

УДК 004.8

О.О. ВОВЧЕНКО*, М.П. ГРАМА*, С.В. ГРИБКОВ*, О.Л. СЕДИХ*

ПОРІВНЯЛЬНИЙ АНАЛІЗ ТРАДИЦІЙНИХ ЕКОНОМЕТРИЧНИХ МОДЕЛЕЙ ТА МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ПРОГНОЗУВАННЯ КРИВОЇ ДОХІДНОСТІ ОБЛІГАЦІЙ

*Національний університет харчових технологій, м. Київ, Україна

Анотація. Стаття присвячена порівняльному аналізу ефективності традиційних економетричних моделей та моделей штучного інтелекту для прогнозування кривої дохідності облігацій шляхом оцінки точності прогнозування та стійкості до змін ринкових умов. У дослідженні використано дані Federal Reserve Economic Database (FRED) з квартальною періодичністю за період 2011–2024 років для 10 термінів погашення державних облігацій США. Дослідження актуальне в контексті зростаючої волатильності на фінансових ринках та необхідності точних прогнозів для прийняття інвестиційних рішень і формування монетарної політики. Проведено порівняння трьох традиційних моделей (DNS, Svensson, VAR) та трьох моделей штучного інтелекту (DNN, CNN-LSTM, XGBoost). Параметризація традиційних моделей здійснювалася методом найменших квадратів, а для моделей штучного інтелекту застосовувався підхід із використанням ковзного вікна та перехресної валідації. Для забезпечення статистичної значущості результатів моделі штучного інтелекту навчалися та тестувалися в рамках 20 незалежних експериментів. Оцінка моделей здійснювалася за критеріями RMSE, MAE та коефіцієнтом детермінації R^2 . Моделі штучного інтелекту продемонстрували значно вищу точність прогнозування: DNN показала найкращі результати з RMSE 0,960 та MAE 0,832, що втричі краще за традиційні моделі. Водночас традиційні моделі виявили більшу стійкість до змін ринкових умов із коефіцієнтами стійкості, близькими до 1, тоді як моделі штучного інтелекту показали вищу чутливість до волатильності. Статистичний аналіз підтвердив значущість відмінностей у точності прогнозування між моделями ($p < 0,05$). Найнижчу точність серед усіх моделей продемонструвала VAR модель, що пояснюється її обмеженими можливостями щодо захоплення нелінійних залежностей. Запропоновані моделі характеризуються різним співвідношенням точності та стабільності, що дозволяє обирати оптимальний підхід залежно від конкретних практичних завдань.

Ключові слова: крива дохідності, прогнозування, нейронні мережі, економетричні моделі, машинне навчання.

Abstract. The article is devoted to a comparative analysis of the effectiveness of traditional econometric models and artificial intelligence models for forecasting the bond yield curve by assessing the forecasting accuracy and resilience to changes in market conditions. The study uses quarterly data from the Federal Reserve Economic Database (FRED) for the period 2011–2024 for ten maturities of US government bonds. The study is relevant in the context of growing volatility in financial markets and the need for accurate forecasts for investment decisions and monetary policy making. Three traditional models (DNS, Svensson, VAR) and three artificial intelligence models (DNN, CNN-LSTM, XGBoost) were compared. Traditional models were parameterized using the least squares method, and artificial intelligence models were parameterized using a sliding window approach and cross-validation. To ensure the statistical significance of the results, artificial intelligence models were trained and tested in 20 independent experiments. The models were evaluated by RMSE, MAE, and the coefficient of determination R^2 . Artificial intelligence models demonstrated significantly higher forecasting accuracy: DNN showed the best results with RMSE of 0.960 and MAE of 0.832, which is three times better than traditional models. At the same time, traditional models showed greater resilience to changes in market conditions with resilience coefficients close to 1, while AI models showed higher sensitivity to volatility. Statistical analysis confirmed the significance of differences in forecasting accuracy between the models ($p < 0.05$). The lowest accuracy among all the models

was demonstrated by the VAR model, which is explained by its limited ability to capture nonlinear dependencies. The proposed models are characterized by a different ratio of accuracy and stability, which allows for choosing the optimal approach depending on specific practical tasks.

