2 may confirm a statistical relationship between a forecast and the target variable, but do not establish that it can be converted into an economically viable trading decision. The study aims to develop an integrated framework combining predictive advantage, signal executability, transaction costs, risk-adjusted performance, and stability across market conditions. The methodology is based on the analysis of research in financial machine learning, market microstructure, backtesting, and risk-adjusted performance evaluation. A conceptual approach distinguishes statistical predictive quality from economic value within the trading process. The study considers forecast transformation into a signal, order submission, execution, net result, and performance stability. It also generalizes requirements for temporal data partitioning, out-of-sample testing, walk-forward validation, stress testing, and experiment documentation. The main result is a five-level framework of economic validation. The first level tests out-of-sample and incremental predictive advantage relative to a benchmark. The second assesses technical execution under liquidity, bid-ask spread, order-book depth, available volume, and latency. The third determines whether expected gross advantage remains positive after commissions, spread costs, slippage, market impact, partial execution, and rebalancing costs. The fourth evaluates whether the net result is acceptable in terms of drawdown, PnL volatility, and tail risk. The fifth verifies whether the model retains value across changing market regimes. A negative result at any level blocks transition to the next stage until the inadequacy is addressed. The framework provides a structured procedure for developing, comparing, and testing trading systems. It helps distinguish statistically attractive but economically impractical models from those retaining value after execution conditions, transaction costs, risk constraints, and regime instability are considered.

Key words: algorithmic trading, machine learning, economic validation, predictive quality, market microstructure, trading risk.

УДК 004.8:336.76

DOI: <https://doi.org/10.32782/dees.24-11>

Бабич О.А.¹

доктор філософії з міжнародних економічних відносин,
асистент кафедри міжнародного бізнесу,
Навчально-науковий інститут міжнародних відносин
Київського національного університету імені Тараса Шевченка

Babych Oleksandr

Educational and Scientific Institute of International Relations
of the Taras Shevchenko National University of Kyiv

Постановка проблеми. Розвиток цифрових фінансових ринків і поширення алгоритмічних торгових систем посилюють використання моделей машинного навчання для прогнозування цінової динаміки та формування торгових рішень. Водночас висока прогнозна точність моделі не є достатньою підставою для визнання її практично придатною. Показники accuracy, precision, recall, AUC, RMSE та інші метрики характеризують відповідність прогнозу обраній цільовій змінній, але не підтверджують, що цей прогноз створює економічно значущу перевагу після його реалізації на ринку.

У алгоритмічному трейдингу між прогнозом і фінансовим результатом існує послідовність

взаємопов'язаних етапів: формування сигналу, вибір напрямку та обсягу позиції, подання заявки, її фактичне виконання та оцінювання нетто-результату. На кожному з них очікувана перевага може бути зменшена або втрачена під впливом ліквідності, ширини bid-ask spread, доступного обсягу, комісій, прослизання, затримки та часткового виконання. Тому статистично коректний прогноз не гарантує прибутковості, якщо сигнал є технічно невиконуваним, а його очікуваний результат не покриває ринкових фрикцій і прийнятого рівня ризику.

Проблема ускладнюється нестаціонарністю фінансових ринків і методологічними ризиками

¹ ORCID: <https://orcid.org/0009-0004-5996-6931>

тестування ML-моделей. Зміна волатильності, ліквідності, структури учасників і конкурентного середовища може послаблювати або усувати залежності, на яких ґрунтується модель. Одночасно багаторазовий добір ознак, параметрів і торгових правил підвищує ризики переоптимізації, витоку інформації, data snooping та хибної інтерпретації історичних результатів. Унаслідок цього модель може демонструвати привабливі показники у backtest, але втрачати економічну цінність після врахування витрат виконання, ризику та зміни ринкового режиму.

Отже, науково-практична проблема полягає у відсутності узгодженої процедури, яка дозволяла б оцінювати ML-модель не лише як інструмент прогнозування, а як елемент цілісного торгового процесу. Її розв'язання потребує комплексної економічної валідації, що поєднує перевірку прогнозової переваги, технічної виконуваності сигналу, нетто-результату після врахування ринкових фрикцій, ризик-скоригованої результативності та стабільності за різних ринкових режимів.

