

УДК 004.8

DOI 10.30837/bi.2025.1(102).05

В. В. Кириї¹, О. В. Калиниченко²¹Харківський національний університет радіоелектроніки, м. Харків, Україна,
valentyna.kyriy@nure.ua²Харківський національний університет радіоелектроніки, м. Харків, Україна,
olga.kalynychenko@nure.ua

ПРОГНОЗУВАННЯ ФІНАНСОВИХ ПОКАЗНИКІВ РИНКУ КІНОВИРОБНИЦТВА ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ

Предметом дослідження є проблема прогнозування фінансових показників на послугоорієнтованих ринках за надзвичайних умов.

Метою роботи є розробка ефективного підходу для прогнозування фінансових показників ринку кіновиробництва, що базується на рекурентних та згорткових нейронних мережах і використовує засоби обробки природної мови для конвертації описових даних в числовий вигляд.

У статті вирішуються такі завдання: визначення набору індикаторів, які здатні описати стану ринку кіновиробництва з точки зору компанії, зовнішнього оточення та цільової аудиторії; формування алгоритмів для передобробки числової та текстової інформації; визначення переліку цільових нейромереж та розкриття особливостей їх імплементації; визначення найбільш ефективного прогностичного підходу за допомогою вирішення задачі лінійної оптимізації.

Використовуються такі методи: аналітичний метод для визначення набору нейромереж; експертне оцінювання для формування найбільш важливих незалежних індикаторів та визначення факторів ефективності; експериментальний, багатокритеріальний оцінювання для визначення найбільш ефективної моделі.

Отримано наступні результати: сформовано набір алгоритмів передобробки даних для подальшого їх використання в рекурентних та згорткових нейронних мережах. Імплементовано декілька поширених архітектур із залученням технології MapReduce. Визначено, що найефективнішою моделлю є двонаправлена рекурентна нейромережа з підтримкою довго- та короткострокової пам'яті. Показана доцільність використання технології паралелізації та визначено набір відкритих питань для подальшого дослідження.

Висновки: проведення аналізу щодо алгоритмів прогнозування фінансових показників заснованих на штучному інтелекті з подальшою експериментальною перевіркою дозволило сформулювати відносно ефективний спосіб передбачення стану індикаторів ринку кіноіндустрії за надзвичайних умов. Отримані результати дозволяють стверджувати доцільність впровадження запропонованого підходу, що може вплинути на формування політики компанії кіновиробника чи фонду, який оперує фінансовими інструментами компанії. Водночас передбачено шляхи подальшого покращення результатів із залученням альтернативних підходів як до паралелізації, так і до прогнозування загалом.

АНАЛІЗ НАСТРОЇВ, НЕЙРОННІ МЕРЕЖІ, ОБРОБКА ПРИРОДНИХ МОВ, СОЦІАЛЬНІ ЗСУВИ, ЧАСТОТНІ ПАТЕРНИ.

V. V. Kyrii, O. V. Kalynychenko. Forecasting financial indicators of the film production market using artificial intelligence. The subject of the research is the problem of forecasting financial indicators for service-oriented markets under extraordinary conditions.

The goal of the work is to develop an effective approach for forecasting the financial indicators of the film production market, which is based on recurrent and convolutional neural networks and uses natural language processing tools to convert descriptive data into numerical form.

The article addresses the following tasks: defining a set of indicators capable of describing the state of the film production market from the point of view of the company, the external environment, and the target audience; development of algorithms for processing numerical and textual information; determining the list of target neural networks and revealing the specifics of their implementation; determination of the most effective prognostic approach by solving the problem of linear optimization.

The following methods used are — an analytical method for determining a set of neural networks; expert assessment for the formation of the most important independent indicators and determination of efficiency factors; experimental, multi-criteria evaluation to determine the most effective model.

The following results were obtained: a set of data reprocessing algorithms was formed for their longer use in recurrent and convolutional neural networks. Several common architectures involving MapReduce technology have been implemented. It was determined that the most effective model is a bidirectional recursive neural network with support for long- and short-term memory. The expediency of using parallelization technology is shown and a set of open questions for further research is defined.

Conclusions: conducting an analysis of algorithms for forecasting financial indicators based on artificial intelligence with subsequent experimental verification made it possible to form a relatively effective way of predicting the state of indicators of the film industry market under extraordinary conditions. The obtained results allow us to affirm the feasibility of implementing the proposed approach, which can influence the policy formation of the film producer's company or the fund that operates the company's financial instruments. At the same time, there are ways to further improve the results with the involvement of alternative approaches both to parallelization and to forecasting in general.

SENTIMENT ANALYSIS, NEURAL NETWORKS, NATURAL LANGUAGE PROCESSING, SOCIAL SHIFTS, FREQUENCY PATTERNS

Вступ

Протягом останніх років світова економіка поступово цифровізується. У дослідженні [1] автори всебічно аналізують це питання, зосереджуючись на тому, що, крім появи нових благ і послуг, розвиток технологій істотно вплинув на вже розвинені ринки. Зокрема, вони зазначають, що "комп'ютерна революція" сприяла появі нових бізнес-моделей, полегшила доступ до глобальних ринків і підвищила ефективність виробничих процесів. Одним з таких є субринок розважальних послуг — кіноіндустрія. З початку минулого століття цей вид дозволя швидко здобув популярність і поступово сформувався як один із найбільш прибуткових [2]. Завдяки інноваціям, таким як звукове кіно, кольорове зображення і спеціальні ефекти, кіноіндустрія постійно оновлювалася, залучаючи все більше глядачів. З появою стабільного і швидкого інтернет-з'єднання по всьому світу почали з'являтися нелегальні ресурси для перегляду фільмів, що створило нові виклики для цього ринку. Втрата доходів через піратство, розвиток соціальних мереж та загальне здешевлення технологій і кіновиробництва стимулювали появу легальних стрімінгових платформ, таких як Netflix, Amazon Prime, Disney+ чи HBO Max [3].