Keywords: yield curve, forecasting, neural networks, econometric models, machine learning.

DOI: 10.34121/1028-9763-2025-2-45-52

1. Вступ

Крива дохідності державних облігацій є одним із важливих індикаторів фінансового ринку, що відображає взаємозв'язок між термінами погашення та дохідністю цінних паперів. Її форма та динаміка надають важливу інформацію про очікування учасників ринку щодо майбутніх процентних ставок, інфляції та економічного зростання. В сучасних умовах високої волатильності фінансових ринків та зростаючої складності економічних взаємозв'язків точне прогнозування кривої дохідності є важливим завданням для центральних банків, інвестиційних компаній та інших фінансових інституцій.

Прогнозування кривої дохідності представляє собою складну задачу через низку факторів. По-перше, фінансові часові ряди характеризуються високою волатильністю та наявністю нелінійних залежностей. По-друге, на формування кривої дохідності впливають макроекономічні фактори, взаємозв'язки між якими можуть змінюватися з часом. По-третє, традиційні економетричні моделі, хоча і мають серйозне теоретичне обґрунтування, часто виявляються недостатньо гнучкими для захоплення складних патернів у даних. З розвитком методів машинного навчання та штучного інтелекту з'являються нові можливості для вирішення цих проблем, що робить актуальним порівняльний аналіз класичних та сучасних підходів до моделювання.

Метою даного дослідження є порівняльний аналіз ефективності традиційних економетричних моделей та моделей на основі штучного інтелекту для прогнозування кривої дохідності державних облігацій. Для досягнення цієї мети поставлено такі завдання:

- протестувати три класичні економетричні моделі: модель Свенссона, динамічну модель Нельсона-Сігеля та VAR модель;
- розробити та навчити три моделі на основі штучного інтелекту: глибоку нейронну мережу (DNN), гібридну модель CNN-LSTM та модель градієнтного бустингу XGBoost;
- провести комплексне порівняння моделей за критеріями точності прогнозування, здатності захоплювати нелінійні залежності та стійкості до змін ринкових умов.

2. Моделювання та результати

У дослідженні використовуються дані з Federal Reserve Economic Database (FRED), що охоплюють період із січня 2011 року по жовтень 2024 року з квартальною періодичністю. Набір даних включає дохідності державних облігацій США для 10 термінів погашення: 3 місяці, 6 місяців, 1 рік, 2 роки, 3 роки, 5 років, 7 років, 10 років, 20 років та 30 років. Вибір цих термінів погашення забезпечує комплексне представлення всієї кривої дохідності від короткострокового до довгострокового сегментів.

Для забезпечення комплексного порівняння моделей використовуються три основні критерії оцінки.

Точність прогнозування оцінюється за допомогою стандартних метрик. Root Mean Square Error (RMSE) використовується для оцінки загальної точності прогнозів, тоді як Mean Absolute Error (MAE) надає оцінку середньої абсолютної помилки. Ці метрики обчислюються як для всієї кривої дохідності, так і окремо для різних сегментів строковості.

Здатність захоплювати нелінійні залежності оцінюється через декілька показників. Основним є коефіцієнт детермінації (R^2), який розраховується для різних сегментів кривої дохідності. Додатково проводиться аналіз точності прогнозування в періоди значних змін

форми кривої та оцінюється здатність моделі відтворювати різні типові форми кривої дохідності.

Стійкість до ринкових змін визначається комплексним підходом. Він містить порівняння RMSE в періоди високої та низької волатильності, обчислення коефіцієнта стійкості як відношення помилок у різних ринкових умовах, а також аналіз стабільності прогнозів при різких змінах ринкових умов.