Аналіз останніх досліджень і публікацій. Сучасні дослідження у сфері фінансового машинного навчання поступово зміщують акцент від ізольованого оцінювання прогнозової точності до аналізу економічної придатності моделей. Показники out-of-sample R^2 , RMSE, MAE, accuracy чи AUC характеризують відповідність прогнозу цільовій змінній, але самі по собі не підтверджують здатності моделі формувати економічно значущі торгові рішення. Шіхао Гу, Браян Келлі та Дачен Сю поєднують оцінювання позавибіркової прогнозової здатності з аналізом економічних результатів портфельних стратегій [1]. Аналогічно Даніеле Б'янкі, Матіас Бюхнер та Андреа Тамоні розглядають ML-прогнози премій за ризик облігацій крізь призму їхньої інвестиційної корисності. У подальшому corrigendum автори уточнили результати після усунення використання інформації, недоступної на момент прогнозування, підтвердивши збереження статистичної та економічної значущості моделі [6; 7]. Емпіричні результати Вольфганга Дробеца та Тіціана Отто також підтверджують можливість використання ML-підходів для оцінювання факторів дохідності на європейському фондовому ринку [8].

У дослідженнях розмежовуються прогнозна якість, економічна цінність сигналу, нетто-результативність стратегії та її ризик-скоригована ефективність. Туран Г. Балі, Гайнер Бекмаєр, Матіс Мерке та Флоріан Вайґерт оцінюють ML-моделі на опціонному ринку не лише за прогнозовою здатністю, а й за результатами сформованих на їх основі стратегій з урахуванням транзакційних витрат [2]. Теїс Інгерслев Єнсен, Браян Келлі, Семен Маламуд і Лассе Гейє Педерсен підкреслюють, що торгові витрати та обмеження

перебалансування можуть істотно змінювати практичну оцінку модельних рішень [3].

Окремий напрям літератури зосереджений на ринковій мікроструктурі та виконанні заявок. Рама Конт, Арсеній Куканов і Саша Стоїков пов'язують короткострокову цінову динаміку зі станом книги заявок і дисбалансом потоку заявок [15]. Роберт Алмгрен і Ніл Крісс розглядають виконання операцій як компроміс між витратами виконання та ризиком несприятливої зміни ціни [14], тоді як Альваро Картеа, Себастьян Джаймунгал і Хосе Пенальва систематизують взаємозв'язок між швидкістю, ціною, ліквідністю та ймовірністю виконання заявки [16]. Отже, прогнозна перевага може не трансформуватися у нетто-результат за несприятливих умов ліквідності, широкого спреда, прослизання, затримки або часткового виконання.

Важливу групу становлять праці щодо нестабільності результатів і достовірності історичного тестування. Джонатан Бруґард і Абальфазл Зареї демонструють часову мінливість результативності окремих ML-підходів на фондовому ринку [4]. Девід Бейлі, Джонатан Борвейн, Маркос Лопес де Прадо та Цзі Цзім Чжу обґрунтовують ризик переоптимізації backtest за багаторазового пошуку найкращої конфігурації [9]. Гел Вайт і Пітер Райнхардт Гансен розвивають методи контролю data snooping та множинних порівнянь [10; 11], а Роб Арнотт, Кемпбелл Гарві та Гаррі Марковіц наголошують на необхідності дисциплінованого протоколу тестування ML-моделей у фінансах [5]. Ризик-скориговану результативність доцільно оцінювати не лише за коефіцієнтом Шарпа, а й з урахуванням вибіркового відбору та ненормальності розподілу результатів, для чого Девід Бейлі та Маркос Лопес де Прадо пропонують deflated Sharpe ratio [18].

Отже, наявні дослідження формують вагоме підґрунтя для оцінювання ML-моделей у фінансах, однак прогнозу якість, виконуваність сигналу, ринкові фрикції, ризик і режимну стабільність переважно розглядають як окремі аспекти. Недостатньо систематизованою залишається послідовна процедура, у межах якої модель перевіряється як елемент цілісного торгового процесу: від підтвердження прогнозової переваги – до оцінювання технічної виконуваності, нетто-результату, ризик-скоригованої ефективності та стабільності за різних ринкових умов. Саме цей методичний розрив визначає необхідність розроблення комплексної рамки економічної валідації ML-моделей в алгоритмічному трейдингу.

Постановка завдання. Метою статті є розроблення комплексної рамки економічної валідації моделей машинного навчання в алгоритмічному трейдингу, яка забезпечує послідовне оцінювання прогнозової переваги моделі, технічної виконуваності торгового сигналу, нетто-результату після

врахування транзакційних витрат і ринкових фрикцій, ризик-скоригованої результативності та стабільності за різних ринкових режимів.

