

УДК 004.8

Д. А. Ткачик, Р. Н. Кветний

АСИМЕТРИЧНИЙ АНАЛІЗ ТОНАЛЬНОСТІ ДІЛОВИХ НОВИН В СИСТЕМАХ ПРОГНОЗУВАННЯ ДАНИХ НА ФІНАНСОВИХ РИНКАХ

Вінницький національний технічний університет, Вінниця

Анотація. Прогнозування даних на фінансових ринках - актуальне завдання у сучасному світі. Здатність передбачити напрямки руху ринку допомагає інвесторам уникнути очевидних ризиків і позбавити себе додаткових витрат. Було розроблено багато різних торгових платформ, щоб швидко отримувати доступ до великих обсягів історичних даних, що дозволяє аналізувати фінансовий ринок з будь-якого куточка планети в режимі реального часу, використовуючи лише ноутбук або персональний комп'ютер. Такі платформи дозволяють розробляти унікальні стратегії та підходи на основі фундаментального або технічного аналізу, які враховують новини про певну компанію, її прибуток, капіталізацію та кількість дивідендів, які вона повинна виплачувати вчасно. Ділові новини є важливим джерелом інформації про стан економіки та ринків. Вони можуть використовуватися для прогнозування майбутніх подій. Одним із методів прогнозування на основі ділових новин є аналіз тональності. Аналіз тональності дозволяє оцінити позитивність або негативність ділових новин. Традиційні методи аналізу тональності використовують симетричний підхід. Це означає, що позитивні та негативні новини однаково враховуються при прогнозуванні. Однак у реальному світі позитивні новини можуть мати більший вплив на ринки, ніж негативні новини. Це пов'язано з тим, що позитивні новини можуть стимулювати економічну активність, а негативні новини можуть її гальмувати.

У статті розглядається застосування асиметричного аналізу тональності ділових новин у системах прогнозування фінансових ринків. Проаналізовано різні методи аналізу тональності ділових новин, їхні переваги та недоліки. Запропоновано новий підхід до аналізу тональності ділових новин, який комплексно використовує штучні нейронні мережі та метод головних компонент.

Ключові слова: історичні дані, ділові новини, аналіз тональності, прогнозування асиметрія, фундаментальний аналіз

Abstract. Forecasting data in financial markets is a pertinent task in the modern world. The ability to predict market movements helps investors avoid obvious risks and spare themselves additional expenses. Numerous trading platforms have been developed to quickly access extensive historical data, enabling real-time analysis of the financial market from any corner of the planet using only a laptop or personal computer. Such platforms allow the development of unique strategies and approaches based on fundamental or technical analysis, taking into account news about a particular company, its earnings, capitalization, and the amount of dividends it is expected to pay on time. Business news is a crucial source of information about the state of the economy and markets. They can be used for forecasting future events.

One method of forecasting based on business news is sentiment analysis. Sentiment analysis allows assessing the positivity or negativity of business news. Traditional sentiment analysis methods employ a symmetric approach. This means that positive and negative news are equally considered in forecasting. However, in the real world, positive news may have a greater impact on markets than negative news. This is because positive news can stimulate economic activity, while negative news may hinder it.

The article explores the application of asymmetric sentiment analysis of business news in financial data forecasting systems. Various methods of sentiment analysis of business news, their advantages, and disadvantages are analyzed. A new approach to sentiment analysis of business news is proposed, which comprehensively utilizes artificial neural networks and principal component analysis.

Key words: historical data, business news, sentiment analysis, asymmetry forecasting, fundamental analysis

DOI: <https://doi.org/10.31649/1999-9941-2023-58-3-65-75>.

Вступ

Ділові новини є важливим джерелом інформації про стан економіки та ринків. Вони можуть використовуватися для прогнозування майбутніх подій. Одним із методів прогнозування на основі ділових новин є аналіз тональності. Аналіз тональності дозволяє оцінити позитивність або негативність ділових новин. Традиційні методи аналізу тональності використовують симетричний підхід. Це означає, що позитивні та негативні новини однаково враховуються при прогнозуванні. Однак у реальному світі позитивні новини можуть мати більший вплив на ринки, ніж негативні новини. Це пов'язано з тим, що позитивні новини можуть стимулювати економічну активність, а негативні новини можуть її гальмувати [1].

Асиметричний аналіз тональності враховує асиметрію тональності, яка вказує на нерівномірність вираження позитивних та негативних відтінків в тексті. У контексті аналізу текстів, таких як новини або відгуки, асиметрія тональності виражається у тому, що кількість та інтенсивність позитивних або негативних висловлювань неоднакові. При цьому позитивні новини враховуються з більшим вагомистію, ніж негативні новини. Це дозволяє поліпшити точність прогнозування.

Нижче наведено основні етапи асиметричного аналізу тональності ділових новин (рисунок 1).