Процедура тестування моделей містить кілька послідовних етапів. Спочатку проводиться валідація даних та навчання моделей на тренувальній вибірці, що охоплює період із 2011 по 2021 роки. Наступним кроком є тестування на тестовій вибірці (2022–2024 роки), де для візуалізації та порівняння використовується остання дата часового ряду (2024-10-01), що дозволяє наочно продемонструвати точність прогнозування для кожного терміну погашення облігацій. Для моделей штучного інтелекту додатково проводиться 20 незалежних запусків, що забезпечує оцінку стабільності результатів. Завершальним етапом є усереднення результатів по всіх запусках для отримання фінальних оцінок ефективності моделей.

Для моделей штучного інтелекту проводиться 20 незалежних запусків із декількох важливих причин. По-перше, через стохастичну природу процесу навчання нейронних мереж під час ініціалізації ваг мережі використовуються випадкові значення і процес оптимізації може приводити до дещо різних локальних мінімумів функції втрат. По-друге, при навчанні використовуються такі техніки, як dropout та mini-batch sampling, які також вносять елемент випадковості. Крім того, для XGBoost моделі застосовується випадковий підбір підмножини ознак та спостережень для кожного дерева.

Багаторазові запуски дозволяють оцінити стабільність роботи моделей через обчислення середніх значень та стандартних відхилень метрик якості. Це також допомагає зменшити вплив випадкових флуктуацій на фінальні результати та отримати більш надійні оцінки реальної ефективності моделей. Ба більше, такий підхід забезпечує статистичну значущість при порівнянні різних моделей між собою.

Традиційні економетричні моделі

Динамічна модель Нельсона-Сігеля (DNS) була представлена як розширення оригінальної статичної моделі Нельсона-Сігеля для кращого врахування часової динаміки кривої дохідності. Відповідно до цієї моделі, миттєва форвардна ставка $f(t)$ описується рівнянням (1)

$$f(t) = \beta^0 + \beta^1 \left(\frac{1 - e^{-\frac{t}{\tau}}}{\frac{t}{\tau}} \right) + \beta^2 \left(\frac{1 - e^{-\frac{t}{\tau}}}{\frac{t}{\tau}} - e^{-\frac{t}{\tau}} \right), \quad (1)$$

де $\beta_1, \beta_2, \beta_3$ — параметри моделі, що інтерпретуються як довгостроковий, короткостроковий та середньостроковий компоненти відповідно, а τ — параметр масштабування, що визначає швидкість загасання експоненціальних членів [1].

Модель Свенссона є розширенням динамічної моделі Нельсона-Сігеля (DNS) шляхом додавання четвертого компонента, що забезпечує більшу гнучкість у моделюванні форми кривої дохідності. Форвардна ставка в моделі Свенссона описується рівнянням (2)

$$f(t) = \beta^0 + \beta^1 \left(\frac{1 - e^{-\frac{p}{\tau}}}{\frac{p}{\tau}} \right) + \beta^2 \left(\frac{1 - e^{-\frac{p}{\tau}}}{\frac{p}{\tau}} - e^{-\frac{p}{\tau}} \right) + \beta^3 \left(\frac{1 - e^{-\frac{p}{\tau^1}}}{\frac{p}{\tau^1}} - e^{-\frac{p}{\tau^1}} \right), \quad (2)$$

де параметри τ та τ^1 є розрахунковими величинами, що визначають форму та рівень кривої безкупонної дохідності. Вони впливають на зміну середньострокових відсоткових ставок, відображаючи їхню динаміку та напрям розвитку, а також відповідають за розташування вигинів кривої дохідності [2].

Vector Autoregression (VAR) моделі представляє альтернативний підхід, що базується на аналізі часових рядів. VAR модель розглядає дохідності різних термінів погашення як взаємопов'язану систему, де кожна змінна залежить від власних лагових значень та лагових значень інших змінних. Загальна форма VAR(p) моделі(3):

$$Y_t = c + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t, \quad (3)$$

де Y_t — вектор дохідностей різних термінів погашення, c — вектор констант, A_i — матриці коефіцієнтів, а ε_t — вектор випадкових помилок [3].