1. Аналіз останніх досліджень і публікацій

Основою подібних сервісів є підписна модель ціноутворення за якої люди платять за певний час користування. Ця модель має багато переваг як для компаній, так і для споживачів. Дослідники з університету міста Грац зазначають, що поширення цього підходу пояснюється можливістю компаній краще планувати свої трансферти, а споживачам — отримувати доступ до різноманітних контентів без значних одноразових витрат. Водночас, ця модель стимулює компанії постійно оновлювати і покращувати свої сервіси, щоб утримати підписників. Не зважаючи на подібні позитивні риси, вагомим недоліком для бізнесу є чутливість рівня доходів від поведінки споживачів. У випадку якихось надзвичайних ситуацій, як-от соціальних зсувів чи катастроф, відтік клієнтів може стимулювати процес банкрутства. У той же час при суттєвому збільшенні користувачів цифрові платформи можуть не витримати навантаження.

В якості підтвердження можна навести роботу групи азієських вчених, яка дослідила швидке зростання популярності стрімінгових сервісів у 2020 році під час активної фази пандемії COVID-19. Науковці довели, що подібна ситуація ускладнює аналіз інформації для поліпшення роботи зі споживачами і потребує обробки значних обсягів даних [4], оскільки соціальний зсув може впливати на поведінку цільової аудиторії і після її завершення. У той же час, пандемія спричинила зупинку виробництва нових фільмів,

що створило дефіцит контенту для стрімінгових платформ, змушуючи їх переглядати свої стратегії та інвестувати в оригінальний контент.

Крім того, сплески користувацької активності можуть виникати без соціальних катастроф, наприклад, у відповідь на певні новини або появу популярного фільму на платформі. Непрогнозовані ситуації можуть викликати перенавантаження, негативно впливаючи на користувацький досвід і зменшуючи ймовірність продовження підписки. Тому компанії змушені інвестувати в інфраструктуру та технології для забезпечення надійної роботи своїх сервісів. Прогнозування обсягів цільового ринку є складним завданням і потребує застосування сучасних підходів до інтелектуального аналізу даних

Загалом при прогнозуванні показників прийнято виділити наступні три великі групи алгоритмів [5, 6]:

- машинне навчання;
- ймовірнісний підхід;
- авторегресійний підхід.

Ймовірнісний підхід дозволяє отримати високу точність за обмежених обсягів даних, але вимагає значної кількості інформації про зовнішні фактори. Представники Корнелівського університету зазначають, що за обмеженої зовнішньої інформації точність цього підходу суттєво знижується [7]. Крім того, ймовірнісний підхід може бути менш надійним у разі раптових змін або кризових ситуацій, коли минулі дані можуть не відображати поточних тенденцій.

Авторегресійні моделі за своєю сутністю не мають вищезазначеної проблеми з обмеженістю вхідних даних. Це було доведено в ході досліджень проведених на базі університетів Шанхаю та Пекіна [8], а також ученими з Японії та Бангладешу [9]. Вони показали, що авторегресійні моделі можуть ефективно враховувати часові залежності і дозволяють отримувати точні прогнози навіть за умов нестабільності. Однак дослідження присвячені цьому класу моделей вже існують у вітчизняних працях. Зокрема, українські дослідники показали, що авторегресія рухомого середнього забезпечує точність понад 95% за нестабільних економічних умов [10]. Ці моделі також дозволяють швидко адаптуватися до нових даних і враховувати зміни в поведінці споживачів.

Розглядаючи інші можливі варіанти вирішення споріднені з авторегресією, слід зазначити, що періоди соціальної нестабільності ускладнюють використання альтернативних методів. Група китайських учених описала подібну проблему для багатовимірної лінійної регресії [11]. Вони відзначили, що такі методи можуть бути ефективними в стабільних умовах, але їхня точність суттєво знижується за наявності непередбачуваних змін. З огляду на це, авторегресійні моделі виглядають найбільш перспективними для прогнозування в умовах соціальної нестабільності.

Стосовно машинного навчання, то подібний підхід може потребувати значних обсягів даних для досягнення точності прогнозування понад 95%, і є чутливим до викидів. Це показали економісти з Пакистану та Південної Кореї [12]. Однак подібні проблеми можна вирішити шляхом додаткової передобробки даних та їх часткової аугментації.

2. Визначення не вирішених раніше частин загальної проблеми. Мета роботи, завдання

Враховуючи вказане вище, можна зазначити, що питання застосування засобів машинного навчання, залишається відкритим для задачі прогнозування стану специфічних ринків послуг. Прикладом такого ринку можна вважати ринок стримінгових сервісів. Особливо потребує опрацювання проблема врахування поведінки цільової аудиторії за надзвичайних ситуацій, які здатні викликати ажіотажний рух попиту. Базуючись на описаному науковому міжнародному досвіді, можна визначити, що найбільш ефективними засобами штучного інтелекту для вирішення поставлених питань є нейронні мережі. Особлива увага варто приділити згортковим та рекурентним архітектурам, пристосованим для обробки великих обсягів числової та текстової інформації за рахунок можливості врахування контексту. У чинній роботі вирішено зробити акцент на чотирьох різних моделях:

- класична згорткова нейромережа (CNN);
- класична рекурентна нейромережа (RNN);
- рекурентна мережа з підтримкою довго- та короткострокової пам'яті (LSTM);
- двонаправлена рекурентна мережа з підтримкою довго- та короткострокової пам'яті (BiLSTM);

Однак треба зауважити, що вказані моделі потребують значних обсягів вихідної інформації а також часу для навчання моделі. Задля вирішення зазначених проблем можна використовувати принципи аугментації даних та паралелізму відповідно. Для останнього було вирішено використати технологію MapReduce, яка дозволяє розподілити навантаження між вузлами. Ефективність цього підходу була доведена групою українських вчених [13].

