

ПОРІВНЯННЯ ЕФЕКТИВНОСТІ АРХІТЕКТУР LSTM ТА MLP ДЛЯ ПРОГНОЗУВАННЯ КОРОТКИХ ЧАСОВИХ РЯДІВ

Кубик Лілія Ярославівна

здобувачка вищої освіти магістерського рівня

<https://orcid.org/0000-0002-3652-6835>

Кафедра інформаційно-обчислювальних систем і управління
Західноукраїнський національний університет, м. Тернопіль, Україна

Прогнозування часових рядів та складних системних процесів є однією з найактуальніших проблем сучасної науки та прикладних досліджень. У цьому контексті, найперспективнішим методом прогнозування визнано використання нейронних мереж. Особливо важливим є для таких сфер, як охорона здоров'я (діагностика захворювань за зображеннями (мрт, рентген), прогнозування ефективності лікування, розробка ліків); метеорологія (деталізоване прогнозування погоди та кліматичних змін); інженерія (прогнозування терміну служби обладнання, оптимізація виробничих процесів, моніторинг структурної цілісності); вивчення ринку (прогнозування попиту, аналіз споживчої поведінки та динаміки трендів); фінанси (прогнозування динаміки цін на акції та валюту, оцінка кредитних ризиків, виявлення шахрайства) тощо [1]. Точність передбачень має вирішальне значення для прийняття стратегічних рішень, мінімізації ризиків та оптимізації ресурсів.

Нейронні мережі здобули широке визнання завдяки низці суттєвих переваг над традиційними статистичними та регресійними алгоритмами [1, 2]:

- здатність до нелінійного моделювання, через вміння ефективно фіксувати складні нелінійні взаємозв'язки у даних та здатність узагальнювати будь-яку неперервну і нелінійну функцію, забезпечуючи високу точність, недосяжну для простих статистичних моделей;

- автоматичне вилучення ознак, особливо моделі глибокого навчання, мають унікальну властивість автоматично визначати та витягувати відповідні ознаки без потреби у ручній, трудомісткій інженерії ознак, що значно спрощує етап попередньої підготовки даних та підвищує загальну ефективність моделі;

- стійкість до «зашумлених» даних, завдяки своїй паралельній структурі та здатності до узагальнення, нейронні мережі навчаються ідентифікувати приховані патерни та ігнорувати окремі аномалії (викиди), а не просто запам'ятовувати їх, що робить прогнози більш надійними в умовах реальних даних;

- масштабованість, нейронні мережі здатні використовувати великі масиви інформації для підвищення точності моделювання, що є ключовою перевагою над багатьма класичними алгоритмами, які мають значні обмеження при роботі з великими даними.

Незважаючи на ці переваги, для вирішення задач прогнозування немає універсальної архітектури нейронної мережі. Вибір моделі нейронної мережі є індивідуальним, який залежить від природи процесу та типу даних.

В задачі прогнозування вступу кількості здобувачів освіти на бакалаврат у 2025, 2026 та 2027 роках на спеціальність F3 «Комп'ютерні науки» Національного університету «Львівська політехніка» був вирішальним саме етап вибору моделі нейронної мережі. Розглянемо етапи вирішення даної задачі, які були розроблені на основі запропонованого алгоритму Харченком В. О. [1, с. 24]:

1. Завантаження даних: імпортуємо архівні значення кількості зарахованих студентів у 2016–2024 роках (табл. 1).

2. Масштабування: виконуємо нормалізацію даних в діапазоні від 0 до 1.

3. Формування послідовностей: для прогнозу одного року використовуються значення попередніх трьох років, тобто для прогнозу на 2019 рік використовуються дані за 2016, 2017, 2018 роки.

4. Побудова моделі: обираємо рекурентну нейромережу – LSTM, використовуємо два приховані шари LSTM по 50 нейронів і 2 для запобігання перенаванчання.

5. Навчання: модель навчається на послідовностях, застосовуємо оптимізатор і функцію втрат.

6. Прогнозування: після навчання мережа робить послідовні прогнози на 2025–2027 роки, використовуючи останні 3 відомі значення як вхід.