Кожна з цих моделей має свої переваги та обмеження. DNS та модель Свенссона забезпечують економічно інтерпретовані параметри та гладкі криві дохідності, але можуть бути недостатньо гнучкими для захоплення складних ринкових умов. VAR моделі краще враховують динамічні взаємозв'язки між різними термінами погашення, але можуть страждати від проблеми «прокляття розмірності» при збільшенні кількості змінних [4].

На рис. 1 представлено порівняння прогнозів трьох традиційних моделей (DNS, Svensson, VAR) із фактичними значеннями кривої дохідності. Графік демонструє, що традиційні моделі мають тенденцію до недооцінки дохідності, особливо в короткостроковому сегменті кривої. Модель VAR показує найбільше відхилення від фактичних значень, як-от моделі DNS та Svensson демонструють ідентичні результати.

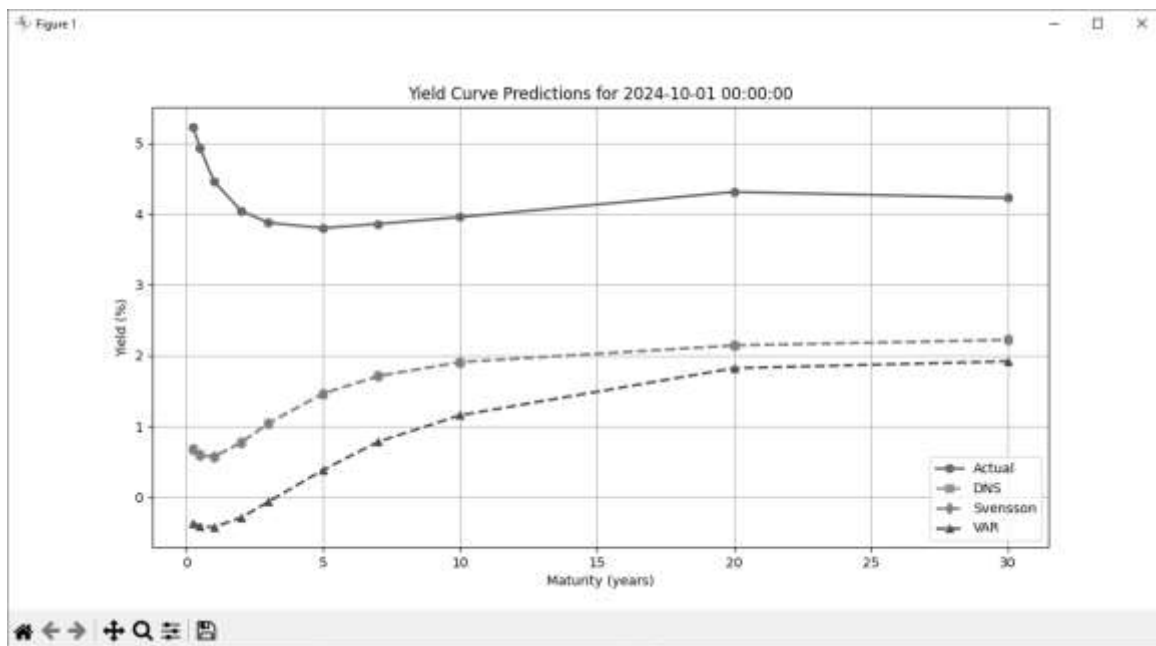


Рисунок 1 — Порівняння прогнозів традиційних моделей

Моделі штучного інтелекту

Розвиток методів машинного навчання відкрив нові можливості для прогнозування фінансових часових рядів. Глибокі нейронні мережі (DNN) представляють собою багатoshарові структури, здатні виявляти складні нелінійні залежності в даних. У контексті прогнозування кривої дохідності DNN використовують послідовність шарів нейронів для перетворення вхідних даних (історичних значень дохідності та макроекономічних

показників) у прогнози майбутніх значень. Особливістю DNN є їхня здатність автоматично вивчати ієрархічні представлення даних, де кожен наступний шар будує все більш абстрактні характеристики [5].