Метою роботи є підвищення точності прогнозування фінансових ринків, з використанням аналізу тональності ділових новин

Результати досліджень

Асиметричний аналіз тональності ділових новин - це метод прогнозування майбутніх подій, який враховує асиметрію впливу позитивних і негативних новин на ринки.



Рисунок 1 - Основні етапи асиметричного аналізу тональності ділових новин

Традиційні методи аналізу тональності використовують симетричний підхід, тобто позитивні та негативні новини однаково враховуються при прогнозуванні. Однак у реальному світі позитивні новини можуть мати більший вплив на ринки, ніж негативні новини. Це пов'язано з тим, що позитивні новини можуть стимулювати економічну активність, а негативні новини можуть її гальмувати. Асиметричний аналіз тональності враховує цю асиметрію. При цьому позитивні новини враховуються з більшим вагомистю, ніж негативні новини. Це дозволяє поліпшити точність прогнозування [2].

Асиметричний аналіз тональності ділових новин може бути використаний для прогнозування різних економічних показників, таких як фондовий ринок, курс валют та інфляція.

Асиметричний аналіз тональності ділових нових складається з наступних етапів:

1. Збір даних: зібрати дані ділових новин за певний період часу.
2. Аналіз тональності: оцінити позитивність або негативність кожної новини.
3. Врахування асиметрії: врахувати асиметрію тональності при прогнозуванні.
4. Прогнозування: зробити прогноз майбутніх подій.

На першому етапі необхідно зібрати дані ділових новин. Ці дані можна зібрати з різних джерел, таких як веб-сайти новин, соціальні мережі та агенції новин.

На другому етапі необхідно оцінити позитивність або негативність кожної новини. Це можна зробити за допомогою різних методів, таких як:

- Лексичний аналіз: використання словників для визначення позитивних та негативних слів.

- Статистичний аналіз: використання статистичних методів для визначення позитивності або негативності тексту.

- Штучний інтелект: використання штучних інтелектуальних методів, таких як машинне навчання та штучні нейронні мережі.

На третьому етапі необхідно врахувати асиметрію тональності при прогнозуванні. Це можна зробити за допомогою різних методів, таких як:

- Використання вагових коефіцієнтів: позитивні новини можна враховувати з більшими ваговими коефіцієнтами, ніж негативні новини.

- Використання спеціальних алгоритмів: розробити спеціальні алгоритми прогнозування, які враховують асиметрію тональності.

На четвертому етапі необхідно зробити прогноз майбутніх подій. Це можна зробити за допомогою різних методів, таких як:

- Лінійне прогнозування: використовувати лінійні моделі для прогнозування майбутніх значень.

- Нелінійне прогнозування: використовувати нелінійні моделі для прогнозування майбутніх значень.

- Глибоке навчання: використовувати штучні нейронні мережі для прогнозування майбутніх значень.

Існує кілька різних методів асиметричного аналізу тональності ділових новин.

Ось деякі з них:

- Метод розкладу на основні компоненти (МГК) МГК - це статистичний метод, який використовується для зменшення розмірності даних. У контексті аналізу тональності ділових новин МГК може бути використаний для виділення двох основних компонентів: позитивного тону та негативного тону. Метод штучних нейронних мереж

- Штучні нейронні мережі - це тип машинного навчання, який може бути використаний для навчання на складних даних. У контексті аналізу тональності ділових новин штучні нейронні мережі можуть бути навчені на наборі даних ділових новин, щоб визначати позитивні та негативні тони.

- Метод обробки природної мови (NLP) NLP - це галузь комп'ютерних наук, яка займається обробкою природної мови. У контексті аналізу тональності ділових новин NLP може використовуватися для виявлення позитивних та негативних слів і фраз у ділових новинах.

- Метод статистичного аналізу Статистичний аналіз може використовуватися для оцінки тона ділових новин за допомогою статистичних методів, таких як частотний аналіз або аналіз дисперсії.

У цій статті розглядається застосування штучних нейронних мереж (ШНМ) у комплексі з методом головних компонент (МГК) для асиметричного аналізу тональності ділових новин. МГК використовується для зменшення розмірності даних, що покращує продуктивність ШНМ [3].

Нейронні мережі — це комп'ютерні моделі, які імітують роботу нейронних систем у мозку. Вони використовуються для вирішення завдань у сферах розпізнавання образів, класифікації, прогнозування, та багатьох інших областях.

Тут ми розглянемо декілька типів нейронних мереж та їх особливості:

1. Персептрон (рисунк 2) — це найпростіший вид нейронної мережі, яка складається з входів, ваг, суматора та функції активації. Його основна функція — класифікація вхідних даних на два класи (наприклад, "так" чи "ні").

2. Рекурентні нейронні мережі (РНМ) - мають здатність зберігати інформацію в попередніх станах, що робить їх ефективними для роботи з послідовностями даних, такими як мовлення або часові ряди.

