

ЕМОЦІОНАЛЬНА ОЦІНКА НОВИН ДЛЯ ПРОГНОЗУВАННЯ КУРСУ БІТКОЇНУ

Наумов А. Б., Смеляков К. С.

Науковий керівник – к.т.н., доцент Чуприна А. С.

Харківський національний університет радіоелектроніки, каф. ПІ,
м. Харків, Україна

e-mail: anton.naumov.cpe@nure.ua

This article showcases the use of news sentiment analysis to predict Bitcoin prices. It offers a comprehensive analysis at the intersection of finance and machine learning. The research includes data collection and pre-processing, sentiment analysis, feature development, and model training. The article highlights the effectiveness of improved efficiency in identifying complex relationships in cryptocurrency markets. Tuning and estimating hyperparameters further improves the accuracy of the model. Although market volatility presents challenges, the results emphasize the importance of forecasting for investors and cryptocurrency researchers.

У світі фінансових ринків впевнене прогнозування курсів криптовалют є ключовим чинником для інвесторів і трейдерів. Одним з факторів, що впливають на динаміку цін на біткоїн, є громадська думка, виражена через новинні та соціальні медіа. Новини та публікації у соціальних медіа можуть мати значний вплив на психологію учасників ринку, формуючи тренди та напрямки руху цін. У цьому контексті аналіз емоціональних оцінок у джерелах новин набуває важливого значення для прогнозування курсу біткоїну.

Незважаючи на активні дослідження криптовалютних ринків, відносно мало робіт присвячено дослідженню впливу емоціональних оцінок новин на цінову динаміку біткоїну. У зв'язку з цим актуальність цього дослідження полягає у необхідності розробки моделей, здатних враховувати та аналізувати інформацію з джерел новин для прогнозування курсу біткоїну з високою точністю.

Метою даної роботи є дослідження використання методів аналізу емоціональних оцінок новин для прогнозування курсу біткоїну. Ми розглянемо методи збирання, обробки та аналізу даних, що використовуються для отримання емоціональних оцінок новин, а також розробимо модель прогнозування курсу біткоїну на основі цих даних та оцінимо її точність.

У цьому дослідженні використовувалася така методологія:

1. Збір даних:
 - збір даних зі стрічок новин Google і Yahoo (рисунок 1).
2. Обробка даних:

– очищення даних (видалення пунктуації, стоп-слів, нормалізація стемером та лематайзером, видалення цифр).

3. Аналіз емоціональних оцінок:

– поляризація сентиментів за допомогою бібліотек Vader та TextBlob (рисунок 2).

4. Отримання історичних даних для біткоїна:

– використання відкритого джерела <https://finance.yahoo.com/>

5. Навчання моделі з урахуванням отриманих метрик сентиметів:

– поділ даних на тренувальну та тестову частину.

– застосування бібліотеки TensorFlow та високорівневого API Keras для навчання моделі.

6. Оцінка метрик:

– використання метрик MAE та RMSE для оцінки точності.

Дана модель, заснована на нейронних мережах з використанням шарів LSTM [1], продемонструвала високу точність у прогнозуванні динаміки ціни біткоїну на основі емоціональних оцінок (рисунок 3), виражених у новинах (RMSE: train – 0.01009, test – 0.00876, MAE: train – 0.00840, test – 0.00783).

У порівнянні з прогнозуванням курсу біткоїна за допомогою моделі на базі алгоритму XGBoost [2], за допомогою даної методики було отримано на порядок вищу точність.

Цей успіх можна пояснити кількома факторами: використання глибокого навчання дозволяє моделі враховувати часові залежності даних, а також попередня обробка текстів та векторизація даних дозволили ефективно подати текстову інформацію у вигляді числових ознак, що покращило процес навчання моделі.