Гібридні моделі CNN-LSTM об'єднують переваги двох типів нейронних мереж. Згорткові нейронні мережі (CNN) ефективні у виявленні локальних патернів та особливостей у даних, тоді як довга короткочасна пам'ять (LSTM) спеціально розроблена для обробки послідовностей та захоплення довгострокових залежностей. У контексті прогнозування кривої дохідності CNN шари спочатку виявляють важливі патерни в історичних даних, а LSTM шари потім використовують ці патерни для прогнозування майбутніх значень, враховуючи часову динаміку. Дослідження показують, що така архітектура особливо ефективна для фінансових часових рядів, де важливо враховувати як короткострокові флуктуації, так і довгострокові тренди [6].

Гradientний бустинг, зокрема, його реалізація XGBoost, представляє собою ансамблевий метод машинного навчання, який послідовно будує набір простих моделей (дерев рішень), де кожна наступна модель намагається виправити помилки попередніх. XGBoost відрізняється від інших методів gradientного бустингу використанням регуляризації та оптимізованих алгоритмів навчання, що робить його особливо ефективним для структурованих даних. При прогнозуванні кривої дохідності XGBoost здатний автоматично виявляти важливі фактори та їх взаємодії, забезпечуючи баланс між точністю прогнозів та обчислювальною ефективністю [7].

Важливою перевагою моделей штучного інтелекту є їхня здатність адаптуватися до змін у ринкових умовах та враховувати нелінійні взаємозв'язки між різними факторами. Однак, на відміну від традиційних економетричних моделей, вони часто страждають від проблеми «чорного ящика», коли складно інтерпретувати механізми прийняття рішень моделлю.

На рис. 2 представлено порівняння прогнозів моделей штучного інтелекту (DNN, CNN-LSTM, XGBoost) порівняно з фактичними значеннями. DNN демонструє найкращу здатність відтворювати форму кривої дохідності, особливо в короткостроковому сегменті. CNN-LSTM показує дещо гірші результати в короткостроковому сегменті, але кращі в середньостроковому та довгостроковому сегментах, ніж DNN, тоді як XGBoost має найбільше відхилення серед моделей ШІ.