Метою роботи є дослідження ефективності використання засобів штучного інтелекту, зокрема рекурентних та згорткових нейронних мереж, для прогнозування фінансових показників ринку кіновиробництва. Задля досягнення цієї мети, визначено наступний перелік *завдань*:

- визначити набір базових характеристик ринку кіновиробництва та зовнішнього оточення і сформулювати алгоритми передобробки цих даних;
- розглянути існуючі архітектури рекурентних та згорткових нейромереж;
- визначити задачу багатокритеріального вибору задля порівняння запропонованих підходів;
- сформулювати план експерименту та провести

його для виокремлення найбільш ефективної моделі для досягнення поставленої мети.

3. Матеріали та методи

3.1. Обмеження та припущення

Перед наведенням опису нейронних мереж та алгоритмів передобробки, необхідно зазначити ряд обмежень, які накладаються на цільові показники та компанії в рамках чинного дослідження:

- тип компанії: ті, що надають послуги стримінгу кіно, наприклад: Netflix, Hulu, Amazon Prime;
- тип ціноутворення: сервіси, які притримуються моделі підписки з необмеженим використанням за фіксованою ставкою;
- поширеність: враховуючи описаний глобальний контекст, було вирішено розглядати лише ті платформи, які функціонують на декількох національних ринках; тому вітчизняні сервіси такі як Sweet.tv не враховуються;
- спеціалізація: лише платформи розраховані на аудиторії різного віку та гендеру, без неясності певної специфіки в контенті.

3.2. Вибір цільових показників

Задля вибору набору показників зовнішнього становища та алгоритмів передобробки було вирішено скористатися методом експертного оцінювання. Загалом опитування було проведено серед 200 IT-фахівців, менеджерів з поширення кінопродукції та менеджерів із впровадження інновацій з Харкова, Києва, Львова, Відня, Лісабона, Праги та Кракова. Анкета складалася з 3 блоків: цільові фактори фінансового стану компанії, фактори зовнішнього впливу та фактори цільової аудиторії. Для респондентів пропонувалося обрати з наданого переліку від 3 до 5 найбільш важливих індикаторів кожної групи. Якщо індикатор не набрав 10 балів (тобто хоча б 10 людей його не обрали), він відсіювався. Самі ж переліки формувалися на основі аналізу останніх міжнародних наукових публікацій [14, 15], що стосувалися прогнозування економічних показників діяльності бізнесу (без обмеження на кіноіндустрію чи ринок розважальних послуг).

Для цільових факторів фінансового стану компанії було обрано:

- індикатор фінансової стабільності компанії;
- ступінь впровадження інновацій;
- індекс інфляції на вартість підписки в період за 5 років;
- індекс збільшення наданого контенту;
- наявність лістингу на американській чи європейській біржі.

Зазначені показники є суто числовими, однак задля пом'якшення проблеми чутливості даних, було вирішено додати в загальний пайплайн з обробки операцію обрізки даних — занадто великі чи малі

показники відносно середнього будуть замінені на певні константи. Після додаткової консультації з експертами поза проведенням опитуванням, було вирішено здійснювати заміну всіх значень, які більші за 97 перцентиль чи менші за 3 перцентиль. Окрім цього вирішено застосувати операцію скейлінгу відносно стандартного відхилення та арифметичного середнього.

Серед індикаторів зовнішнього впливу було обрано наступні 5 показників:

- ступінь гостроти соціального зсуву визначений за алгоритмом, який було запропоновано групою українських вчених [10];

- ступінь монополізації ринку визначений за індексом Герфіндаля-Гіршмана;

- агрегований зважений показник стабільності національних економік в яких працює платформа, в якості вагів використовується обсяг доходу компанії в країні відносно загального значення;

- показник фінансової стабільності світової економічної системи.

Три останні показники розраховуються з використанням класичного математичного апарату. Перший же, використовує засоби обробки природних мов, аби сформувати профіль цільової надзвичайної ситуації і ступінь готовності компанії до можливого ажіотажного росту цін на комплементарні товари чи збільшення навантаження на цифрові ресурси. Треба зауважити, що навіть умовно незначні події для світу чи національних економік можуть ставати соціальним зсувом в рамках кіноіндустрії. Як приклад, слугує ситуація зі страйком спілки акторів США, що суттєво вплинуло на виробництво нових фільмів і серіалів, а відповідно, і на обсяги нового контенту на платформах.

Заключною групою індикаторів є ті, що стосуються цільової аудиторії:

- середній індекс платоспроможності населення по відношенню до ціни підписки на стримінгову платформу;

- показник настрою населення сформований на основі алгоритму пошуку частотних патернів серед 200 найбільш популярних новин у соціальних мережах, які стосуються відповідного сервісу;

- показник дезінформації, щодо роботи кіноплатформи.

Задля визначення показника настрою населення було вирішено використати алгоритм Априорі для пошуку частотних патернів. Вибір цього способу ґрунтувався на простоті алгоритму, легкості його імплементації за допомогою мови програмування Python та можливостям паралелізації. Він має всього чотири кроки:

- перший: визначити рівень підтримки для кожного елемента в тексті. Підтримкою називають

частоту появи елемента в певному наборі даних;

- другий: використавши обмеження на рівень підтримки відфільтрувати набір слів;

- третій: побудувати асоціативні правила на основі найбільш частих закономірностей;

- четвертий: відсортувати отримані значення у порядку спадання ліфту.