Для досягнення поставленої мети визначено такі завдання: встановити обмеження традиційних метрик прогнозу як самостійного критерію практичної придатності ML-моделей; визначити параметри ринкової мікроструктури, що зумовлюють виконувальність торгового сигналу; обґрунтувати необхідність оцінювання нетто-переваги після врахування витрат виконання; систематизувати показники ризику та критерії стабільності результатів; розробити послідовність рівнів економічної валідації й процедурні вимоги до їх застосування.

Виклад основного матеріалу дослідження.

Оцінювання ML-моделей у фінансовому аналізі зазвичай починається з перевірки прогнозу здатності за показниками assigasy, AUC, RMSE, MAE та позавибіркового R^2 . Проте ці метрики відображають узгодженість прогнозу з фактичною динамікою цільової змінної, але не доводять наявності економічно значущої торгової переваги. Шіхао Гу, Браян Келлі та Дачен Сю поєднують оцінювання позавибіркової прогнозу здатності з аналізом результативності портфельних стратегій, сформованих на основі моделей [1]. Подібно Даніеле Б'янкі, Матіас Бюхнер та Андреа Тамоні оцінюють ML-прогнози премій за ризик облігацій через їхню корисність для інвестиційних рішень, а не лише через статистичні характеристики [6].

Економічний результат виникає не безпосередньо з прогнозу, а через його перетворення на торговий сигнал, вибір напрямку й обсягу позиції, подання та фактичне виконання заявки. Тому правильне визначення напрямку руху ціни не гарантує прибутковості: очікуваний рух може не покривати комісії, спред, прослизання та інші витрати або супроводжуватися неприйнятним ризиком.

У алгоритмічному трейдингу необхідно розмежовувати загальну прогнозу точність та інкрементальну економічну цінність сигналу. Високі значення AUC чи assigasy можуть формуватися переважно завдяки коректній класифікації типових спостережень, що впливають на середню метрику, але не забезпечують додаткової переваги порівняно з доступними ринковими даними чи наперед визначеним бенчмарком. Відтак практично значущою є не лише середня точність, а умовна точність і калібрування моделі у ситуаціях, де її прогнозована перевага є достатньою для компенсації витрат виконання та ризику.

Такий підхід узгоджується з теорією інформаційної ефективності ринку. Євген Фама пов'язує ефективність ринку з відображенням доступної інформації у цінах [12]. Водночас Санфорд Гроссман і Джозеф Стігліц наголошують, що повна інформаційна ефективність є проблематичною через витрати на отримання й опрацювання

інформації [13]. Отже, потенційна торгова цінність моделі визначається не самим фактом правильного прогнозу, а її здатністю формувати додаткову інформаційну перевагу, ще не повністю відображену в ринкових очікуваннях.

Необхідність переходу від статистичної до економічної оцінки підтверджують Туран Г. Балі, Гайнер Бекмаєр, Матіс Мерке та Флоріан Вайґерт, які оцінюють ML-моделі на опціонному ринку також за результатами стратегій після врахування транзакційних витрат [2]. Теїс Інгерслев Енсен, Браян Келлі, Семен Маламуд і Лассе Гейє Педерсен також показують, що валова результативність може втрачати практичну цінність після врахування обороту портфеля та витрат перебалансування [3].

Отже, прогнозна точність є необхідною, але недостатньою умовою практичної придатності ML-моделі. Її оцінювання має доповнюватися перевіркою інкрементальної інформаційної цінності, умовної точності та можливості збереження очікуваного результату після реалізації торгового сигналу.

Підтверджена прогнозна перевага ще не гарантує практичної придатності моделі. Наступним етапом є перевірка виконувальності торгового сигналу – можливості реалізувати заявку у визначеному напрямі, обсязі та часовому горизонті за поточних ринкових умов. Виконувальність визначається не самим прогнозом, а взаємодією між сигналом, параметрами заявки та станом ринкової мікроструктури.

До ключових характеристик мікроструктури належать ліквідність, ширина bid–ask spread, глибина книги заявок, доступний обсяг на релевантних цінових рівнях, короткострокова волатильність, швидкість оновлення котирувань і затримка між формуванням сигналу та поданням заявки. Рама Конт, Арсеній Куканов і Саша Стоїков показують зв'язок короткострокової динаміки цін із дисбалансом потоку заявок та зниженням глибини ринку [15]. Отже, стан книги заявок безпосередньо впливає на якість виконання та реалізований результат сигналу.