Рекурентні нейронні мережі (РНМ) є особливим класом нейронних мереж, які призначені для роботи з послідовнісними даними, такими як текст, часові ряди або мовлення. Одна з головних відмінностей РНМ від інших типів нейронних мереж полягає в їхній здатності утримувати та

використовувати інформацію з попередніх кроків часу. Важливим елементом РНМ є зворотні зв'язки, які дозволяють передавати інформацію від одного часового кроку до іншого [4].

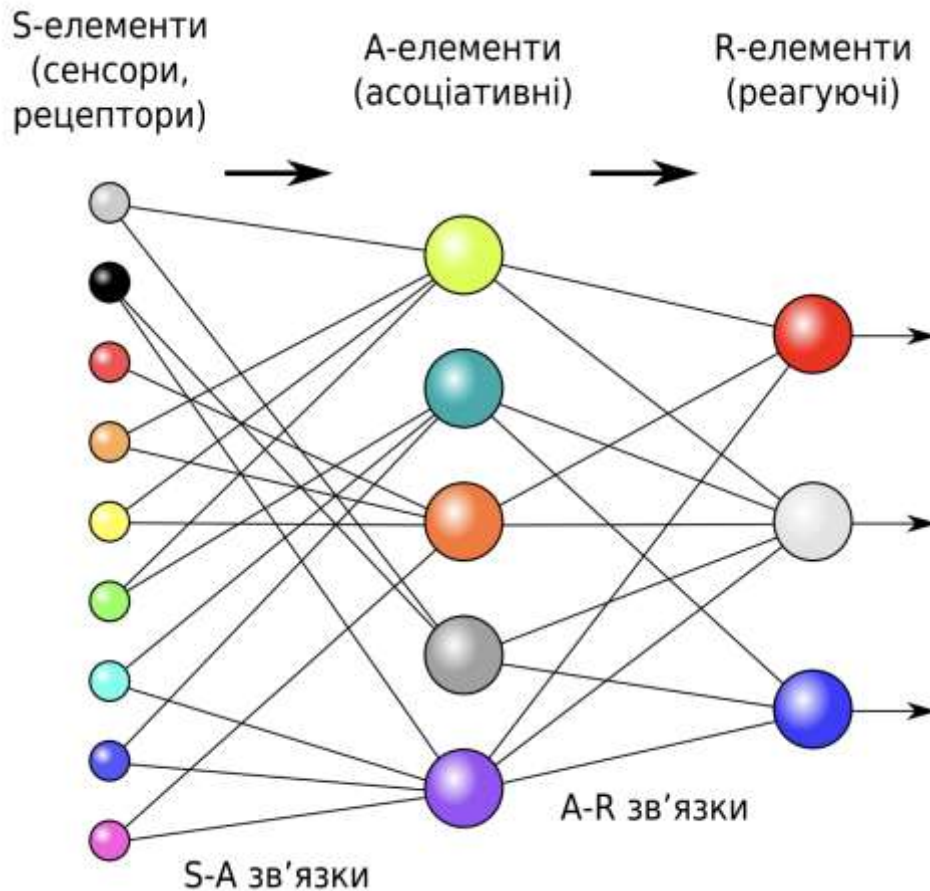


Рисунок 2 - Логічна схема перцептрону з трьома виходами

Основна ідея РНМ полягає в тому, що вона має внутрішні стани, які оновлюються на кожному кроці часу від вхідних даних та попереднього стану. Основна формула для одного часового кроку РНМ може бути виражена наступним чином:

$$h_t = \tanh(W_{hh}h_t - 1 + W_{xh}x_t + b_h), \quad (1)$$

де h_t - вектор прихованого стану на часі t , x_t - вектор вхідних даних на часі t , W_{hh} - матриця ваг для прихованого стану, W_{xh} - матриця ваг для вхідних даних, b_h - зміщення для прихованого стану, \tanh - гіперболічний тангенс, який може використовуватися як функція активації.

Ця формула показує, як оновлюється прихований стан h_t на кожному кроці часу, враховуючи попередній прихований стан h_{t-1} та вхідні дані x_t . Функція активації \tanh допомагає контролювати значення, які може приймати h_t , утримуючи їх у діапазоні від -1 до 1.

3. Згорткові нейронні мережі (ЗНМ) - азвичай використовуються для обробки зображень. Мають здатність до виявлення просторових шаблонів через згорткові шари.

Згорткові нейронні мережі є ефективними для обробки візуальної інформації, такої як зображення. Вони використовують операцію згортки для виявлення локальних патернів та особливостей відображення, що дозволяє їм автоматично вивчати властивості зображень [5]. Основні компоненти згорткових нейронних мереж включають в себе згорткові шари підсемплінгу, і повнозв'язані шари, що використовуються для класифікації.