	Title	Source	Description	Link	Time
0	Bitcoin support levels to watch as BTC price a...	The Cointelegraph	Bitcoin bulls should be able to defend \$28,000...	https://cointelegraph.com/news/bitcoin-support...	4 hours ago
1	Bitcoin bulls risk trading range loss as BTC p...	The Cointelegraph	Bitcoin falls on Fed minutes, inflation outloo...	https://cointelegraph.com/news/bitcoin-bulls-r...	9 hours ago

Рисунок 1 – Формат даних зібраних зі стрічок новин

date	source	subject	text	title	url	Subjectivity	Polarity	Analysis
11/15/2022 11:00	CoinTelegraph	blockchain	roster full entertainment industry notable	Aether Games Sets Sights on Dark Fantasy DeFi	https://cointelegraph.com/news/aether-games-se...	0.362500	0.112500	1

Рисунок 2 – Формат даних після обробки та поляризації

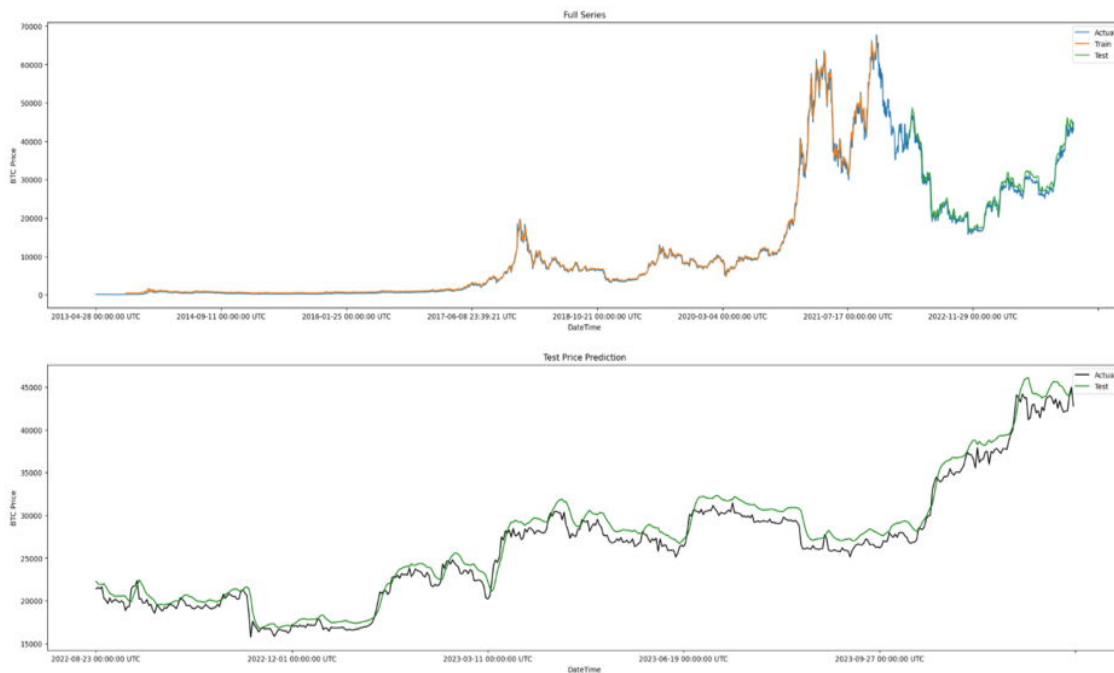


Рисунок 3 – Графіки реального та спрогнозованого курсу біткоїна, побудованого на основі аналізу емоціональних оцінок

На завершення, наше дослідження показало, що аналіз емоціональних оцінок новин із застосуванням інструментів машинного навчання, таких як нейронні мережі, що використовують TensorFlow та Keras, має значний потенціал для прогнозування курсу біткоїну. Було продемонстровано високу точність моделі у передбаченні ціни біткоїну на основі емоціональних оцінок, що підтверджує ефективність цього підходу. Однак, для подальшого розвитку цієї методики та підвищення можливостей її застосування у реальних умовах, необхідно продовжити дослідження у напрямку покращення застосованих методів обробки тексту, збільшення обсягів та різноманітності джерел даних та оптимізації моделей для роботи в реальному часі.

Список використаних джерел:

1. Smelyakov, K., Bizkrovnyi, O., Sharonova, N., Smelyakov, S., Chupryna, A. Building of Regression Models for Cryptocurrency Price Prediction CEUR Workshop Proceedings, 2022, 3171, pp. 1216–1232.

2. Afanasieva I.V., Naumov A.B., Onyshchenko K.G. Bitcoin price prediction using the boosting algorithm. The 2nd International scientific and practical conference “Science and society: modern trends in a changing world” (January 22-24, 2024) MDPC Publishing, Vienna, Austria. 2024. 662 p.