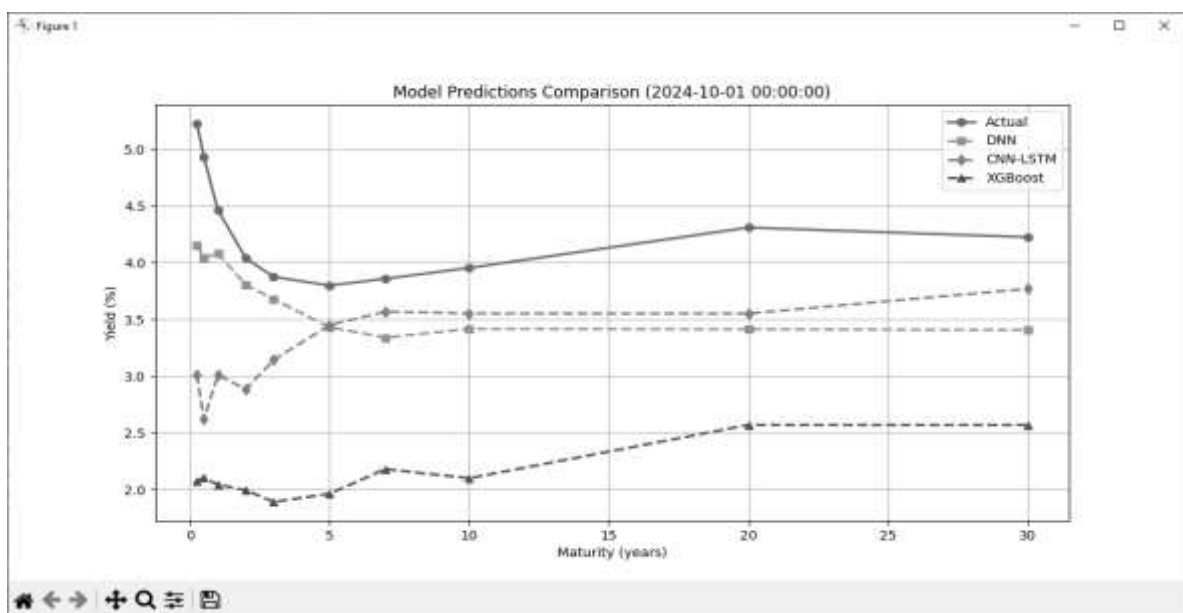


Рисунок 2 — Порівняння прогнозів моделей штучного інтелекту

Візуальне порівняння обох графіків чітко демонструє перевагу моделей штучного інтелекту, особливо DNN, у точності прогнозування кривої дохідності. Традиційні моделі мають значно більші відхилення від фактичних значень у порівнянні з будь-якою моделлю ШІ.

Комплексне порівняння моделей

Результати оцінки моделей представлені в табл. 1 та на рис. 3, які демонструють ключові метрики для всіх досліджуваних моделей.

Таблиця 1 — Порівняння метрик моделей

Модель	RMSE	MAE	R ²
DNN	0,960	0,832	0,054
CNN-LSTM	1,460	1,231	-0,624
XGBoost	2,067	1,869	-3,625
DNS	2,920	2,549	-4,044
Svensson	2,920	2,549	-4,044
VAR	3,385	2,937	-6,136

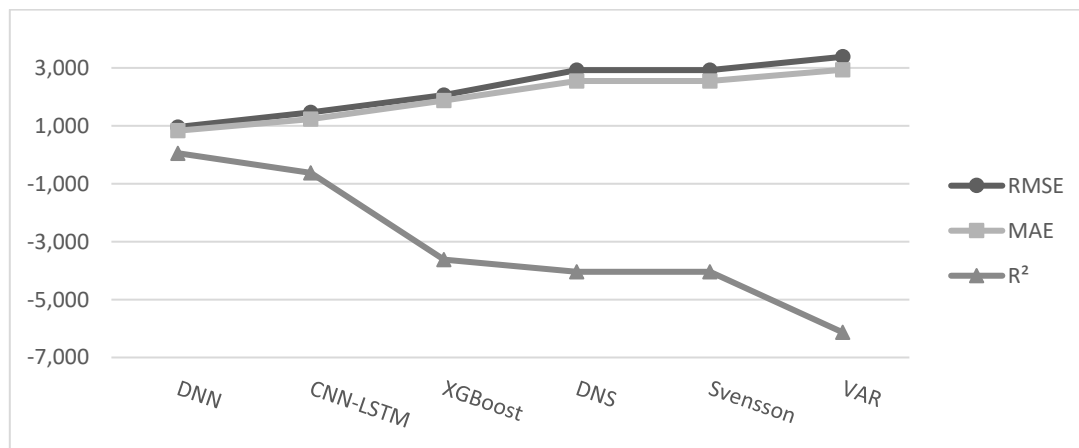


Рисунок 3 — Діаграма порівняння метрик моделей

Порівняльний аналіз точності моделей показує значну перевагу підходів на основі штучного інтелекту. DNN демонструє найкращі результати (RMSE 0,960, MAE 0,832), що майже втричі перевищує точність традиційних моделей. Друге місце посідає CNN-LSTM (RMSE 1,460, MAE 1,231), а XGBoost, хоч і показує гірші результати серед ШІ моделей (RMSE 2,067, MAE 1,869), все ж перевершує класичні підходи. Моделі DNS та Svensson показали ідентичні результати (RMSE 2,920, MAE 2,549), тоді як VAR модель демонструє найнижчу точність (RMSE 3,385, MAE 2,937).