Основна формула для операції згортки може бути виражена так:

$$C(i, j) = \sigma(\sum_m \sum_n I(i + m, j + n) \cdot K(m, n) + b), \quad (2)$$

де $C(i, j)$ - вихідний піксель після операції згортки, $I(i + m, j + n)$ - значення пікселя вхідного зображення, $K(m, n)$ - вага фільтра (ядро) для операції згортки, b - зміщення, σ - функція активації.

Операція згортки відбувається шляхом руху фільтра (ядра) по вхідному зображенню і обчислення взваженої суми значень пікселів [6]. Це дозволяє виявляти локальні патерни та особливості вхідного зображення. Шар підсемплінгу використовується для зменшення просторового розміру ознак та зниження кількості параметрів в мережі. Загальна формула для операції підсемплінгу може виглядати так:

$$P(i, j) = \max_{m, n}(I(s \cdot i + m, s \cdot j + n)), \quad (3)$$

де $P(i, j)$ - вихідний піксель після підсемплінгу, $I(s \cdot i + m, s \cdot j + n)$ - значення вхідного пікселя, s - фактор підсемплінгу.

Оскільки в даному дослідженні проводиться аналіз тональності новин (текстів), то доцільніше обрати саме рекурентні нейронні мережі.

Метод аналізу головних компонентів (МГК) є статистичним методом зменшення розмірності даних, зберігаючи при цьому їх варіативність. Основна ідея полягає в перетворенні оригінальних змінних у новий набір лінійно некорельованих змінних, які називаються головними компонентами.

Принцип дії МГК:

1. Стандартизація даних: Дані стандартизуються для забезпечення однакового масштабу.
2. Розрахунок матриці коваріацій: Визначається матриця коваріацій для оригінальних змінних.
3. Розрахунок головних компонент: Визначаються власні значення та власні вектори матриці коваріації, і вони використовуються для створення головних компонент.
4. Вибір головних компонент: Вибираються головні компоненти, які пояснюють найбільшу частку варіативності.

Використання МГК:

1. Зменшення розмірності даних: Використовується для візуалізації та аналізу даних.
2. Видалення кореляції між змінними: Допомагає уникнути мультиколінеарності у моделях.
3. Покращення ефективності моделей: Використовується для відбору важливих змінних та покращення роботи моделей машинного навчання.

Загальна схема математичних перетворень в МГК зображена на рисунку 3.

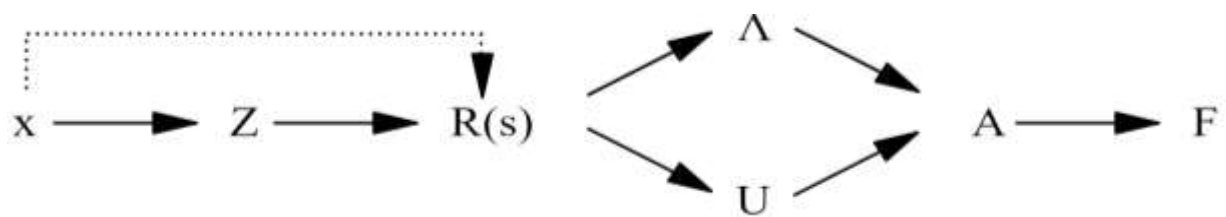


Рисунок 3 - Загальна схема математичних перетворень в МГК

На рисунку 3 позначення: X — матриця вихідних даних розмірністю $n \times m$ (n — число об'єктів спостереження, m — число елементарних аналітичних ознак); Z — матриця центрованих і нормованих значень ознак; R — матриця парних кореляцій. Λ — діагональна матриця власних (характеристичних) чисел. A — матриця факторного відображення, її елементи a_{ij} — вагові коефіцієнти. Спочатку A має розмірність $m \times m$ — за кількістю елементарних ознак X_j , потім в аналізі залишається r найвагоміших компонент, $r \leq m$. Обчислюють матрицю A за відомими даними матриці власних чисел Λ і нормованими власними векторами V за формулою $A = V\Lambda^{1/2}$. F — матриця значень головних компонент розмірністю $r \times n$.

Тексти новин для аналізу потрібно токенизувати на окремі слова або токени, і потім перетворити їх у вектори [7]. Один з підходів - використання словника, де кожне унікальне слово має відповідний

вектор. Це можна реалізувати за допомогою предтренованих векторів слів (наприклад, GloVe або Word2Vec) або з нуля, використовуючи ембедінги слів в нейронній мережі.

Модель для аналізу тональності може бути заснована на рекурентних нейронних мережах (RNN), довільних різновидностях Long Short-Term Memory (LSTM) або Gated Recurrent Unit (GRU), або трансформерних архітектурах (наприклад, BERT).