Нижче наведено фрагмент псевдокоду для реалізації алгоритму (рис. 1).

```
while Lk-1 is not empty
  Ck ← Apriori_gen(Lk-1, k)
  for transactions t in T
    Dt ← {c in Ck : c ⊆ t}
    for candidates c in Dt
      count[c] ← count[c] + 1
  Lk ← {c in Ck : count[c] ≥ ε}
```

Рис. 1. Псевдокод для реалізації алгоритму Априорі

Стосовно показника дезінформації, як і у випадку з настроєм, було вирішено здійснювати аналіз тексту 200 найбільш поширених новин. Визначення "фейковості" по відношенню до роботи певного сервісу здійснювалась на основі наступних особливостей:

- використання риторичних запитань у суспільно значущих темах: Лінгвістичні дослідження показують, що у офіційно-діловому та публіцистичному стилях, які використовуються засобами масової інформації, риторичні запитання застосовуються рідко;

- відсутність заперечних конструкцій для зниження когнітивного навантаження: В поєднанні з песимістично забарвленими словами, наприклад, заміна слова "негаразд" на "катастрофа", відсутність заперечних конструкцій знижує когнітивне навантаження на читача. Слід зазначити, що у подальших текстах ненормативна лексика буде виключена, оскільки вона ускладнює процес аналізу емоційного забарвлення;

- використання закликів і заохочень у неналежному контексті та надмірна кількість займенників: Імітація публіцистичного стилю досягається через заклики і заохочення, вжиті у неналежному контексті, а також через надмірну кількість займенників;

- висока частота використання коротких речень та граматичних помилок: Часте використання коротких речень і слів з граматичними помилками.

Хоча подібні характеристики не є вичерпними, як показують міжнародні дослідження [13], вони можуть слугувати апроксимуючою функцією до функції дезінформації.

Розглянувши особливості визначення ключових індикаторів кожної групи перейдемо до розгляду нейронних мереж, які були використані в ході дослідження.

3.3. Огляд архітектур нейронних мереж

Почнемо з рекурентних нейромереж (рис. 2). Їхня сутність полягає у використанні декількох послідовних прихованих шарів. Це дозволяє враховувати

результат попереднього шару в наступного. Подібну особливість прийнято називати короткостроковою пам'яттю.

В середині кожного прихованого шару відбувається оптимізація градієнту. Через це у випадку не передоброблених даних може виникнути проблема вибохового (того, що прямує до нескінченності) та зникаючого (того, що прямує до 0) градієнту. Задля уникнення подібної проблеми було вирішено використати вже вказані функції скейлінгу та обрізки.

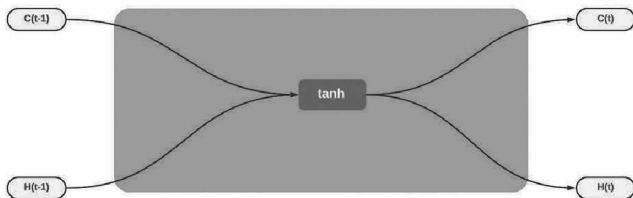


Рис. 2. Схематичне зображення прихованого шару архітектури RNN

Однак, з точки зору архітектурної складової можна використовувати більш просунуту модель з підтримкою довго- та короткострокової пам'яті (рис. 3) — її було вирішено обрати в якості другого алгоритму для дослідження. Сутність його математичного апарату полягає у поступовому використанні декількох сигмод та гіперболічних тангенсів, які дозволяють корегувати значення аргументу градієнтної функції. Послідовність складається з трьох ключових етапів — Forget Gate, Input Gate і Output Gate. На першому етапі відбувається множення сігма-функцію активації, що дозволяє отримати результат в межах від 0 до 1 і помножити його на значення з довгострокової пам'яті. Другий етап подібний до першого, однак включає в себе корекцію із застосуванням гіперболічного тангенсу. Подібна операція дозволяє нівелювати проблему спрямування до нескінченності при здійсненні додавання до значення з каналу довгострокової пам'яті. У результаті роботи двох вказаних етапів формується стан пам'яті, що разом із вхідними та попередніми вихідними даними, слугує базисом для визначення нового показника короткотривалої пам'яті на третьому етапі.

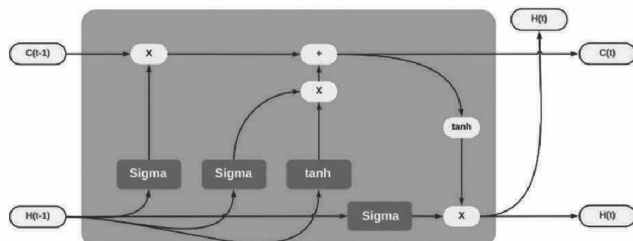


Рис. 3. Схематичне зображення прихованого шару архітектури LSTM

Хоча вказана архітектура дозволяє уникнути проблеми градієнтів, вона все ж має один істотний недолік — неможливість врахувати майбутній контекст. Зазначені вище цільові індикатори для кожної групи

є передобробленими числовими значеннями, проте для кращого розуміння вказаного недоліку розглянемо речення "Apple is something that...". Для обраної архітектури нейромережі невідомо, що саме мається на увазі під "Apple" — фрукт чи компанія, бо вона не має інформації про закінчення цього речення. Для неї варіанти "Apple is something that competitors simply cannot reproduce" та "Apple is something that I like to eat" є ідемпотентними.

Вирішення описаної проблеми з точки зору архітектури відбувається додавання другої мережі, яка спрямована в інший бік. Подібна модель називається Двонаправлена рекурентна нейромережа з підтримкою довго- та короткострокової пам'яті. Після відпрацювання двох підмереж результат обох рівнів поєднується, спочатку шляхом простої конкатенації, а після цього за допомогою лінійних трансформацій, визначених за допомогою крос-валідації. У нашому випадку — використання усереднених значень.

Перейдемо до заключної архітектури — згорткової нейромережі (рис. 4). У порівнянні з попереднім випадком, ця модель не запам'ятовує результати попередніх шарів. Натомість вона використовує шар згортки, що дозволяє суттєво зменшити розмірність вихідних даних. Редукція відбувається за допомогою проходження фільтра поміж даних і заміною цих даних на скалярний добуток значень між фільтром і даними.