Виконання заявки передбачає компроміс між швидкістю, ціною, обсягом і ризиком реалізації. Роберт Алмгрен і Ніл Крісс розглядають його як вибір між витратами виконання та ризиком несприятливої зміни ціни впродовж періоду реалізації [14]. Швидке виконання зменшує ризик втрати актуальності сигналу, але може посилювати ціновий вплив; повільніше – обмежує безпосередній вплив на ціну, але збільшує ризик зміни ринкової ситуації.

Вибір типу заявки також змінює профіль виконання: ринкова заявка підвищує ймовірність негайної реалізації, але збільшує ризик прослизання; лімітна контролює граничну ціну, проте може бути

виконана частково або залишитися невиконаною. Альваро Картеа, Себастьян Джаймунгал і Хосе Пенальва розглядають алгоритмічне виконання як компроміс між швидкістю, ціною, ризиком та імовірністю реалізації заявки [16].

Один і той самий прогноз може бути виконуваним за вузького спреда, достатньої глибини та стабільних котирувань, але втрачати придатність за дефіциту ліквідності, підвищеної волатильності або швидкої зміни книги заявок. Терренс Гендершотт, Чарльз Джонс і Альберт Менквелд пов'язують алгоритмічну торгівлю зі змінами ліквідності та спредів [17], що підтверджує змінність умов реалізації торгових рішень.

Відтак попередня перевірка сигналу має охоплювати цінові умови, доступний обсяг і глибину книги заявок, часові характеристики виконання та ризики часткового чи повного невиконання. Така перевірка встановлює технічну реалізованість сигналу; його чиста економічна результативність після врахування всіх фрикцій оцінюється на наступному етапі.

Технічна можливість виконання заявки не гарантує її економічної доцільності. На наступному етапі очікуваний валовий результат має бути скоригований на комісії, витрати спреда, просливання, ринковий вплив, затримку та часткове виконання. Сигнал може розглядатися як економічно придатний лише тоді, коли його очікувана нетто-перевага залишається додатною за базового та консервативного сценаріїв торгових умов [2; 3].

Після підтвердження прогнозової переваги та виконуваності сигналу необхідно встановити, чи зберігає стратегія економічну цінність у часі та чи є її результат прийнятним з погляду ризику. Позитивна дохідність на історичній або окремій позавибірковій вибірці не підтверджує надійності моделі, оскільки зміни волатильності, ліквідності, структури учасників і конкурентного середовища можуть порушувати зв'язки між ознаками та майбутньою динамікою цін.

Ринковий режим доцільно розуміти як відносно однорідний стан ринку, визначений поєднанням волатильності, ліквідності, напряму цінової динаміки та інтенсивності інформаційних змін. Джонатан Бруґаард і Абальфазл Зареї виявили часове зниження позавибіркової прибутковості окремих ML-підходів до технічних торгових правил [4]. Це свідчить про можливу ерозію економічної цінності окремих моделей, а не про універсальну втрату їх ефективності, та зумовлює потребу аналізувати результати в часових і режимних підвибірках.

Оцінку моделі можуть штучно завищувати методологічні викривлення. Переоптимізація виникає, коли серед великої кількості моделей, ознак і параметрів відбирається конфігурація з найкращим історичним результатом, який частково

відображає випадкові властивості вибірки. Девід Бейлі, Джонатан Борвейн, Маркос Лопес де Прадо та Цзі Цзім Чжу показують, що багаторазовий пошук найкращої конфігурації може поєднуватися зі слабкою позавибірковою ефективністю [9]. Подібно, data snooping, витік інформації, зміщення виживання та адаптація до шуму створюють ілюзію переваги; Гел Вайт і Пітер Райнхардт Гансен обґрунтовують необхідність контролю множинних порівнянь під час перевірки стратегій [10; 11].

Від ризику помилкової оцінки моделі слід відрізнити ризик торгового результату. Сукупна дохідність має доповнюватися аналізом максимальної просадки, волатильності PnL, асиметрії розподілу, tail risk та стабільності в часових і режимних підвибірках. Частка прибуткових угод є лише допоміжним індикатором, оскільки не відображає масштабу рідкісних великих втрат. Коефіцієнт Шарпа також потребує обережної інтерпретації: Девід Бейлі та Маркос Лопес де Прадо пропонують deflated Sharpe ratio для коригування впливу вибіркового відбору, переоптимізації та ненормальності розподілу результатів [18].