7. Денормалізація та округлення: прогнози повертаються до реальної шкали (кількість студентів), округлюються до цілих.

8. Візуалізація результатів: будуємо графіки фактичних даних (синім кольором), навченого прогнозу (зеленим) та прогнозу моделі LSTM на 2025–2027 роки (червоним) (рис. 1).

9. Аналіз похибки обраної моделі та визначення адекватності обраної моделі.

Прогноз здійснювався на основі відкритих даних, а саме на кількості зарахованих осіб на навчання для здобуття ступеня бакалавра (основа вступу – повна загальна середня освіта) у 2016–2024 роках, з реєстру Єдиної державної електронної бази з питань освіти [4] (табл. 1).

Таблиця 1 – Кількість зарахованих осіб на бакалаврат у 2016–2024 роках

№ з/п	Рік	Кількість зарахованих осіб
1	2016	402
2	2017	472
3	2018	537
4	2019	586
5	2020	547
6	2021	675
7	2022	644
8	2023	684
9	2024	514

*дані сформовано [4]

Для створення прогнозу спершу було обрано модель глибокого навчання з довгою короткочасною пам'яттю – LSTM, яка ефективно себе проявила у задачі прогнозування погоди [3], але виявилася не ефективною для передбачення кількості майбутніх студентів на спеціальності F3 «Комп'ютерні науки».

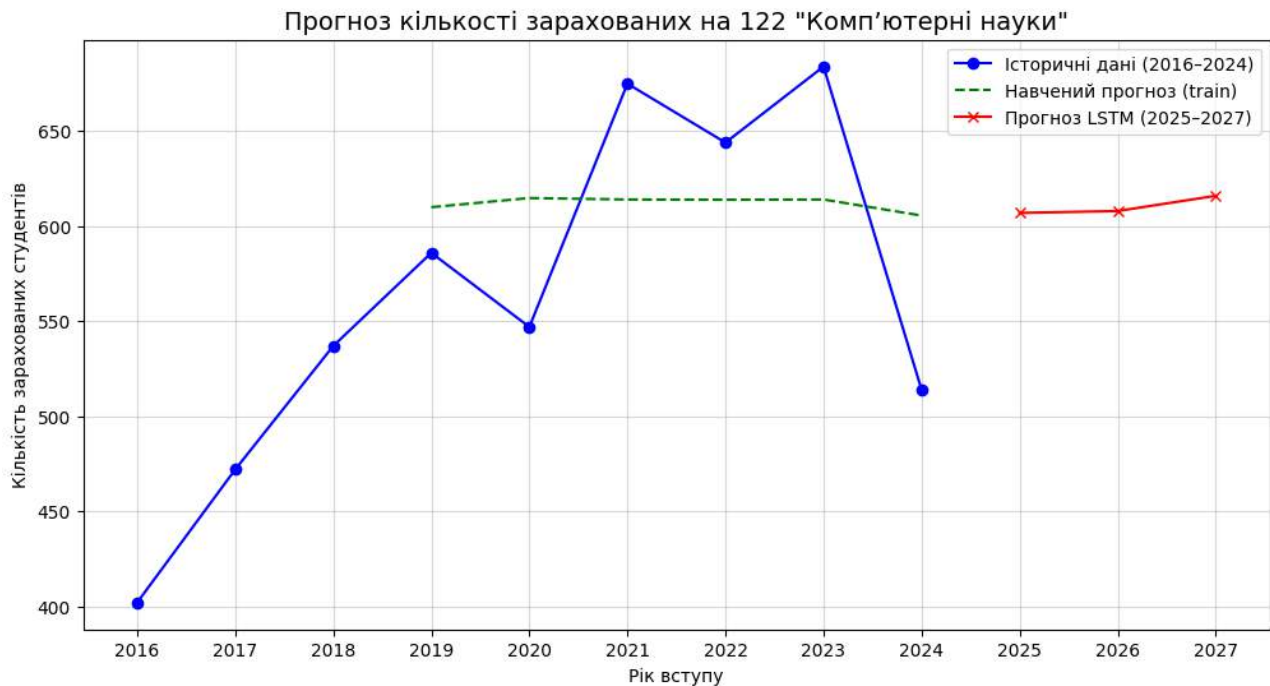


Рисунок 1. Візуалізація результатів прогнозу моделі LSTM

Прогнозовані результати моделі LSTM щодо кількості зарахованих осіб на спеціальність F3 «Комп'ютерні науки» наведені в табл. 2.