Оскільки моделі типу BERT мають обмеження, нам треба скорочувати розмір тексту, не втративши нічого корисного. У цій частині завдання є кілька підходів. Наведемо кілька можливих:

За допомогою синтаксичного дерева:

- використовуючи синтаксичний аналізатор із бібліотеки SpaCy, ми будемо синтаксичне дерево, у якого в кожному вузлі знаходиться слово або знак пунктуації;

- починаючи з знайденого вузла, що містить сутність, запускаємо на дереві «алгоритм забарвлення»: спочатку помічаємо вузли дерева на m рівнів вище знайденого, потім вузли на n рівнів нижче позначених (m і n - параметри, що налаштовуються) [8];

- всі помічені вузли об'єднуємо в текст, якщо між вузлами є перепустки, ставимо три крапки.

Простіший варіант: брати всі пропозиції, в яких згадується сутність, яка нас цікавить.

У новинах ми активно використовували набір бібліотек SpaCy як допоміжний інструмент:

- для отримання іменованих сутностей та їх нормалізації;

- розбиття статей на пропозиції;

- вилучення інформації з синтаксичного дерева для автоматичного виділення ключових іменованих сутностей та скороченого варіанта тексту.

Варто зазначити, що в тексті часто та сама сутність може зустрічатися у вигляді скорочень, псевдонімів або абревіатур. Це відома проблема різного написання назв організацій залежно від видавництва, настрою автора, погоди, дня тижня тощо.

Для конкретного контрагента, наприклад Монобанк, вдасться знайти у статтях безліч варіантів написання: «Моно», «Монобанк», «Банк Моно» тощо.

Для приведення сутностей до єдиної форми ми використовуємо словник можливих написань для кожної сутності та нечіткий пошук за допомогою бібліотеки FuzzyWuzzy [9]. Це також важливо для завдання матчінгу новини та бази контрагентів, за якими ми хочемо проводити аналіз.

Для кожної новини:

1. Виділяємо сутності та робимо оцінку тональності новини щодо кожної;

2. Наводимо різні варіанти вживання однієї сутності єдиного виду;

3. Робимо заміну ключової сутності на X ;

4. Утискаємо текст вихідної новини до 512 токенів щодо X .

У результаті ми отримуємо N^* прикладів та навчаємо BERT із додатковим вихідним шаром класифікації.

Підготовка даних і навчання моделі зображено на рисунку 4.

$$\text{Точність} = \frac{\text{Кількість правильних прогнозів}}{\text{Кількість всіх прогнозів}} \quad (4)$$

Для бінарної класифікації точність також може бути обчислена за допомогою критерію k щодо позитивних і негативних рейтингів, наступним чином:

$$k = \frac{TP+TN}{TP+TN+FP+FN}, \quad (5)$$

де TP - правильно визначені позитивні результати, TN - правильно визначені негативні результати, FP - неправильно визначені позитивні результати і FN - неправильно визначені негативні результати.

Отримані результати слід розглядати як обнадійливі. Отримана точність 0,7301. Якість моделі можна покращити за рахунок збільшення навчальної вибірки та підбору параметрів моделі. Незважаючи на помилки, модель помиляється в сусідні класи, при цьому явно не плутаючи негативні та позитивні приклади.

