

ІДЕНТИФІКАЦІЯ ІНВЕСТИЦІЙНИХ РИЗИКІВ ЗА ДОПОМОГОЮ АНАЛІЗУ ЕМОЦІЙНОГО ОКРАСУ В ТЕКСТАХ

Вінницький національний технічний університет

Анотація

Проведено аналіз методів для визначення емоційного забарвлення в текстах. Проведено дослідження взаємозв'язків між настроями в текстах отриманих з мережі Твіттер та зміною цін на акції компанії Тесла. Розглянуто можливість ідентифікації інвестиційних ризиків на основі аналізу емоційного окрасу текстів.

Ключові слова: аналіз настрою тексту; нейронні мережі; ідентифікація ризиків; аналіз ризиків.

Abstract

An analysis of methods for determining the emotional color in the texts. A study was conducted on the relationship between sentiment in texts received from Twitter and changes in Tesla's share prices. The possibility of identifying investment risks based on the analysis of the emotional color of the texts is considered.

Keywords: analysis of emotional coloring of texts; neural networks; risk identification; investment risk analysis.

Вступ

Ідентифікація інвестиційних ризиків є однією з актуальних задач як для аналітиків так і для дослідників. Можливі ризики в поведінці ринку важко спрогнозувати через високу волатильність, яка залежить від різних політичних та економічних факторів. А прогнозування цін на акції на основі тільки історичних даних або текстової інформації здебільшого є недостатнім [1].

Настрій ринку — це якісна міра ставлення та настрої інвесторів до фінансових ринків загалом та конкретних секторів чи активів зокрема. Позитивні та негативні настрої впливають на поведінку ціни, а також створюють торговельні та інвестиційні можливості для активних трейдерів та довгострокових інвесторів. Застосування аналізу емоційного окрасу в текстах новин про відповідні ринки, сектори чи активи, може допомогти ідентифікувати нові ризики, які неможливо ідентифікувати аналізом історичних даних [2].

Метою роботи є дослідження взаємозв'язків між настроями ринку та зміною трендів в історичних даних.

Результати дослідження

Для дослідження взаємозв'язків фінансових ринків з емоційним забарвленням текстів новин пов'язаних з ними, джерелом текстів було обрано соціальну мережу Twitter, яка дозволяє отримати тексти останніх новин пов'язаних з певними секторами чи активами через відкритий API (рисунок 1).

1	id	created_at	text
2	1,53E+18	Thu May 19 15:03:16 +0000 2022	@JHochderffer As far as price increases we've seen in the EV space, it has been largely driven by material & trans... https://t.co/P9WZwdSb7e
3	1,53E+18	Thu May 19 15:03:13 +0000 2022	@elonmusk ... and just like that my next car will not be a Tesla, you spoiled brat.
4	1,53E+18	Thu May 19 15:03:12 +0000 2022	RT @Logically_JC: I've never seen a MAGA flag on a Tesla.
5	1,53E+18	Thu May 19 15:03:12 +0000 2022	RT @DeanObeidallah: Real world impact of Elon Musk's of not acting in best interest of Tesla: Since he went "alt-right" Tesla stock has dro...
6	1,53E+18	Thu May 19 15:03:11 +0000 2022	@PatriotJourney @DavMicRot Nah, there are many other options now. I will never buy a Tesla. They are good for the plant BYW.
7	1,53E+18	Thu May 19 15:03:11 +0000 2022	@elonmusk I'm so disappointed in you. You were doing a lot of good things for the world with your companies, and as... https://t.co/Wf6wsPuKIS
8	1,53E+18	Thu May 19 15:03:10 +0000 2022	RT @TheTeslaLife: Here's why #Tesla says your Dashcam and Sentry Mode aren't working right now https://t.co/q4y9LdUch9 https://t.co/TOfwVkg...
9	1,53E+18	Thu May 19 15:03:09 +0000 2022	RT @RonFilipkowski: Normally not my practice to spend large amounts of money on businesses with an owner who detests me. I personally think...
10	1,53E+18	Thu May 19 15:03:08 +0000 2022	Elon played the long game. Kissed libtard ass to get billions from gov to fund Tesla and is now using the profits t... https://t.co/MVBVX2jJOp
11	1,53E+18	Thu May 19 15:03:08 +0000 2022	RT @engineers_feed: On May 14, 1918, Nikola Tesla patented Lightning-Protector. https://t.co/3wEBdEyw0P
12	1,53E+18	Thu May 19 15:03:07 +0000 2022	@NCaDaPi @TheRickWilson @Tesla He's pissed all over the Tesla brand. And when its overpriced stock collapses furthe... https://t.co/PkMM7ZlQU8
13	1,53E+18	Thu May 19 15:03:07 +0000 2022	RT @YourAnonNews: us Houston Texas: The batteries onboard the Tesla continued to ignite despite efforts to douse the flames, authorities sa...
14	1,53E+18	Thu May 19 15:03:06 +0000 2022	RT @MayolsSpicy: Elon Musk is spiraling out of control faster than Tesla stock.
15	1,53E+18	Thu May 19 15:03:05 +0000 2022	RT @dutu2warn: Shares of Tesla have plunged again today, over 7% so far. Today's drop is more about the cars than Musk. Two analysts were c...