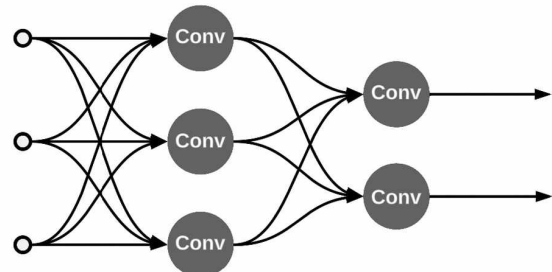


Рис. 4. Схематичне зображення архітектури CNN

Для задавання цієї моделі потрібні значення її гіперпараметрів встановлених в ході крос валідації, зокрема:

- було обрано фільтр розмірності $5 \times 5 \times 12$, де 12 — це кількість дескрипторів, тобто цільових змінних, які описані вище;
- глибина, як і випадку останнього показника розмірності фільтра має бути рівна кількості факторів — 12;
- розмір ядра (при крос-валідації було використано ядро в межах від 2 до 5 і встановлено оптимальне значення рівне 4);
- розмір кроку при розгляді. Виходячи із рекомендацій, які вказують на небажаність використання кроку більше 3 для тексту, було визначено оптимальне значення рівне 1;
- виходячи із встановленого кроку, параметр додавання неістотних нулів не застосовуватиметься;

— виходячи із специфіки предметної області було вирішено не застосовувати параметр зміщення.

Варто зауважити, що кількість шарів згорткової мережі для аналізу текстової інформації має бути рівна 1.

Проблемою зазначеної архітектури при її використанні для обробки природних мов є обмеженість у врахуванні контексту. Звичайно, проходження фільтра дозволяє взяти до уваги окіл проте часові межі певного соціального зсуву можуть бути істотними. Таким чином існує ймовірність, що визначний контекст знаходитиметься поза фільтром CNN моделі.

3.4. Розпаралелювання алгоритмів

Як зазначалося вище, для реалізації паралелізації було вирішено застосувати технологію MapReduce, яка базується на розподілі вихідного набору даних на окремі вузли. Виходячи з цієї концепції, ключовими компонентами моделі є функції мапінгу та редукції. Найбільш поширеними варіантами реалізації є платформи Spark та Hadoop. У цій роботі було обрано Hadoop, оскільки він включає додаткові функції всередині кожного вузла, що прискорює роботу з базами даних. Це є значною перевагою обраного підходу з огляду на велику кількість різноманітної інформації, що потребує обробки. Графічне зображення запропонованого рішення представлено нижче (рис. 5).

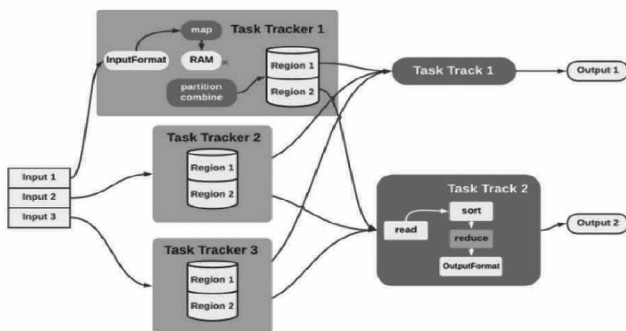


Рис. 5. Схематичне зображення MapReduce

У випадку передобробки даних усі показники трансформуватимуться паралельно один до одного на окремих вузлах. Особливості реалізації роботи з регіонами пам'яті пришвидшать розрахунок показника дезінформації та настрою населення. Зокрема для алгоритму Априорі ми розподіляємо зчитану з бази даних інформацію по різних блоках. Кожен із цих блоків виконує послідовну форму алгоритму. Далі об'єднувач агрегує значення в кожному блоці та передає їх редуктору, який, у свою чергу, фільтрує дані. Варто зазначити, що процес упорядкування, необхідний для коректної роботи алгоритму, виконується автоматично.

У випадку нейронних мереж використання паралелізації повноцінно доступне лише для BiLSTM, аби надати можливість працювати двом різнонаправленим мережам одночасно

Загалом запропонований підхід має наступний набір переваг:

- масштабованість, яка поліпшується завдяки розподілу навантаження по вузлах;
- відмовостійкість: MapReduce автоматично керує відмовами вузлів, що забезпечує надійність в обробці даних, оскільки при виникненні проблем навантаження буде перенесено на допоміжний вузол;
- простота програмування: у випадку Hadoop реалізації достатньо створити функції редукції та мапінгу а також додати конфігуруючий файл;
- суттєве покращення швидкодії: як вже було зазначено, технологія MapReduce дозволяє досягти пришвидшення в декілька разів.

Розглянувши ключові показники та алгоритми обробки й прогнозування перейдемо до формування плану експерименту.

3.5. План експерименту

Задля реалізації описаних алгоритмів було вирішено використати мову програмування Python3 і відповідні пакети:

- tensorflow — задля реалізації пайплайнів нейронних мереж;
- polars — задля структуризації даних та пришвидшення векторних операцій (обрізання, скейлінг);
- scikit-learn — задля реалізації кастомних трансформерів;
- nltk — задля конвертації текстових даних, зокрема знаходження частоти слів та інших операцій;
- numpy — задля допоміжної конвертації даних, оскільки на момент написання чинної статті polars підтримується не всіма зазначеними вище бібліотеками;
- datetime — для фіксації часу роботи.

Кількість вузлів MapReduce була встановлена на рівні 10, при цьому кожен мав наступну конфігурацію:

- частота процесора — 2.8 ГГц;
- кількість ядер — 1;
- оперативна пам'ять — 8 Гб;
- оточення — Linux Ubuntu 22.04.