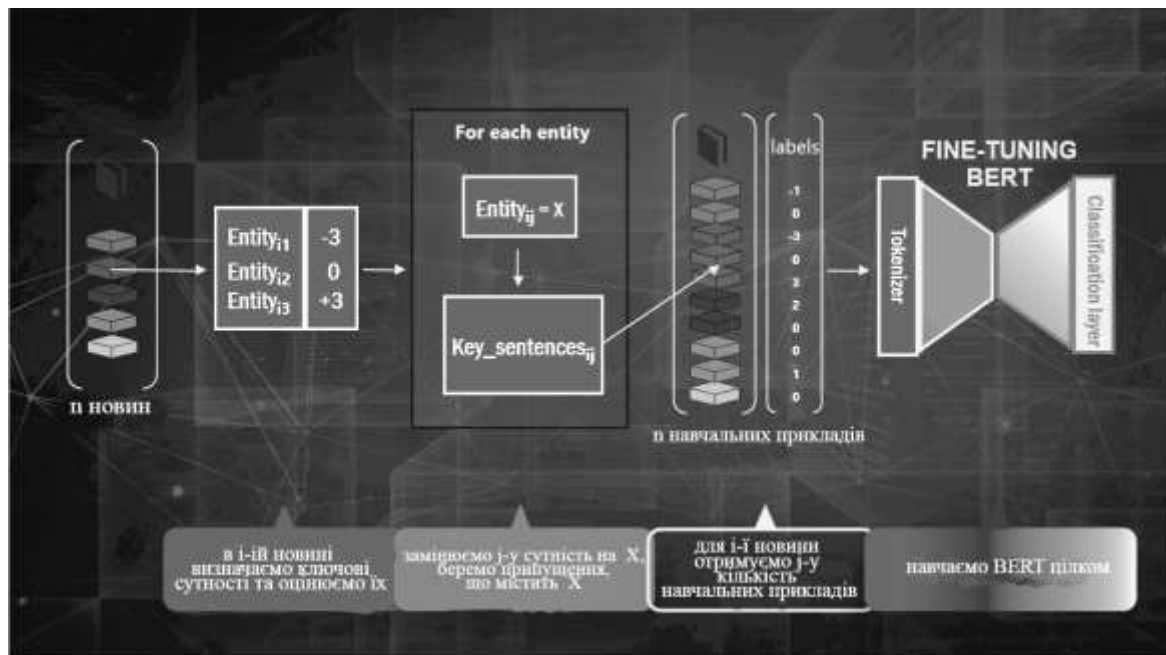


Рисунок 4 – Підготовка даних і навчання моделі BERT

Результати навчання моделі зображено на рисунку 5 у вигляді матриці помилок.



Рисунок 6 – Матриця помилок

Точність є одним із показників для оцінки моделей класифікації. Неформально, точність - це частка прогнозів, які модель вірно визначила. Формально точність має наступне визначення:

При використанні навченої моделі (рисунку 6) ми видаємо оцінки для кількох найважливіших сутностей новини. При цьому важливість визначається як зважена кількість згадок про сутність у тексті новини. Коли вона зустрічається як підмет, ми беремо її з великою вагою. Зрештою, ми просто відсікаємо всі неважливі сутності, частка згадки яких не перевищує заданий поріг.

Чи є суть підметом чи його частиною, ми визначаємо за допомогою синтаксичного дерева, що видається SpaCy.

Беремо новину та отримуємо ключові сутності.

Для кожної ключової сутності:

1. Наводимо різні варіанти вживання однієї сутності єдиного виду;
2. Робимо заміну ключової сутності на X;
3. Утискаємо текст прикладу до 512 токенів щодо X;
4. Подасмо в навчену модель;
5. Отримуємо можливості класів шкали тональностей, вибираємо найбільш можливий клас.

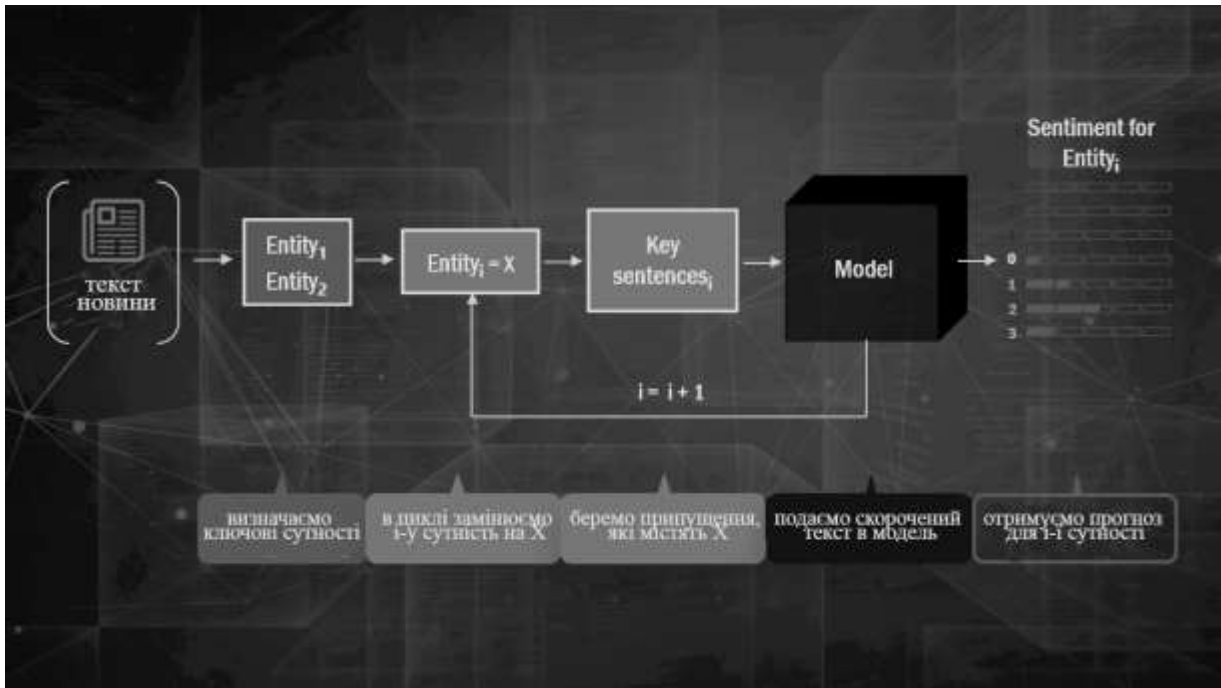


Рисунок 6 – Використання навченої моделі

Після отримання класів тональності за допомогою Штучних Нейронних Мереж (ШНМ), можна використовувати метод головних компонент (МГК), щоб зменшити розмірність цих класів та виокремити головні компоненти, які найкраще описують варіативність даних. Детальний процес буде описано нижче.