Рисунок 1 – Приклад даних отриманих з API Twitter з останніми новинами про компанію Tesla

Для проведення аналізу отриманих даних, необхідно відфільтрувати їх від спеціальних символів, посилань на сайти, відміток користувачів та відміток про відповідь. Виконати дану задачу можна використавши стандартні методи для роботи з рядковими типами даних, задавши шаблони для фільтрації відповідних випадків.

На сьогоднішній день існує багато методів для аналізу емоційного окрасу текстів. Вони базуються на нейронних мережах, що дозволяє максимально ефективно навчити модель для розпізнавання текстів конкретних обраних напрямків. Для дослідження взаємозв'язків між поведінкою фінансових ринків та емоційного забарвлення текстів пов'язаних з ними новин, було обрано попередньо підготовлену модель класифікатора текстів з бібліотеки Flair для Python. Даний метод дозволяє отримати оцінку тексту в вигляді значень «Позитивна» та «Негативна», а також ймовірність даної оцінки [3].

	A	B	C	D	E
1	id	created_at	text	probability	sentiment
2	1,53E+18	Thu May 19 15:03:16 +0000 2022	As far as price increases we've seen in the EV space, it has been largely c	0,887823164	NEGATIVE
3	1,53E+18	Thu May 19 15:03:13 +0000 2022	... and just like that my next car will not be a Tesla, you spoiled brat.	0,986812055	NEGATIVE
4	1,53E+18	Thu May 19 15:03:12 +0000 2022	I've never seen a MAGA flag on a Tesla.	0,926846743	NEGATIVE
5	1,53E+18	Thu May 19 15:03:12 +0000 2022	Real world impact of Elon Musk's of not acting in best interest of Tesla: S	0,999765456	NEGATIVE
6	1,53E+18	Thu May 19 15:03:11 +0000 2022	Nah, there are many other options now. I will never buy a Tesla. They ar	0,955424368	NEGATIVE
7	1,53E+18	Thu May 19 15:03:11 +0000 2022	I'm so disappointed in you. You were doing a lot of good things for the w	0,999955773	NEGATIVE
8	1,53E+18	Thu May 19 15:03:10 +0000 2022	Here's why #Tesla says your Dashcam and Sentry Mode aren't working ri	0,999961495	NEGATIVE
9	1,53E+18	Thu May 19 15:03:09 +0000 2022	Normally not my practice to spend large amounts of money on business	0,997812629	NEGATIVE
10	1,53E+18	Thu May 19 15:03:08 +0000 2022	Elon played the long game. Kissed libtard ass to get billions from gov to	0,759497285	POSITIVE
11	1,53E+18	Thu May 19 15:03:08 +0000 2022	On May 14, 1918, Nikola Tesla patented Lightning-Protector.	0,991738915	POSITIVE
12	1,53E+18	Thu May 19 15:03:07 +0000 2022	Tesla He's pissed all over the Tesla brand. And when its overpriced stock	0,999397397	NEGATIVE
13	1,53E+18	Thu May 19 15:03:07 +0000 2022	us Houston Texas: The batteries onboard the Tesla continued to ignite d	0,999818742	NEGATIVE
14	1,53E+18	Thu May 19 15:03:06 +0000 2022	Elon Musk is spiraling out of control faster than Tesla stock.	0,996204436	NEGATIVE
15	1,53E+18	Thu May 19 15:03:05 +0000 2022	Shares of Tesla have plunged again today, over 7% so far. Today's drop is	0,989903569	NEGATIVE

Рисунок 2 – Результат проведення аналізу емоційного забарвлення отриманих текстів

Для подальшого дослідження впливу отриманих оцінок на зміну тренду руху цін, необхідно завантажити історичні дані для даної компанії. Оптимальним джерелом для отримання історичних даних було обрано Yahoo Finance, який дозволяє отримувати дані з частотою від однієї хвилини.

Отримані оцінки емоційного забарвлення та історичні данні необхідно привести до однієї частоти та часового поясу, для порівняння руху ціни та тренду настрою ринку. Приведення таблиць до однієї частоти проводиться за допомогою обрахунку середнього значення всіх точок даних на кожному з діапазонів обраної частоти.