Отже, валідація має передбачати часовий поділ даних, фінальний позавибірковий тест, walk-forward або rolling-window перевірку та аналіз за різних режимів волатильності, ліквідності й цінової динаміки. Роб Арнотт, Кемпбелл Гарві та Гаррі Марковіц підкреслюють, що достовірність backtesting у ML залежить від дисциплінованого протоколу, який відокремлює розроблення моделі, відбір параметрів і фінальну перевірку [5]. Таким чином, стійкість моделі визначається не сукупною дохідністю, а здатністю зберігати прийнятний ризик-скоригований результат після контролю методологічних викривлень і зміни ринкових режимів.

Оцінювання ML-моделей в алгоритмічному трейдингу не може обмежуватися перевіркою прогнозової точності або сукупної прибутковості. Практична придатність моделі залежить від поєднання її прогнозової переваги, виконуваності торгового сигналу, впливу ринкових фрикцій, прийнятності ризик-скоригованого результату та стабільності за різних ринкових режимів. На цій основі пропонується комплексна рамка економічної валідації ML-моделей, метою якої є не максимізація кількості показників, а встановлення послідовності обов'язкових перевірок, кожна з яких усуває окремий тип хибного висновку щодо практичної цінності моделі.

Рамка складається з п'яти послідовних рівнів: прогнозової якості моделі; виконуваності торгового сигналу; урахування транзакційних витрат і ринкових фрикцій; ризик-скоригованої результативності; стабільності за різних ринкових режимів. Кожний рівень є окремим фільтром. Негативний результат на будь-якому з них блокує перехід моделі до наступного етапу до усунення встановленої

причини невідповідності. Мінімальні критерії прийнятності мають визначатися ex ante, тобто до фінального тестування, з урахуванням класу активу, горизонту прогнозування, частоти угод, ліквідності, очікуваного обсягу позиції та допустимого ризикового бюджету.

Склад рівнів валідації, показники їх оцінювання, мінімальні критерії прийнятності, потенційні ризики та можливі рішення систематизовано в табл. 1.

Як показано в табл. 1, запропонована рамка діє як послідовний механізм відсіву моделей, непридатних до практичного застосування. Проходження кожного рівня є необхідною, але недостатньою умовою переходу до наступного: прогнозна перевага не гарантує технічної виконуваності сигналу; виконуваний сигнал може втратити нетто-цінність після врахування витрат; прибуткова стратегія може виявитися неприйнятною за профілем ризику або нестійкою за зміни ринкового режиму.

На першому рівні перевіряється наявність позавибіркової та інкрементальної прогносної переваги відносно наперед визначеного бенчмарку, яким може бути наївний прогноз, buy-and-hold стратегія, базова статистична модель або ринковий орієнтир. Другий і третій рівні послідовно розмежовують технічну реалізованість заявки та її економічну доцільність: спочатку встановлюється можливість виконання у потрібному напрямі,

обсязі й часовому інтервалі, а потім – збереження додатної нетто-переваги після врахування комісій, спреду, прослизання та інших фрикцій.

Наступні рівні визначають придатність стратегії для управління капіталом і її стійкість у часі. Ризик-скоригована оцінка має підтвердити, що нетто-дохідність не досягається ціною неприйнятної просадки, нестабільності результатів або хвостового ризику. Перевірка за ринковими режимами, своєю чергою, встановлює, чи не зосереджена ефективність у короткому періоді або вузьких умовах волатильності, ліквідності чи цінової динаміки. Модель, залежна від конкретного режиму, може застосовуватися лише в межах заздалегідь визначених умов.

Застосування рамки потребує процедурної коректності. Дані мають поділятися виключно у часовій послідовності на вибірки для розроблення, валідації та фінального позавибіркового тестування; випадкове перемішування спостережень не допускається через ризик витоку майбутньої інформації. До початку фінальної перевірки необхідно зафіксувати склад ознак, специфікацію моделі, правила формування й виконання сигналу, припущення щодо витрат, ризикові обмеження та критерії переходу між рівнями.