Аналіз здатності моделей захоплювати нелінійні залежності показав суттєві відмінності між традиційними та ШІ-підходами. DNN демонструє найкращу здатність до моделювання нелінійностей із позитивним показником R² (0,054). Традиційні моделі показали значно гірші результати з негативними значеннями R², що вказує на їх обмежену здатність захоплювати складні нелінійні патерни.

Результати аналізу стійкості моделей представлені в табл. 2 та на рис. 4.

Таблиця 2 — Показники стійкості моделей у різних ринкових умовах

Модель	RMSE висока волат.	RMSE низька волат.	Стійкість
DNN	0,225	0,047	4,782
CNN-LSTM	0,262	0,060	4,389

Продовження таблиці 2

XGBoost	0,488	0,039	12,507
DNS	2,939	3,599	0,817
Svensson	2,939	3,599	0,817
VAR	3,159	4,300	0,735

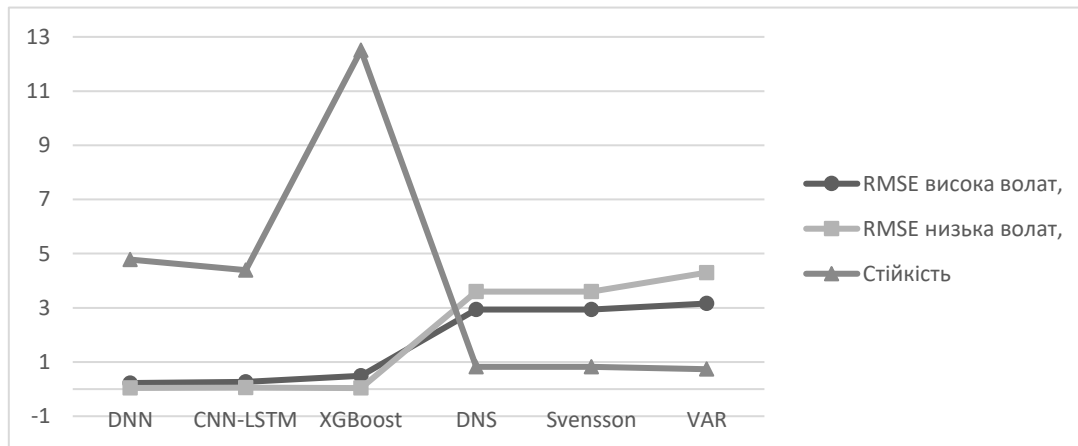


Рисунок 4 — Діаграма порівняння стійкості моделей

Аналіз стійкості моделей до ринкових змін виявляє цікаву закономірність. Традиційні моделі (DNS, Svensson, VAR) демонструють коефіцієнти стійкості, близькі до 1 (0,817, 0,817 та 0,735 відповідно), що вказує на їх відносно стабільну роботу в різних ринкових умовах, хоча з високими показниками RMSE. Натомість, моделі штучного інтелекту показують значно нижчі абсолютні помилки, але більшу чутливість до волатильності ринку. DNN та CNN-LSTM мають схожі показники стійкості (4,782 та 4,389), тоді як XGBoost демонструє найвищий коефіцієнт (12,507), що свідчить про його значну чутливість до ринкових умов, незважаючи на кращі показники RMSE порівняно з традиційними моделями.

Статистичний аналіз виявив, що відмінності в точності прогнозування між моделями є статистично значущими ($p < 0,05$). DNN показує найкращі результати за всіма метриками, за винятком стійкості в періоди високої волатильності, де традиційні моделі демонструють більшу стабільність.