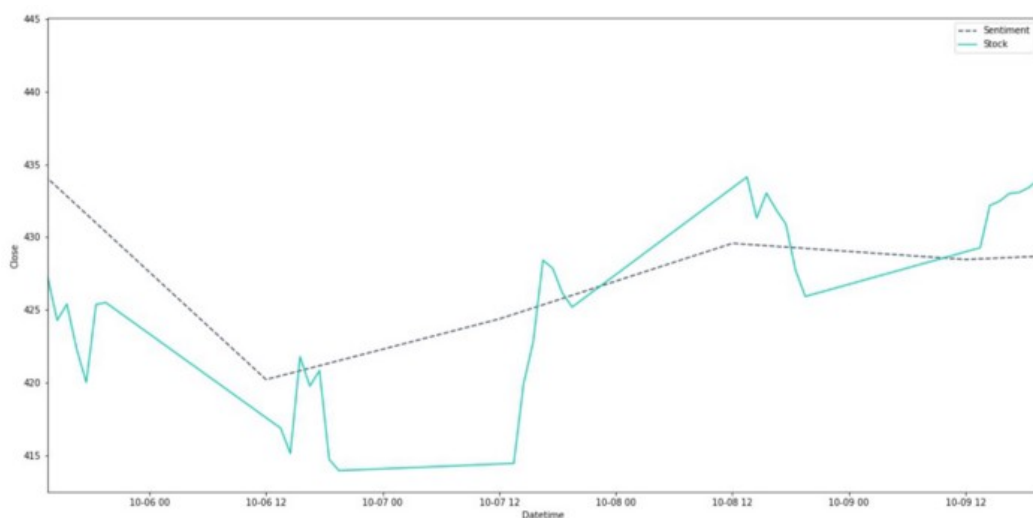


Рисунок 3 – Середнє значення оцінки емоційного забарвлення текстів, накладене на графік зміни цін

На рисунку 3 показано трансформований графік оцінок емоційного забарвлення текстів з соціальної мережі Twitter, накладений на графік зміни цін на акції даної компанії. Даний рисунок демонструє наявність прямо пропорційного зв'язку між настроями ринку до певної компанії та зміною ціни акції у відповідні проміжки часу.

Виходячи з результатів проведеного дослідження можна зробити висновок, що емоційне забарвлення текстів новин відносно фінансових ринків, має прямий вплив на них і може бути використане, як додатковий інструмент для ідентифікації прихованих ризиків, які не можливо ідентифікувати прогнозуванням на основі історичних даних.

Висновки

В даній роботі було проведено аналіз методів для визначення емоційного забарвлення текстів новин та розглянуто їх застосування для ідентифікації інвестиційних ризиків на фінансових ринках. Було досліджено вплив емоційного забарвлення текстів новин на ціну акції відповідної компанії та визначено, що оцінка емоційного забарвлення ринку має безпосередній вплив на тренд зміни цін і може застосовуватися як додатковий інструмент для ідентифікації прихованих ризиків.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Захарчук О. В. Розпізнавання емоційного забарвлення тексту для фундаментального аналізу на фінансових ринках [Електронний ресурс] / О. В. Захарчук, Д. А. Ткачик // Конференції ВНТУ електронні наукові видання. – 2021. – Режим доступу до ресурсу: <https://ojs.ukrlogos.in.ua/index.php/scientia/article/view/13531>.

2. Bellini T., Stress Testing and Risk Integration in Banks. / T. Bellini – Academic Press, 2016. – 316с

3. Іванов О. В. Класичний контент-аналіз та аналіз тексту: термінологічні та методологічні відмінності / Іванов Олег Валерійович // Вісник Харківського національного університету імені В. Н. Каразіна, Харків: Видавничий центр ХНУ імені В. Н. Каразіна, 2013. — № 1045. — С.72.

Захарчук Олександр Васильович — аспірант кафедри автоматизації та інтелектуальних інформаційних технологій, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, Вінниця, e-mail: olekmay@gmail.com;

Керівник: **Бісікало Олег Володимирович** — д-р. техн. наук, професор, завідувач кафедри АІТ, Вінницький національний технічний університет, м. Вінниця.

Zaharchuk Oleksandr V. — graduate student of the Department of Automation and Intelligent Information Technologies, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnitsa, e-mail: olekmay@gmail.com;

Supervisor: **Oleg V. Bisikalo** — Dr. Sc. (Eng.), Professor, Head of the Chair of Automation and Intellectual Information Technologies, Vinnytsia National Technical University, Vinnytsia.