Таблиця 2 – Результат прогнозу моделі LSTM

№ з/п	Рік	Прогноз кількості зарахованих осіб
1	2025	607
2	2026	608
3	2027	616

*дані сформовано автором

Для прогнозу кількості майбутніх вступників на спеціальність F3 використовувалося лише одне значення на рік з 2016 по 2024 – це всього 9 точок (табл. 1), що утворює надзвичайно короткий часовий ряд для моделі LSTM, який є малоінформативним. Також негативно вплинуло те, що зв'язок між роками не є часовий, а причинно-соціальний, тобто залежить від зміни правил вступу чи війни в країні. Результати оцінювання якості моделі LSTM свідчать про низьку точність прогнозування та обмежену здатність до узагальнення даних. Модель лише частково відтворює тенденції у вибірці, що вказує на недостатню ефективність її навчання. Для підвищення якості прогнозів доцільно збільшити навчальний матеріал використавши допоміжні змінні (наприклад, врахувати кількість студентів, які навчаються на даній спеціальності поквартально).

Наступним кроком дослідження було повернення до етапу 4 та заміна моделі LSTM на класичну архітектуру – MLP, а саме багатошаровий перцептрон, який демонструє кращі показники на коротких рядах.

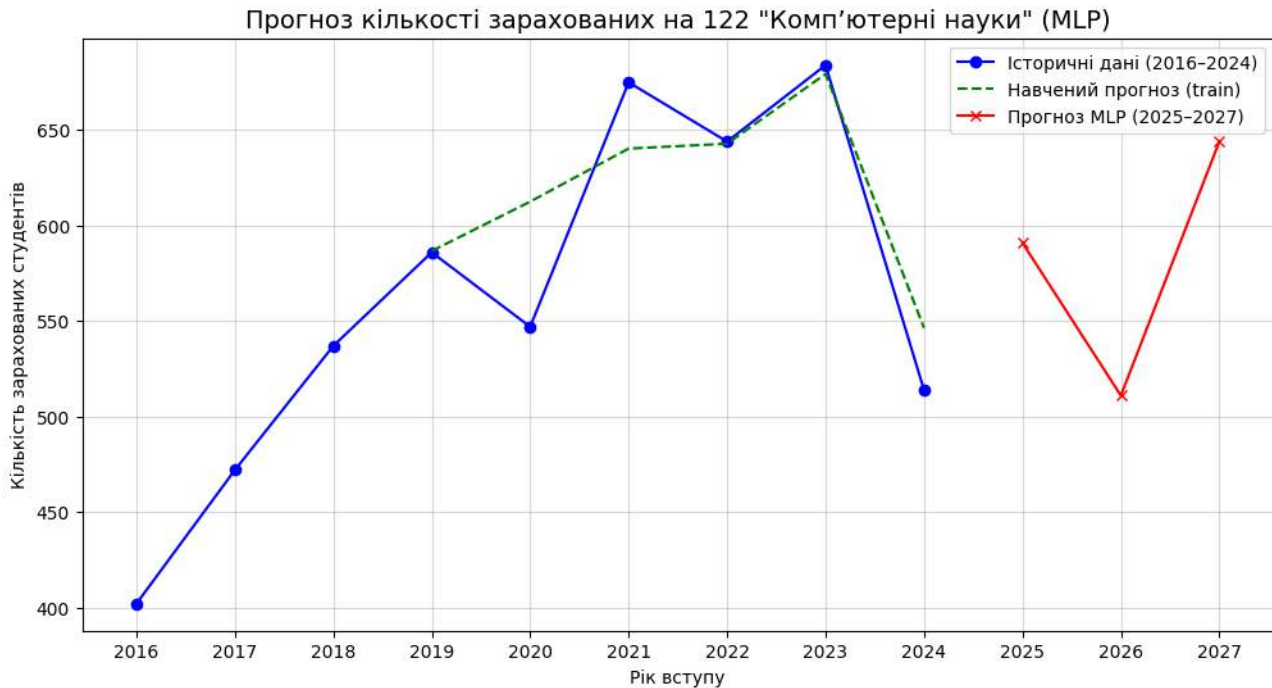


Рисунок 2. Візуалізація результатів прогнозу моделі MLP

Прогнозовані результати моделі MLP щодо кількості зарахованих осіб на спеціальність F3 «Комп'ютерні науки» наведені в табл. 3.

