

Роман Руденський

д.е.н., професор, професор кафедри комп'ютерних наук
Національний університет біоресурсів і природокористування України, м. Київ, Україна
0009-0002-3682-9702
roman.rudensky@nubip.edu.ua

Володимир Кравченко

д.е.н., доцент, професор кафедри економічної кібернетики
Національний університет біоресурсів і природокористування України, м. Київ, Україна
0000-0002-8033-3985
v.kravchenko@nubip.edu.ua

АНАЛІЗ ТЕМ І НАСТРОЇВ У ЦИФРОВОМУ ДИСКУРСІ АГРАРНОГО СЕКТОРУ

Анотація. Дослідження присвячено виявленню ключових тем і наративів, що циркулюють у цифровому медіапросторі аграрного сектору України, а також дослідження емоційної тональності (позитивної, негативної, нейтральної) повідомлень у межах виявлених тематичних кластерів. Розроблено моделі аналізу тем і настроїв в цифровому медіапросторі щодо публічного управління в аграрному секторі України з використанням методів машинного навчання та попередньо навченої моделі з бібліотеки Hugging Face для Python.

Ключові слова: аналіз тем; аналіз настроїв; машинне навчання, цифровий дискурс.

1. ВСТУП

В епоху цифрової трансформації громадська думка відіграє ключову роль у формуванні різних політичних рішень, особливо в таких стратегічних галузях, як аграрний сектор. Разом з тим, в онлайн просторі вона має значний вплив на поведінку та прийняття рішень різними зацікавленими сторонами, зокрема фермерами, посередниками і споживачами [1].

Постановка проблеми. Виходячи з цього, виникає потреба у використанні передових технологій для моніторингу та прогнозування цифрового дискурсу. ІТ-рішення, які базуються на методах машинного навчання та великих мовних моделях (LLM), дозволяють ефективно оцінювати настрої та думки громадян, їх зміни в реальному часі, що дає змогу своєчасно змінювати політику органів державної влади різних рівнів, забезпечуючи її відповідність вимогам та потребам суспільства.

Аналіз останніх досліджень і публікацій. Аналіз великих даних із засобів масової інформації відкриває можливості для створення моделей, що прогнозують громадську думку щодо державної політики [2]. Такий підхід враховує вплив громадської думки в цифровому просторі та активність інтернет-користувачів, спрямовану на демократичну участь, що дозволяє виявляти приховані інсайти, пов'язані з впровадженням нових політичних рішень [3]. З метою підвищення громадської безпеки у соціальних мережах розробили систему моніторингу громадської думки на основі технологій машинного навчання та методів сегментації слів і впровадили її на платформі Sina Weibo [4].

Мета публікації. Метою статті є розробка моделей аналізу тем і настроїв в цифровому медіапросторі щодо публічного управління в аграрному секторі України.

2. МЕТОДИ ДОСЛІДЖЕННЯ

Корпус текстів було сформовано на основі публікацій з провідних українських Telegram-каналів, спеціалізованих на аграрній тематиці: Agro News, Agro Portal, Agro Review, Agro Polit, Super Agronom, Landlord і Latifundist. Перед початком аналізу тексти були технічно підготовлені: очищені від службової інформації (HTML-теги, символи),

токенізовані та лематизовані. Для ідентифікації тематичних кластерів в цифровому дискурсі аграрного сектору в Україні застосовано: методи тематичного моделювання (topic modelling) з використанням алгоритму Latent Dirichlet Allocation (LDA) з бібліотеки gensim; метрику когерентності (c_v) для вибору оптимальних параметрів (кількість тем, кількість проходів); багатовимірною шкалювання (MDS) для візуалізації семантичної близькості. Для аналізу емоційної тональності (sentiment analysis) застосовано: попередньо навчену модель з бібліотеки transformers (Hugging Face) для класифікації повідомлень на позитивні, негативні та нейтральні; ProcessPoolExecutor для паралельної обробки великих обсягів тексту. Варто зазначити, що було здійснено перетворення оцінок настроїв у числову шкалу від -1 (негативний) до +1 (позитивний) для кількісного аналізу.

3. РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ

3.1. Тематичне моделювання

Для визначення найкращих параметрів моделі випробовуються різні комбінації кількості тем (num_topics) і кількості проходів (passes). Найвищий показник когерентності (0.517) досягнуто при num_topics = 5 і passes = 50. Метод тематичного моделювання до зібраного корпусу текстів із використанням ймовірнісного алгоритму дозволив ідентифікувати п'ять провідних тематичних кластерів: Тема 0 "Новини українського агробізнесу: фінансування, розвиток, міжнародна інтеграція"; Тема 1 "Інтеграція технологій і агросистем: від захисту рослин до стратегій сталого землеробства"; Тема 2 "Врожайність сільськогосподарської продукції в Україні: ринкові та природні фактори"; Тема 3 "Український експорт агропродукції через морські та сухопутні коридори: виклики та можливості на тлі геополітичних змін"; Тема 4 "Державна політика й аграрні реформи в Україні".

3.2. Аналіз настроїв

Щомісячні зміни настроїв проілюстровані за допомогою діаграми "candlestick" (рис. 1). За допомогою цієї діаграми можна побачити зміну настрою з бурхливої волатильності (2016–2019 рр.) у стійке ослаблення (2022–2024 рр.). Це видно по довших "свічках" (тілах і тінях), що вказує на великі коливання настрою як протягом місяця (діапазон між 1-м і 3-м квантилем, показаний тінями), так і між початком та кінцем місяця (показано тілом свічки).

Довге тіло свідчить про сильну зміну, коротке – про незначну зміну. Колір тіла показує напрямок руху настрою за період: зелене – настрої в кінці місяця кращі ніж на початку, тобто оцінка зросла; червоне – настрої в кінці місяця гірші ніж на початку, тобто оцінка спала. Верхня тінь – лінія від верхньої межі тіла до 3-го квантилю оцінки настрою, тобто показує, наскільки високо піднімалася оцінка вище кінцевого рівня для зеленої свічки або початкового рівня для червоної свічки. Нижня тінь – лінія від нижньої межі тіла до 1-го квантилю оцінки настрою за місяць. Показує, наскільки низько вона опускалася нижче початкового рівня для зеленої свічки або кінцевого для червоної свічки. Довжина тіней вказує на волатильність настрою поза діапазоном початок-кінець. Довгі тіні свідчать про значні коливання настрою протягом періоду, які не втрималися до закриття. Тіні свічок (між квантилями 1 і 3) показують розкид настроїв протягом місяця. Навіть у місяці, коли загальна зміна настрою від початку до кінця місяця була невеликою (коротке тіло), могли бути значні коливання протягом місяця (довгі тіні).