Для підтвердження стійкості висновків доцільно застосовувати walk-forward або rolling-window перевірку, а також stress testing за погіршених,

Таблиця 1

Комплексна рамка економічної валідації ML-моделей в алгоритмічному трейдингу

Рівень	Ключове питання	Основні показники	Критерій і рішення	Основний ризик
Прогнозна якість	Чи має модель позавибіркову та інкрементальну перевагу?	Accuracy, AUC, precision, recall, RMSE, MAE, out-of-sample R ² , калібрування, бенчмарк	Підтверджена перевага над бенчмарком; за відсутності – доопрацювання або відхилення	Відтворення шуму чи загальнодоступного патерну
Виконуваність сигналу	Чи можна виконати заявку в заданому напрямі, обсязі та часі?	Spread, глибина книги, доступний обсяг, імовірність виконання, latency, оновлення котирувань	Реалістичне виконання за заданими параметрами; інакше – коригування обсягу, типу заявки або умов	Часткове, запізніле або відсутнє виконання
Транзакційні витрати і фрикції	Чи зберігається додатна нетто-перевага після витрат?	Комісії, витрати спреду, прослизання, market impact, часткове виконання, перебалансування	Додатна нетто-перевага в базовому й консервативному сценаріях; інакше – перегляд правил або відмова	Валова перевага поглинається витратами
Ризик-скоригована результативність	Чи є нетто-результат прийнятним для управління капіталом?	Нетто-дохідність, максимальна просадка, волатильність PnL, Sharpe ratio, deflated Sharpe ratio, асиметрія, tail risk	Відповідність ризиковому бюджету; інакше – перегляд лімітів, виходів або логіки	Прибутковість за неприйнятної просадки чи хвостового ризику
Стабільність за ринковими режимами	Чи зберігається результативність у різних ринкових умовах?	Часові підвибірки, walk-forward, результати за волатильністю, ліквідністю й ціновою динамікою	Відсутність критичної залежності від одного періоду чи режиму; інакше – обмеження сфери застосування або відхилення	Ефективність не відтворюється після зміни режиму

Джерело: складено автором

але правдоподібних умов виконання: ширшого спреда, меншої глибини книги заявок, більшого прослизання, затримок, часткового виконання та підвищеної волатильності. Відтворюваність забезпечується документуванням версії й часових меж даних, процедур їх оброблення, параметрів моделі, припущень, результатів кожного рівня та причин допуску, доопрацювання або відхилення моделі.

Таким чином, ML-модель може вважатися такою, що пройшла комплексну економічну валідацію, лише за одночасного підтвердження прогнозної переваги, технічної виконуваності сигналу, додатної нетто-переваги після ринкових фрикцій, прийнятного ризикового профілю та стабільності за різних ринкових режимів.

Висновки. У статті обґрунтовано, що висока статистична точність моделі машинного навчання не є достатньою умовою її економічної придатності в алгоритмічному трейдингу. Прогнозна перевага може не трансформуватися у позитивний фінансовий результат через невиконуваність сигналу, ліквіднісні обмеження, спред, прослизання, комісії, часткове виконання, неприйнятний ризиковий профіль або втрату результативності за зміни ринкового режиму.

Запропоновано комплексну рамку економічної валідації ML-моделей, що передбачає послідовне проходження п'яти рівнів: перевірки позавибіркової та інкрементальної прогнозної переваги; технічної виконуваності торгового сигналу; збереження додатної нетто-переваги після врахування транзакційних витрат і ринкових фрикцій; прийнятності ризик-скоригованого результату; стабільності моделі в різних ринкових режимах. Така послідовність унеможливорює перенесення моделі до наступного етапу за наявності критичної невідповідності на попередньому рівні.

Практичне значення рамки полягає у формуванні стандартизованої процедури розроблення та тестування алгоритмічних торгових систем. Вона дозволяє відокремити статистично привабливі, але практично непридатні моделі від рішень, здатних зберігати економічну цінність після реалістичного врахування умов виконання, витрат, ризиків і режимної нестабільності. Її застосування передбачає часовий поділ даних, фінальне позавибіркове тестування, walk-forward або rolling-window перевірку, stress testing та документування параметрів і результатів експериментів.

Обмеження запропонованого підходу полягають у відсутності універсальних числових критеріїв прийнятності. Порогові значення показників мають визначатися до тестування з урахуванням класу активу, ліквідності, горизонту прогнозування, частоти угод, розміру позиції та допустимого ризикового бюджету. Крім того, якість валідації залежить від повноти історичних даних, реалістичності

припущень щодо виконання заявок і можливості коректно ідентифікувати ринкові режими.

Перспективи подальших досліджень пов'язані з емпіричною апробацією рамки на різних сегментах фінансового ринку та для різних класів ML-моделей за єдиною процедурою економічної валідації. Окремого розвитку потребує адаптація критеріїв до рівня ліквідності, волатильності та часових горизонтів, а також інтеграція адаптивних механізмів виявлення й контролю ринкового режиму в алгоритмічні торгові системи.