Задля визначення найбільш ефективної моделі було вирішено обрати принцип лінійної адитивної згортки з ваговими коефіцієнтами. Хоча подібний підхід має певні недоліки пов'язані з можливістю покращення "малоефективних" альтернатив шляхом корекції лише одного із факторів, він є апікабельним за умови, що ці альтернативи є близькими за всіма критеріями.

Наступний кроком у формування плану експерименту є визначення критеріїв ефективності, як і у випадку з цільовими індикаторами було проведено експертне оцінювання (серед тої самої групи експертів). В наслідок чого обрані наступні фактори (сума показників важливості рівна 100):

— точність прогнозу — оскільки прогнозування фінансових показників прямо впливає на корпоративні рішення щодо політики компанії (або у випадку трейдингу на рішення сторонніх осіб щодо купівлі/продажу цінних паперів) цей показник було обрано найголовнішим у питанні ефективності запропонованого підходу із присвоєнням показника важливості рівним 70. Під точністю у цьому випадку мається на увазі обернений нормалізований показник кореня середньоквадратичної похибки (RMSE);

— економія часу роботи моделі — беручи до уваги специфіку показників, що розглядається та надзвичайність умов, наступний за важливістю є час роботи моделі. Варто зауважити, що поняття "час" не враховує тривалість тренування моделей, оскільки це прямо не впливає на швидкість прийняття рішень. Однак, передобробка даних перед передачею їх на вхід — враховується. Показник важливості цього фактора рівний 20;

— економія обсягу даних — оскільки кожна компанія на ринку кіновиробництва може мати свої особливі риси обсяг даних необхідний для досягнення точності у 80%. Показник важливості цього фактора рівний 10;

Враховуючи зазначене отримуємо наступні вагові коефіцієнти:

- точність прогнозу — 0.7;
- економія часу роботи моделі — 0.2;
- економія обсягу даних — 0.1.

Для подальшого дослідження було сформовано власний набір даних щодо Netflix та Amazon Prime. Усі описані вище характеристики були взяті або з соціальних мереж та новин з BBC (у випадку текстових описів) й з відкритих джерел, зокрема Statista та Yahoo Finance. Для кожного набору даних в якості таргету було вирішено обрати ціну закриття акцій компанії.

Заключним елементом плану експерименту є визначення можливих похибок. Виходячи із описаного плану можна виокремити наступні фактори, що здатні вплинути на результат: під час перевірки економії часу: людський фактор та інструментальна похибка; під час перевірки точності: проблема даних.

Аби пом'якшити зазначені невизначеності вимірювання показників здійснюватиметься декілька разів.

4. Результати досліджень

Почнемо з показника економії часу роботи алгоритмів. Результати наведені нижче (табл. 1). Усі значення економії порашовані відносно найповільнішого алгоритму — двонаправленої нейромережі з підтримкою довго- та короткострокової пам'яті. Варто зауважити, що різниця у часі з LSTM присутня не зважаючи на застосування паралелізації в наслідок накладних витрат з редукції даних.

Таблиця 1

Економія часу роботи алгоритма (у мілісекундах)

Спроба	CNN	RNN	LSTM	BiLSTM
I	58	37	11	0
II	55	55	12	0
III	57	48	10	0
IV	61	42	13	0
V	59	38	9	0
Середнє	58	44	11	0

Як бачимо алгоритм згорткової нейромережі працює швидше, оскільки має незначну глибину і містить в собі простіші алгоритмічні операції. Подібна логіка прослідковується і при порівнянні RNN з LSTM. Стосовно ефективності MapReduce підходу варто зауважити, що після декількох вимірювань було встановлено, що пришвидшення складає близько 1.89. Перейдемо до результатів точності прогнозування. У таблиці нижче (табл. 2) наведені обернене нормалізовані значення кореня середньоквадратичного відхилення нормалізованого в межах від 0 до 1. Подібна нормалізація можна застосовувати оскільки було прийнято рішення здійснювати скейлінг усіх незалежних та залежних змінних.

Таблиця 2

Точність прогнозування

Дані	CNN	RNN	LSTM	BiLSTM
Netflix	0.74	0.81	0.83	0.89
Amazon	0.72	0.79	0.84	0.91
Середнє	0.73	0.8	0.835	0.9

Треба зауважити, що використання MapReduce не вплинуло на точність BiLSTM. З огляду на отриманий результат саме BiLSTM є найефективнішою моделлю для обраного фактора.

Заключною метрикою є економія обсягу даних, необхідних для досягнення точності щонайменше 80%. Для цього було проведено кілька ітерацій з поступовим збільшенням кількості записів від 5000 до 10000 (оскільки дані були згенеровані на щоденній основі проводилась додаткова аугментація за допомогою векторної авторегресії). Результати наведено у таблиці нижче (табл. 3).

Таблиця 3

Економія обсягу даних для досягнення точності у 80%

Дані	CNN	RNN	LSTM	BiLSTM
Netflix	0	100	500	700
Amazon	0	100	600	600
Середнє	0	100	550	650

Тепер можемо систематизувати отримані значення метрик та визначимо альтернативи оптимальні за Парето (табл. 4). Усі значення були нормовані та округлені до сотих.

Таблиця 4

Значення критеріїв при аналізі аудіо оптимальних за Парето

Модель	Економія часу	Точність	Економія обсягу даних	Згортка
CNN	1.00	0.73	0.00	0.711
RNN	0.76	0.80	0.15	0.727
LSTM	0.19	0.84	0.85	0.711
BiLSTM	0.00	0.90	1.00	0.730

Враховуючи наведене вище, можна зазначити, що найбільш ефективною моделлю є двонаправлена нейромережа з підтримкою довго- та короткострокової пам'яті. Вона ж гарантує отримання найбільшою точності щодо прогнозування даних. Варто зауважити, що питання швидкодії цієї моделі можна вирішити за допомогою паралелізації внутрішніх розрахунків, зокрема визначення стану обох типів пам'яті. Важливим є також те, що модель LSTM виявилось гіршою за RNN. Однак це пов'язано з особливостями побудови чинної задачі з оптимізацією функції ефективності. Загалом отриманий результат відповідає міжнародним дослідженням проведеним для інших, не кіно, ринків.