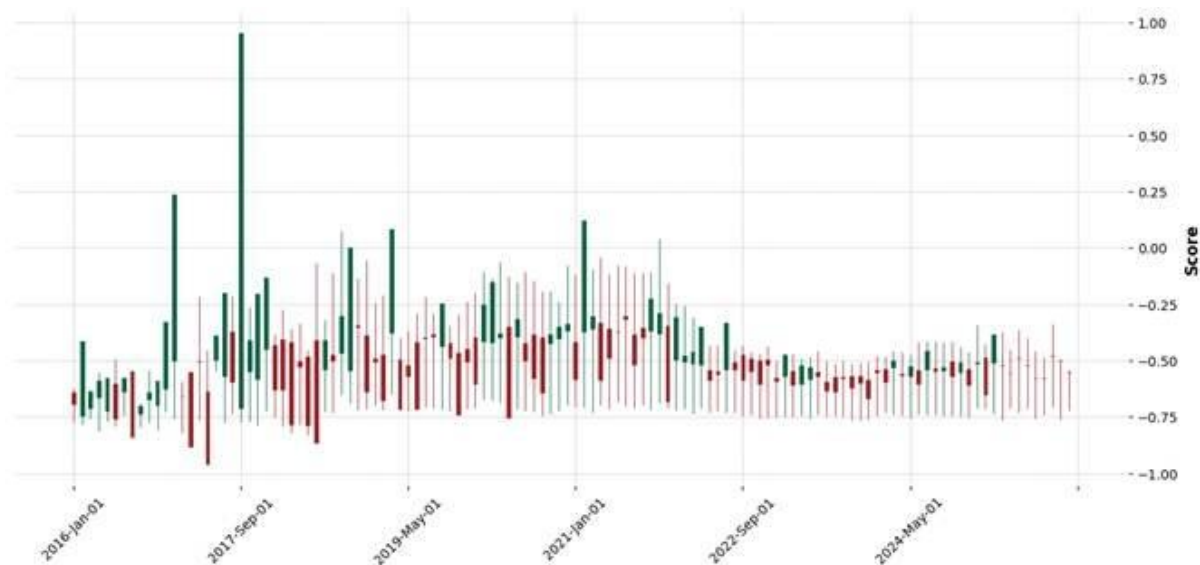


Рис. 1. Діаграма "candlestick" щомісячних змін настроїв

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Отже, дослідження одним з перших застосувало тематичний аналіз і аналіз настроїв до українських аграрних медіа. Результати розкривають роль етично спрямованої аналітики медіа в підвищенні прозорості, передбаченні суспільних занепокоєнь та підтримці науково обґрунтованого керування в умовах національної невизначеності. Виявлено 5 домінантних тематичних кластерів; найбільшу увагу отримали аграрні реформи та політика. Аналіз настроїв показав переважно негативний тон за всіма темами, особливо щодо експорту та стану ринку. Контент про аграрну політику демонстрував відносно більш позитивний настрій. Емоційний тон значно варіювався у часі, що свідчить про динамічні та залежні від теми суспільні позиції.

ПОСИЛАННЯ

- [1] Lv, X. et al. (2024) 'Spillover effect of network public opinion on market prices of small-scale agricultural products', *Mathematics*, 12(4), p. 539. doi:10.3390/math12040539.
- [2] Chen, X. (2021) 'Research on public opinion forecast of public policy based on Big Data Media', 2021 5th Annual International Conference on Data Science and Business Analytics (ICDSBA), pp. 103–106. doi:10.1109/icdsba53075.2021.00030.
- [3] Redjeki, S. and Widyarto, S. (2022) 'Comparison of seven machine learning algorithms in the classification of public opinion', *Tech-E*, 5(2), pp. 143–149. doi:10.31253/te.v5i1.1046.
- [4] Weidong, G. et al. (2022) 'Network public opinion monitoring method based on machine learning', 2022 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), pp. 178–181. doi:10.1109/icitbs55627.2022.00047.

MINISTRY OF EDUCATION
AND SCIENCE OF UKRAINE

NATIONAL UNIVERSITY
OF LIFE AND ENVIRONMENTAL
SCIENCES OF UKRAINE

FACULTY OF INFORMATION
TECHNOLOGY

МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ

PROCEEDINGS

XIII International scientific
and practical conference

**GLOBAL AND
REGIONAL PROBLEMS OF
INFORMATIZATION IN
SOCIETY AND
NATURE USING
'2025**

13-14 November 2025

Kyiv, NULES of Ukraine

Kyiv 2025

МАТЕРІАЛИ

XIII Міжнародної науково-
практичної конференції

**ГЛОБАЛЬНІ ТА
РЕГІОНАЛЬНІ ПРОБЛЕМИ
ІНФОРМАТИЗАЦІЇ В
СУСПІЛЬСТВІ І
ПРИРОДОКОРИСТУВАННІ
'2025**

13-14 листопада 2025 року

Київ, НУБіП України

Київ 2025

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

МАТЕРІАЛИ

ХІІІ Міжнародної науково-практичної конференції

ГЛОБАЛЬНІ ТА РЕГІОНАЛЬНІ ПРОБЛЕМИ ІНФОРМАТИЗАЦІЇ В СУСПІЛЬСТВІ І ПРИРОДОКОРИСТУВАННІ '2025

13-14 листопада 2025 року

Київ, НУБіП України

Київ 2025

УДК 004

Рекомендовано до друку вченою радою факультету інформаційних технологій Національного університету біоресурсів і природокористування України (протокол № 4 від 18.12.2025).

Укладач: д.т.н., доцент Шкарупило В.В.

Збірник матеріалів XIII Міжнародної науково-практичної конференції "Глобальні та регіональні проблеми інформатизації в суспільстві і природокористуванні '2025", 13–14 листопада 2025 року, НУБіП України, Київ. – К.: НУБіП України, 2025. – 206 с.

Відповідальність за зміст публікацій несуть автори.

© Національний університет біоресурсів
і природокористування України, 2025