БІБЛІОГРАФІЧНИЙ СПИСОК:

- Gu S., Kelly B., Xiu D. Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*. 2020. Vol. 33, no. 5. P. 2223–2273. DOI: <https://doi.org/10.1093/rfs/hhaa009>
- Bali T. G., Beckmeyer H., Mörke M., Weigert F. Option Return Predictability with Machine Learning and Big Data. *The Review of Financial Studies*. 2023. Vol. 36, no. 9. P. 3548–3602. DOI: <https://doi.org/10.1093/rfs/hhad017> (дата звернення: 25.06.2026).
- Jensen T. I., Kelly B., Malamud S., Pedersen L. H. Machine Learning and the Implementable Efficient Frontier. *The Review of Financial Studies*. 2026. Article hhag022. DOI: <https://doi.org/10.1093/rfs/hhag022> (дата звернення: 25.06.2026).
- Brogaard J., Zareei A. Machine Learning and the Stock Market. *Journal of Financial and Quantitative Analysis*. 2023. Vol. 58, no. 4. P. 1431–1472. DOI: <https://doi.org/10.1017/S0022109023000571> (дата звернення: 25.06.2026).
- Arnott R. D., Harvey C. R., Markowitz H. A Backtesting Protocol in the Era of Machine Learning. *The Journal of Financial Data Science*. 2019. Vol. 1, no. 1. P. 64–74. DOI: <https://doi.org/10.3905/jfds.2019.1.064> (дата звернення: 25.06.2026).
- Bianchi D., Büchner M., Tamoni A. Bond Risk Premiums with Machine Learning. *The Review of Financial Studies*. 2021. Vol. 34, no. 2. P. 1046–1089. DOI: <https://doi.org/10.1093/rfs/hhaa062> (дата звернення: 25.06.2026).
- Bianchi D., Büchner M., Hoogteijling T., Tamoni A. Corrigendum: Bond Risk Premiums with Machine Learning. *The Review of Financial Studies*. 2021. Vol. 34, no. 2. P. 1090–1103. DOI: <https://doi.org/10.1093/rfs/hhaa098> (дата звернення: 25.06.2026).
- Drobetz W., Otto T. Empirical Asset Pricing via Machine Learning: Evidence from the European Stock Market. *Journal of Asset Management*. 2021. Vol. 22, no. 7. P. 507–538. DOI: <https://doi.org/10.1057/s41260-021-00237-x> (дата звернення: 25.06.2026).
- Bailey D. H., Borwein J. M., López de Prado M., Zhu Q. J. The Probability of Backtest Overfitting. *The Journal of Computational Finance*. 2017. Vol. 20, no. 4. P. 39–69. DOI: <https://doi.org/10.21314/JCF.2016.322> (дата звернення: 25.06.2026).
- White H. A Reality Check for Data Snooping. *Econometrica*. 2000. Vol. 68, no. 5. P. 1097–1126. DOI: <https://doi.org/10.1111/1468-0262.00152> (дата звернення: 25.06.2026).

11. Hansen P. R. A Test for Superior Predictive Ability. *Journal of Business & Economic Statistics*. 2005. Vol. 23, no. 4. P. 365–380. DOI: <https://doi.org/10.1198/073500105000000063> (дата звернення: 25.06.2026).

12. Fama E. F. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*. 1970. Vol. 25, no. 2. P. 383–417. DOI: <https://doi.org/10.1111/j.1540-6261.1970.tb00518.x> (дата звернення: 25.06.2026).

13. Grossman S. J., Stiglitz J. E. On the Impossibility of Informationally Efficient Markets. *The American Economic Review*. 1980. Vol. 70, no. 3. P. 393–408. URL: <https://www.jstor.org/stable/1805228> (дата звернення: 25.06.2026).

14. Almgren R., Chriss N. Optimal Execution of Portfolio Transactions. *Journal of Risk*. 2001. Vol. 3, no. 2. P. 5–39. DOI: <https://doi.org/10.21314/JOR.2001.041> (дата звернення: 25.06.2026).

15. Cont R., Kukanov A., Stoikov S. The Price Impact of Order Book Events. *Journal of Financial Econometrics*. 2014. Vol. 12, no. 1. P. 47–88. DOI: <https://doi.org/10.1093/jfinec/nbt003> (дата звернення: 25.06.2026).

16. Cartea Á., Jaimungal S., Penalva J. *Algorithmic and High-Frequency Trading*. Cambridge : Cambridge University Press, 2015. 343 p. DOI: <https://doi.org/10.1017/CBO9781316481183> (дата звернення: 25.06.2026).