Нехай X буде матрицею, де кожний рядок представляє класів тональності для відповідної новини. Перш за все, слід стандартизувати дані, щоб забезпечити однаковий масштаб. Це важливо для вірного роботи МГК. Використовуємо бібліотеку `scikit-learn` для визначення головних компонент.

Після визначення головних компонент можна також оцінити, яку частку варіативності вони зберігають, що показує, яку частку загальної варіативності виявлено кожним з головних компонент. Можна також оцінити важливість кожного головного компонента за допомогою власних значень, які вказують на зміну варіативності, яку вони представляють.

Отримані приклади даних, що наведені в таблиці 1, можна використовувати як вхід для подальших моделей прогнозування, таких як класифікатори для прогнозування ринкової реакції на основі тональності новин.

Таблиця 1 – Приклад вихідних даних

Новина	Клас тональності	Головна компонента
1	0.8	0.7
2	-0.3	-0.5
3	0.1	0.2

В цілому комплексне використання штучних нейронних мереж та методу головних компонент має кілька переваг, які можуть зробити його кращим варіантом для аналізу тональності тексту та прийняття рішень на фінансових ринках порівняно з іншими методами:

- ШНМ відмінно пристосовані для аналізу складних залежностей у великих текстових даних, що може сприяти точнішій класифікації тональності, а комбінація ШНМ і МГК дозволяє створити комплексну модель, яка здатна краще виявляти та розуміти складні зв'язки між текстом і факторами, що впливають на тональність.

- МГК допомагає зменшити розмірність текстових даних, зберігаючи при цьому інформацію, що є важливою для аналізу. Зменшення розмірності дозволяє підвищити продуктивність моделі більше ніж на 50%, дозволяючи швидше обробляти великі об'єми текстової інформації та знижуючи витрати обчислювальних ресурсів.

- МГК дозволяє визначити головні компоненти, які найкраще виражають варіативність даних, забезпечуючи інтерпретованість зменшених даних. Інтерпретованість є важливою при прийнятті фінансових рішень, і цей метод забезпечує збереження важливих характеристик від зменшених даних.

- ШНМ адаптовані до змін у текстових даних, що може бути важливим для аналізу фінансових новин. Комплексне використання ШНМ та МГК дозволяє ефективно працювати з динамічними текстовими даними, що мають значний вплив на фінансові ринки.

- комбінація точної класифікації ШНМ та зменшення розмірності МГК дозволяє покращити прогнозування ринкової реакції на основі тональності новин. Застосування обох методів допомагає створити модель, яка краще адаптується до змін у тексті та забезпечує більш точні та надійні прогнози.

Таким чином описаний в статті підхід та використані в ньому методи збільшують продуктивність систем для прогнозування фінансових ринків. За рахунок використання штучних нейронних мереж та зменшення розмірності даних для аналізу зменшується час на прийняття рішень та підвищується точність прогнозів.

Висновки

Розвиток систем прогнозування даних є актуальним, оскільки в сучасному комп'ютеризованому світі значна кількість людей має доступ до Інтернету і володіє особистим комп'ютером, що надає швидкий доступ до фондових бірж та дозволяє переглядати котирування в режимі реального часу не залишаючи дім [10].

У даній статті розглянуто використання штучних нейронних мереж (ШНМ) та методу аналізу головних компонент (РСА) для аналізу тональності фінансових новин та прийняття рішень на фінансових ринках.

Процес прийняття рішення при цьому включає такі етапи:

1. Підготовка Даних: Зчитування та підготовка текстових та фінансових даних.
2. Використання ШНМ для Аналізу Тональності:
 - Тренування моделі (наприклад, LSTM або BERT) на текстах новин для аналізу тональності.
 - Отримання векторів тональності для кожної новини.
3. Стандартизація та Використання РСА:
 - Стандартизація отриманих векторів для однакового масштабу.
 - астосування РСА для зменшення розмірності та отримання головних компонент.
4. Прогноз Ринкової Реакції:
 - Використання зменшених даних в якості вхідних для моделі прогнозування ринкової реакції.
5. Прийняття Управлінських Рішень:
 - На основі прогнозу приймаються управлінські рішення, такі як оптимізація портфеля чи розробка стратегій.

Цей підхід дозволяє об'єднати сильні сторони ШНМ у роботі з текстовою інформацією та PCA для зменшення розмірності даних, полегшуючи подальший аналіз та використання в прийнятті рішень [11]. Для подальшого розвитку запропонованого підходу для досягнення найкращих результатів важливо налагоджувати параметри моделей та валідувати їх на реальних історичних даних.