З аналізу видно, що моделі штучного інтелекту, особливо DNN, показують значно кращу точність прогнозування та здатність до моделювання нелінійностей порівняно з традиційними підходами. Однак, традиційні моделі демонструють більшу стабільність в умовах низької волатильності, що може бути фактором при виборі моделі для конкретних практичних застосувань.

3. Висновки

У даному дослідженні було проведено комплексне порівняння традиційних економетричних моделей та моделей штучного інтелекту для прогнозування кривої дохідності державних облігацій. Результати аналізу демонструють суттєву перевагу моделей ШІ у точності прогнозування, де DNN показала найкращі результати з RMSE 0,960 та MAE 0,832, що майже втричі краще за традиційні підходи.

Важливим спостереженням є те, що моделі штучного інтелекту, особливо DNN, показали значно кращу здатність до моделювання нелінійних залежностей, що підтверджується позитивним показником R^2 (0,054). Водночас, традиційні моделі, хоча і

демонструють гірші показники точності, виявилися більш стійкими до змін ринкових умов, про що свідчать їх коефіцієнти стійкості, близькі до 1.

Практична значимість дослідження полягає у формуванні рекомендацій щодо вибору моделей для різних сценаріїв використання. Для завдань, де пріоритетом є точність прогнозування, рекомендується використовувати DNN або CNN-LSTM моделі. Натомість, у ситуаціях із високою волатильністю ринку або коли важлива інтерпретованість результатів, традиційні моделі DNS та Svensson можуть бути більш доцільними.

Подальші дослідження можуть бути спрямовані на розробку гібридних підходів, які б поєднували переваги обох типів моделей — високу точність моделей III та стабільність традиційних економетричних підходів. Також перспективним напрямом є включення додаткових макроекономічних факторів та дослідження їх впливу на якість прогнозування.

СПИСОК ДЖЕРЕЛ

1. Беляєв Д. Математичне моделювання як складова розробки фінансового механізму формування золотовалютних резервів. SumDU Repository: Home. URL: https://essuir.sumdu.edu.ua/bitstream/download/123456789/52520/5/Bieliaiev_Matematychne_modeliuvannia.pdf (дата звернення: 18.02.2025).
2. Paientko T., Rudaia M. Methodological approaches to the discounting rate choice in forensic economics research. *Criminalistics and forensics*. 2021. N 66. URL: <https://doi.org/10.33994/kndise.2020.66.80> (date of access: 18.02.2025).
3. Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis / G. Chen et al. *Computers in biology and medicine*. 2011. Vol. 41, N 12. P. 1142–1155. DOI: <https://doi.org/10.1016/j.compbiomed.2011.09.004> (date of access: 18.02.2025).
4. Ткачик Д.А., Кветний Р.Н. Використання економетричних моделей для прогнозування фінансових показників. Репозитарій Вінницького національного технічного університету. URL: <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/38480/17505.pdf> (дата звернення: 18.02.2025).
5. Bayraci S., Susuz O. A Deep Neural Network (DNN) based classification modelin application to loan default prediction by SelÅşuk BAYRACI, Orkun SUSUZ. *Theoretical and Applied Economics*. URL: https://www.ectap.ro/a-deep-neural-network-dnn-based-classification-modelin-application-to-loan-default-prediction-seluk-bayraci_orkun-susuz/a1421/ (date of access: 18.02.2025).
6. Deng L. et al. Analysis of the Effectiveness of CNN-LSTM Models Incorporating Bert and Attention Mechanisms in Sentiment Analysis of Data Reviews. *Proc. of the 2023 4th International Conference on Big Data and Informatization Education (ICBDIE 2023)*. Dordrecht, 2023. P. 821–829. DOI: https://doi.org/10.2991/978-94-6463-238-5_106 (date of access: 18.02.2025).
7. Nielsen D. Tree boosting with xgboost — why does xgboost win «every» machine learning competition? NTNU Open. URL: <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2433761> (date of access: 18.02.2025).

Стаття надійшла до редакції 22.03.2025