Таблиця 3 – Результат прогнозу моделі MLP

№ з/п	Рік	Прогноз кількості зарахованих осіб
1	2025	591
2	2026	511
3	2027	644

*дані сформовано автором

Отримані результати оцінювання якості моделі MLP свідчать про достатньо високу узгодженість прогнозованих і фактичних значень. Значення показників похибки підтверджують прийнятний рівень точності моделі та її здатність адекватно відображати основні закономірності у вихідних даних. Модель пояснює близько 73% варіації даних, що є досить хорошим результатом для емпіричних або соціально-економічних показників. Це дає підстави вважати побудовану нейронну мережу придатною для подальшого використання у задачах прогнозування.

Модель LSTM традиційно застосовується для аналізу послідовних даних і має здатність вловлювати часові залежності між спостереженнями. Проте для коротких рядів, зокрема таких, що містять лише дев'ять річних спостережень,

LSTM стикається з низкою проблем. По-перше, обсяг даних є недостатнім для ефективного формування внутрішньої пам'яті моделі. По-друге, у даному випадку спостережуваний процес (кількість вступників на певну спеціальність) не є типовим часовим рядом у класичному розумінні, оскільки залежить не лише від попередніх значень, а й від зовнішніх соціально-економічних чинників (зміни правил вступу, демографічна ситуація, воєнний стан, державне замовлення тощо). Результати моделі засвідчили невисоку якість прогнозування.

На відміну від LSTM, модель MLP показала стабільніші результати на тій самій вибірці. На етапі побудови багатошаровий перцептрон розглядав кожне спостереження як незалежну точку, що дозволило уникнути проблеми недостатності часової інформації. Завдяки своїй простішій архітектурі MLP менш схильна до перенавчання на малих наборах даних і здатна узагальнювати основні закономірності навіть за обмеженої кількості прикладів.

Узагальнюючи результати дослідження, можна зробити висновок, що для коротких історичних вибірок (до десяти спостережень) та даних, у яких часові залежності не є домінуючими, MLP виявляється ефективнішою та стабільнішою моделлю, ніж LSTM. Хоча LSTM має значні переваги у випадках довгих часових рядів, її застосування у задачах із короткими і соціально залежними даними є недоцільним. Отже, для прогнозування кількості вступників на спеціальність F3 «Комп'ютерні науки» у 2025–2027 роках оптимальним рішенням є використання моделі MLP з подальшою адаптацією її структури та параметрів під специфіку освітніх даних.

Список використаних джерел

1. Харченко В. О. Основи машинного навчання : навч. посіб. / В. О. Харченко. – Суми : Сумський державний університет, 2023. – 264 с.
2. Бурлеєв, О., Василенко, О., & Іваненко, Р. (2021). ЕФЕКТИВНІСТЬ ВИКОРИСТАННЯ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ В ЕКОНОМІЦІ. *Економіка та суспільство*, (31). <https://doi.org/10.32782/2524-0072/2021-31-27>
3. Кубик Л.Я. Оцінка ефективності архітектури LSTM-моделі у прогнозуванні динаміки температури повітря. Collection of Scientific Papers with the Proceedings of the 5th International Scientific and Practical Conference «Modern Perspectives on Global Scientific Solutions» (October 13-15, 2025, Bergen, Norway). European Open Science Space. 2025. № 57. С. 82–85. URL: <https://www.eoss-conf.com/en/archive/modern-perspectives-on-global-scientific-solutions-13-10-25/>. DOI: 10.70286/EOSS-13.10.2025.001
4. Реєстр суб'єктів освітньої діяльності. – Режим доступу: URL: <https://registry.edbo.gov.ua/opendata/entrant/> (дата звернення: 15.10.2025р.).