Список літератури

- [1] Д. А. Ткачик, Р. Н. Кветний, "Розробка ефективних комбінацій моделей технічного аналізу для прогнозування ринку" Матеріали XLIX Науково-технічної конференції ВНТУ, Вінниця, 27-28 квітня 2020 р.. 2020. [Online]. Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2020/paper/view/9600>.
- [2] Goodfellow, I., Bengio, Y., Courville, A. (2016). "Deep Learning." MIT Press.
- [3] Tsai, C. F., & Wang, C. C. (2018). "Using deep neural network models for risk prediction and classification in real-time financial market." *Expert Systems with Applications*, 95, 12-27.
- [4] D. P. Kroese, S. J. Rubinstein, P. I. Frazier, et al. (2019). "Statistical Machine Learning for Asset Management." *Journal of Financial Data Science*, 1(1), 2-19.
- [5] Tsantekidis, A., Passalis, N., Tefas, A., et al. (2017). "Using deep learning and Google Trends to predict the stock price of a large IT company." *Expert Systems with Applications*, 73, 125-139.
- [6] Stock Market Sentiment Analysis in 2023, 2022. URL: <https://research.aimultiple.com/sentiment-analysis-stock-market/>
- [7] Types of Sentiment Analysis and How Brands Perform Them, 2020. URL: <https://www.analyticsinsight.net/types-of-sentiment-analysis-and-how-brands-perform-them/>
- [8] Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks / B. Liang et al. *Knowledge-Based Systems*. 2022. Vol. 235. P. 107643. URL: <https://doi.org/10.1016/j.knosys.2021.107643>
- [9] Zhang, Y., Watson, J., & Johnson, J. (2019). "Deep learning in finance." *Journal of Economic Dynamics and Control*, 98, 1-16.
- [10] Yu, L., Wang, S., & Lai, K. K. (2017). "A deep learning and cross-domain approach for stock price movement prediction." *Expert Systems with Applications*, 83, 56-66.
- [11] Hagenau, M., Liebmann, M., & Neumann, D. (2013). "Automated news reading: Stock price prediction based on financial news using context-capturing features." *Decision Support Systems*, 55(3), 685-697.

Стаття надійшла: 19.09.2023

References

- [1] D. A. Tkachyk, R. N. Kvyetnyy, "Rozrobka efektyvnykh kombinatsiy modeley tekhnichnoho analizu dlya prohozuvannya rynku" Materialy XLIX Naukovo-tekhnichnoyi konferentsiyi VNTU, Vinnytsya, 27-28 kvitnya 2020 r.. 2020. [Online]. Rezhym dostupu: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2020/paper/view/9600>.
- [2] Goodfellow, I., Bengio, Y., Courville, A. (2016). "Deep Learning." MIT Press.
- [3] Tsai, C. F., & Wang, C. C. (2018). "Using deep neural network models for risk prediction and classification in real-time financial market." *Expert Systems with Applications*, 95, 12-27.
- [4] D. P. Kroese, S. J. Rubinstein, P. I. Frazier, et al. (2019). "Statistical Machine Learning for Asset Management." *Journal of Financial Data Science*, 1(1), 2-19.
- [5] Tsantekidis, A., Passalis, N., Tefas, A., et al. (2017). "Using deep learning and Google Trends to predict the stock price of a large IT company." *Expert Systems with Applications*, 73, 125-139.
- [6] Stock Market Sentiment Analysis in 2023, 2022. URL: <https://research.aimultiple.com/sentiment-analysis-stock-market/>
- [7] Types of Sentiment Analysis and How Brands Perform Them, 2020. URL: <https://www.analyticsinsight.net/types-of-sentiment-analysis-and-how-brands-perform-them/>
- [8] Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks / B. Liang et al. *Knowledge-Based Systems*. 2022. Vol. 235. P. 107643. URL: <https://doi.org/10.1016/j.knosys.2021.107643>
- [9] Zhang, Y., Watson, J., & Johnson, J. (2019). "Deep learning in finance." *Journal of Economic Dynamics and Control*, 98, 1-16.

- [10] Yu, L., Wang, S., & Lai, K. K. (2017). "A deep learning and cross-domain approach for stock price movement prediction." *Expert Systems with Applications*, 83, 56-66.
- [11] Hagenau, M., Liebmann, M., & Neumann, D. (2013). "Automated news reading: Stock price prediction based on financial news using context-capturing features." *Decision Support Systems*, 55(3), 685-697.

Відомості про авторів

Ткачик Денис Анатолійович – аспірант кафедри автоматизації та інтелектуальних інформаційних технологій ВНТУ

Квєтний Роман Наумович – доктор технічних наук, професор, професор кафедри автоматизації та інтелектуальних інформаційних технологій ВНТУ

D.A. Tkachuk, R. N. Kvietnyi

**ASYMMETRIC SENTIMENT ANALYSIS OF BUSINESS NEWS IN HISTORICAL
DATA FORECASTING SYSTEMS**

Vinnitsia National Technical University