Висновки

Метою чинної роботи було дослідження ефективності використання засобів штучного інтелекту, зокрема рекурентних та згорткових нейронних мереж, для прогнозування фінансових показників ринку кіновиробництва. Задля цього здійснено аналіз сучасних наукових публікацій, присвячених обраній тематиці, та проведено ряд експертних опитувань, що дозволило сформулювати набір цільових прогностичних алгоритмів і послідовність незалежних змінних. До останніх відноситься:

- індикатор фінансової стабільності компанії;
- ступінь впровадження інновацій;
- індекс інфляції на вартість підписки в період за 5 років;
- індекс збільшення наданого контенту;
- наявність лістингу на американській чи європейській біржі;
- ступінь гостроти соціального зсуву;
- ступінь монополізації ринку визначений за індексом Герфіндала-Гіршмана;
- агрегований зважений показник стабільності національних економік в яких працює платформа;
- показник фінансової стабільності світової економічної системи;
- середній індекс платоспроможності населення по відношенню до ціни підписки на стрімінгову платформу;
- показник настрою населення сформований на основі алгоритму пошуку частотних патернів;
- показник дезінформації, щодо роботи кіноплатформи.

У якості цільових моделей було розглянуто класичну згорткову та рекурентну нейромережу, а також більш просунуті версії останньої, зокрема, рекурентну нейромережу з підтримкою довго- та короткострокової пам'яті та її двонаправлену варіацію. Для передобробки описаних даних було застосовано алгоритм пошуку частотних патернів Априорі та принципи контент-аналізу.

Задля подолання проблеми пов'язаної з часом роботи та, частково, навчання моделі, було вирішено використати технологію MapReduce. Аби визначити найбільш ефективну нейронну мережу було сформовано набір критеріїв, який дозволив використати принцип лінійної адитивної згортки з ваговими коефіцієнтами.

У ході експериментальної перевірки із залученням даних щодо роботи Netflix та Amazon Prime виявлено, що двонаправлена рекурентна нейромережа з довго та короткостроковою пам'яттю є найбільш ефективною, не зважаючи на повільність по відношенню до інших моделей. При цьому застосування MapReduce було виправдним, наблизивши час роботи вказаного підходу до простої версії.

Таким чином можемо стверджувати, що використання BiLSTM та описаного набору незалежних змінних у поєднанні з алгоритмами їх передобробки і технологією MapReduce дозволяє ефективно прогнозувати фінансові показники діяльності компаній на кіноринку. Це в свою чергу відкриває можливості для корекції політики компанії та передбачення стану фінансового ринку. Відкритим залишається питання подальшого прискорення запропонованого підходу і можливість застосування більш просунутих гібридних моделей.

Список літератури

- [1] Bagnoli C., Biazzo S., Biotto G. Digital business models for Industry 4.0. How innovation and technology shape the future of companies. Springer, Cham. 2022. 268 с. DOI: 10.1007/978-3-030-97284-4.
- [2] Oyewola D. O., Dada E. G. Machine Learning Methods for Predicting the Popularity of Movies. *Journal of Artificial Intelligence and Systems*. 2020. Т. 4. С. 65–82. DOI: 10.33969/AIS.2022040105.
- [3] Wang W., Guo Q. Subscription strategy choices of network video platforms in the presence of social influence. *Electronic Commerce Research*. 2021. Т. 23. С. 577–604. DOI: 10.1007/s10660-021-09504-w.
- [4] Shin Z., Moon J., Rho S. A Comparative Analysis of Ensemble Learning-Based Classification Models for Explainable Term Deposit Subscription Forecasting. *Journal of Society for e-Business Studies*. Т. 3 (26). С. 1–21. URL: <http://www.jsebs.org/jsebs/index.php/jsebs/article/view/457> (дата звернення: 08.07.2024).
- [5] E-commerce Sales Forecast Based on Ensemble Learning / C. Zhan. *IEEE International Symposium on Product Compliance Engineering-Asia*. 2020. С. 1–5. DOI: 10.1109/ISPCE-CN51288.2020.9321858.