17. Hendershott T., Jones C. M., Menkveld A. J. Does Algorithmic Trading Improve Liquidity? *The Journal of Finance*. 2011. Vol. 66, no. 1. P. 1–33. DOI: <https://doi.org/10.1111/j.1540-6261.2010.01624.x> (дата звернення: 25.06.2026).

18. Bailey D. H., López de Prado M. The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality. *The Journal of Portfolio Management*. 2014. Vol. 40, no. 5. P. 94–107. DOI: <https://doi.org/10.3905/jpm.2014.40.5.094> (дата звернення: 25.06.2026).

REFERENCES:

1. Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33 (5), 2223–2273. DOI: <https://doi.org/10.1093/rfs/hhaa009>

2. Bali, T. G., Beckmeyer, H., Mörke, M., & Weigert, F. (2023). Option return predictability with machine learning and big data. *The Review of Financial Studies*, 36 (9), 3548–3602. DOI: <https://doi.org/10.1093/rfs/hhad017>

3. Jensen, T. I., Kelly, B., Malamud, S., & Pedersen, L. H. (2026). Machine learning and the implementable efficient frontier. *The Review of Financial Studies*. Advance online publication. DOI: <https://doi.org/10.1093/rfs/hhag022>

4. Brogaard, J., & Zareei, A. (2023). Machine learning and the stock market. *Journal of Financial*

and Quantitative Analysis, 58 (4), 1431–1472. DOI: <https://doi.org/10.1017/S0022109023000571>

5. Arnott, R. D., Harvey, C. R., & Markowitz, H. (2019). A backtesting protocol in the era of machine learning. *The Journal of Financial Data Science*, 1 (1), 64–74. DOI: <https://doi.org/10.3905/jfds.2019.1.064>

6. Bianchi, D., Büchner, M., & Tamoni, A. (2021). Bond risk premiums with machine learning. *The Review of Financial Studies*, 34 (2), 1046–1089. DOI: <https://doi.org/10.1093/rfs/hhaa062>

7. Bianchi, D., Büchner, M., Hoogteijling, T., & Tamoni, A. (2021). Corrigendum: Bond risk premiums with machine learning. *The Review of Financial Studies*, 34 (2), 1090–1103. DOI: <https://doi.org/10.1093/rfs/hhaa098>

8. Drobetz, W., & Otto, T. (2021). Empirical asset pricing via machine learning: Evidence from the European stock market. *Journal of Asset Management*, 22 (7), 507–538. DOI: <https://doi.org/10.1057/s41260-021-00237-x>

9. Bailey, D. H., Borwein, J. M., López de Prado, M., & Zhu, Q. J. (2017). The probability of backtest overfitting. *The Journal of Computational Finance*, 20 (4), 39–69. DOI: <https://doi.org/10.21314/JCF.2016.322>

10. White, H. (2000). A reality check for data snooping. *Econometrica*, 68 (5), 1097–1126. DOI: <https://doi.org/10.1111/1468-0262.00152>

11. Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, 23 (4), 365–380. DOI: <https://doi.org/10.1198/073500105000000063>

12. Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25 (2), 383–417. DOI: <https://doi.org/10.1111/j.1540-6261.1970.tb00518.x>

13. Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70 (3), 393–408. URL: <https://www.jstor.org/stable/1805228>

14. Almgren, R., & Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3 (2), 5–39. DOI: <https://doi.org/10.21314/JOR.2001.041>

15. Cont, R., Kukanov, A., & Stoikov, S. (2014). The price impact of order book events. *Journal of Financial Econometrics*, 12 (1), 47–88. DOI: <https://doi.org/10.1093/jfinec/nbt003>

16. Cartea, Á., Jaimungal, S., & Penalva, J. (2015). *Algorithmic and high-frequency trading*. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781316481183>

17. Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance*, 66 (1), 1–33. DOI: <https://doi.org/10.1111/j.1540-6261.2010.01624.x>

18. Bailey, D. H., & López de Prado, M. (2014). The deflated Sharpe ratio: Correcting for selection bias, backtest overfitting, and non-normality. *The Journal of Portfolio Management*, 40 (5), 94–107. DOI: <https://doi.org/10.3905/jpm.2014.40.5.094>

Дата надходження статті: 26.06.2026

Дата прийняття статті до публікації: 21.07.2026

Дата публікації статті: 23.07.2026