- [6] Masini R. P., Medeiros M. C., Mendes E. F. Machine learning advances for time series forecasting. *Journal of Economic Surveys*. 2021. T. 37 (1). С. 76–111. DOI: 10.1111/joes.12429.
- [7] Ning C., You F. Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers & Chemical Engineering*. 2019. T. 125. С. 434–448. DOI: 10.1016/j.compchemeng.2019.03.034.
- [8] Qi X.-Z., Ning Z., Qin M. Economic policy uncertainty, investor sentiment and financial stability—an empirical study based on the time varying parameter-vector autoregression model. *Journal of Economic Interaction and Coordination*. 2022. T. 17. С. 779–799. DOI: 10.1007/s11403-021-00342-5.
- [9] Shibasaki M., Witayangkurn A., Rahman M. M. Comparison of life patterns from mobile data in Bangladesh. *Smart Technology & Urban Development*, м. Chiang Mai, 13 груд. — 14 груд. 2019 р. 2019. С. 1–7. DOI: 10.1109/STUD49732.2019.9018795.
- [10] Parallelization of the VAR Algorithm Family to Increase the Efficiency of Forecasting Market Indicators During Social Disaster / A. Khovrat та ін. *Information Technology and Implementation*, м. Київ, 30 лист. — 2 груд. 2022 р. 2022. С. 222–233. URL: https://ceur-ws.org/Vol-3347/Paper_19.pdf (дата звернення: 08.07.2024).
- [11] Wang W., Guo Q. Subscription strategy choices of network video platforms in the presence of social influence. *Electronic Commerce Research*. 2021. T. 23. С. 577–604. DOI: 10.54691/bcpbm.v34i.3018.
- [12] A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector / I. Ullah та ін. *IEEE Access*. 2019. T. 7. С. 60134–60149. DOI: 10.1109/access.2019.2914999.
- [13] Yakovlev S., Khovrat A., Kobziev V. Using Parallelized Neural Networks to Detect Falsified Audio Information in Socially Oriented Systems. *Information Technology and Implementation*, м. Київ, 20 лист. — 21 лист. 2023 р. 2023. С. 220–238. URL: https://ceur-ws.org/Vol-3624/Paper_19.pdf (дата звернення: 08.07.2024).
- [14] Chen Y. C., Huang W. C. Constructing a stock-price forecast CNN model with gold and crude oil indicators. *Applied Soft Computing*. 2021. T. 112. № 107760. DOI: 10.1016/j.asoc.2021.107760.
- [15] Sousa, A. M. J., Braga, A. C., Cunha J. Impact of macroeconomic indicators on bankruptcy prediction models: case of the Portuguese construction sector. *Quantitative Finance and Economics*. 2022. T. 6 (3). С. 405–432. DOI: 10.3934/QFE.2022018.
- plainable Term Deposit Subscription Forecasting", *Journal of Society for e-Business Studies*, Vol. 26(3), P. 1–21, available at: <http://www.jsebs.org/jsebs/index.php/jsebs/article/view/457> (last accessed 08.07.2024).
- [5] Li, J., Zhan, C., Sha, W., Jiang, W., Guo, Y. (2020), "E-commerce Sales Forecast Based on Ensemble Learning", *IEEE International Symposium on Product Compliance Engineering-Asia*, P. 1–5. DOI: 10.1109/ISPCE-CN51288.2020.9321858.
- [6] Masini, R., Medeiros, M., Mendes, E. (2021), "Machine learning advances for time series forecasting", *Journal of Economic Surveys*, Vol. 37 (1), P. 76–111. DOI: 10.1111/joes.12429.
- [7] Ning, C., You, F. (2019), "Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming", *Computers & Chemical Engineering*, No. 125, P. 434–448. DOI: 10.1016/j.compchemeng.2019.03.034.
- [8] Qi, X.-Z., Ning, Z., Qin, M. (2022), "Economic policy uncertainty, investor sentiment and financial stability — an empirical study based on the time varying parameter-vector autoregression model", *Journal of Economic Interaction and Coordination*, Vol. 17, P. 779–799. DOI: 10.1007/s11403-021-00342-5.
- [9] Shibasaki, M., Witayangkurn, A., Rahman, M. M. (2019), "Comparison of life patterns from mobile data in Bangladesh". *Smart Technology & Urban Development (STUD-2019): 1st International Conference, Chiang Mai, 13 December — 14 December 2019: IEEE*. P. 1–7. DOI: 10.1109/STUD49732.2019.9018795.
- [10] Khovrat, A., Kobziev, V., Nazarov, A., Yakovlev, S. (2022), "Parallelization of the VAR Algorithm Family to Increase the Efficiency of Forecasting Market Indicators During Social Disaster". *Information Technology and Implementation (IT&I 2022): 9th International Conference, Kyiv, 30 November — 2 December 2022: CEUR Workshop Proceedings*. No. 3347, P. 222–233. available at: https://ceur-ws.org/Vol-3347/Paper_19.pdf (last accessed: 08.07.2024).
- [11] Wang, G., Wang, Zh., Xie, Y. (2022), "Subscribers forecasting of netflix based on multiple linear models", *BCP Business & Management*, Vol. 34, P. 229–236. DOI: 10.54691/bcpbm.v34i.3018.
- [12] Ullah, I., Raza, B., Malik, A. K., Imran, M., Islam, S. U., Kim, S. W. (2019), "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector", *IEEE Access*, Vol. 7, P. 60134–60149. DOI: 10.1109/ACCESS.2019.2914999.
- [13] Yakovlev, S., Khovrat, A., Kobziev, V. (2023), "Using Parallelized Neural Networks to Detect Falsified Audio Information in Socially Oriented Systems". *Information Technology and Implementation (IT&I 2023): 10th International Conference, Kyiv, 20 November — 21 November 2023: CEUR Workshop Proceedings*. No. 3624, P. 220–238. available at: https://ceur-ws.org/Vol-3624/Paper_19.pdf (last accessed: 08.07.2024).
- [14] Chen, Y. C., Huang W. C. (2021), "Constructing a stock-price forecast CNN model with gold and crude oil indicators", *Applied Soft Computing*, Vol. 112, Article 107760. DOI: 10.1016/j.asoc.2021.107760.
- [15] Sousa, A. M. J., Braga, A. C., Cunha J. (2022), "Impact of macroeconomic indicators on bankruptcy prediction models: case of the Portuguese construction sector", *Quantitative Finance and Economics*, Vol. 6 (3), P. 405–432. DOI: 10.3934/QFE.2022018.

References

- [1] Bagnoli, C., Biazio, S., Biotto, G. (2022), *Digital business models for Industry 4.0. How innovation and technology shape the future of companies*, Springer, Cham. 268 p. DOI: 10.1007/978-3-030-97284-4
- [2] Oyewola, D. O., Dada, E. G. (2022), "Machine Learning Methods for Predicting the Popularity of Movies", *Journal of Artificial Intelligence and Systems*, Vol. 4, P. 65–82. DOI: 10.33969/AIS.2022040105.
- [3] Wang, W., Guo, Q. (2021), "Subscription strategy choices of network video platforms in the presence of social influence", *Electronic Commerce Research*, Vol. 23, P. 577–604. DOI: 10.1007/s10660-021-09504-w.
- [4] Shin, Z., Moon, J., Rho, S. (2021), "A Comparative Analysis of Ensemble Learning-Based Classification Models for Ex-