

Abstract: *The paper examines the use of AI-based chatbots in agriculture for providing informational support to farmers. Their role in delivering advisory services, collecting field data, and assisting in monitoring food security risks is considered. The main advantages and limitations of implementing such systems in the agricultural sector are outlined.*

Науковий керівник:

Ємельянов С. І.,

*викладач кафедри економічної кібернетики, комп'ютерних наук та інформаційних технологій,
Миколаївський національний аграрний університет*

УДК 004.4:631.95

Виявлення маніпулятивного контенту про аграрну галузь на основі аналізу метаданих Youtube

Заярнюк Нікіта,

здобувач вищої освіти спеціальності 122 «Комп'ютерні науки»

Миколаївський національний аграрний університет,

м. Миколаїв, Україна

Анотація: *У роботі досліджено проблему поширення маніпулятивного контенту щодо аграрної галузі на відеоплатформі YouTube. Запропоновано підхід до автоматизованого виявлення маніпуляцій на основі аналізу текстових метаданих відео. Для збору даних використовується інструмент yt-dlp (із інтеграцією API "Return YouTube Dislike"), для транскрибування аудіоконтенту – модель Whisper, для аналізу тексту – методи обробки природної мови та машинного навчання. Запропонований підхід дозволяє здійснювати автоматизований моніторинг інформаційного простору та сприяти підвищенню інформаційної безпеки аграрного сектору.*

Ключові слова: *маніпулятивний контент, YouTube, аграрна галузь, метадані, yt-dlp, OpenAI Whisper, обробка природної мови, машинне навчання, інформаційна безпека.*

Сучасна архітектура аграрного сектору України зазнає докорінних трансформацій під впливом глобальних викликів та інтеграції передових інформаційних технологій. В умовах воєнного стану, де кожна одиниця інформації може впливати на ринкову стабільність, продовольчу безпеку та міжнародні торговельні відносини, захист інформаційного простору набуває критичного значення. Використання відеоплатформи YouTube як потужного інструменту інформаційно-комунікаційних процесів створює можливості для поширення освітніх матеріалів, новин та аналітичної інформації, у тому числі у

сфері сільського господарства. Однак рекомендаційні механізми YouTube та орієнтація на «залучення» підсилюють видимість сенсаційних повідомлень, що збільшує ризики маніпуляцій (клікбейт, підміна причинно-наслідкових зв'язків, апеляція до страху, перекручення статистики) у чутливих для агросектору темах – від експорту зерна до ГМО чи цін на пальне й добрива.

Процес цифровізації агропромислового комплексу (АПК) вимагає впровадження інтелектуальних систем моніторингу, здатних автоматизувати виявлення ознак маніпулятивного впливу у величезних масивах відеоконтенту. Традиційні методи контролю інформації стають неефективними через швидкість поширення контенту та складність його модальності, де текст, аудіо та візуальні образи переплітаються для створення певного нарративу. Саме тому аналіз текстових метаданих – заголовків, описів, тегів та коментарів у поєднанні з транскриптацією аудіовізуального ряду стає фундаментом для побудови систем інформаційної безпеки в аграрній сфері.

Розробка інтелектуальної системи для аналізу контенту YouTube базується на інтегрованому використанні засобів автоматизованого збору даних, обробки природної мови (NLP) та машинного навчання. Запропонована система аналізу відеоконтенту передбачає багатоступеневу обробку інформації.

На першому етапі здійснюється збір даних із відеоплатформи YouTube. Для цього використовується утиліта `yt-dlp` – відкритий інструмент командного рядка, який дозволяє вилучати багаторівневі метадані (заголовок, опис, теги, ідентифікатор каналу) у структурованому форматі JSON без прямого звернення до платного API. Для подолання обмеження прихованих дизлайків інтегрується API проєкту "Return YouTube Dislike" через відповідний `post-processor` плагін, що дозволяє отримувати оціночні показники незадоволення аудиторії.

Другим етапом є отримання текстового представлення аудіоконтенту, що є критично важливим, оскільки значна частина маніпулятивного впливу реалізується через усне мовлення, яке не завжди відображається в заголовку. Для цього використовується OpenAI Whisper – нейромережева модель, яку часто використовують для відео, що не мають автоматичних субтитрів або вони неякісні. Використання Whisper дозволяє виявити невідповідність між «клікбейтним» заголовком та реальним змістом відео, що є однією з ознак маніпуляції.

На третьому етапі проводиться попередня обробка текстових даних. Вона включає очищення тексту, нормалізацію, видалення стоп-слів та інші процедури, спрямовані на підготовку тексту для подальшого лінгвістичного аналізу.

На завершальному етапі застосовуються математичні моделі та алгоритми NLP для виявлення ознак маніпуляції у тексті. Використання моделей сімейства BERT дозволяє визначати нюанси емоційного забарвлення та логічні протиріччя у текстах, що дозволяє класифікувати текст транскрипту за категоріями (правда/фейк) та проводити аналіз тональності коментарів.

Маніпулятивний контент часто характеризується специфічними цифровими відбитками. Ознаками дезінформації в аграрному контенті виступають як статистичні аномалії, так і лінгвістичні маркери. Математично

ймовірність маніпуляції оцінюється через вагові коефіцієнти ознак, де аномальний рівень дизлайків DL_x (отриманий через RYD) залишається важливим індикатором:

$$DL_x = \frac{D_x}{L_x},$$

де D_x – оціночна кількість дизлайків, L_x – кількість вподобань. Аномальні значення DL_x часто супроводжують відео з неправдивими даними, які викликають негативну реакцію спільноти.

Однак після рішення платформи YouTube у 2021 році приховати публічне відображення кількості дизлайків, традиційні підходи до оцінювання реакції аудиторії на відеоконтент зазнали суттєвих змін. У зв'язку з цим для аналізу потенційно маніпулятивного контенту використовують набір альтернативних метрик, що базуються на поведінкових характеристиках взаємодії користувачів із відео:

Engagement-to-View Ratio (коефіцієнт взаємодій до переглядів) – ця метрика відображає інтенсивність взаємодії аудиторії з відео відносно кількості його переглядів. Низьке значення показника EVR за високої кількості переглядів може інтерпретуватися як потенційна ознака аномальної поведінки аудиторії або маніпулятивного поширення контенту.

Semantic Consistency (семантична узгодженість) – метрика семантичної узгодженості використовується для оцінки відповідності між заголовком відео та фактичним змістом його транскрипту. Низька семантична подібність може свідчити про використання клікбейтових або маніпулятивних заголовків.

Comments Fakeness Index (індекс «фейковості» коментарів) – цей показник розраховується як відношення коментарів, що містять ключові слова-маркери ("фейк", "брехня", "маніпуляція"), до загальної кількості відгуків.

Для забезпечення практичної цінності дослідження розроблено прототип системи на мові програмування Python, що об'єднує згадані інструменти у єдину систему обробки даних.

Таблиця 1 Програмно-технологічний інструментарій моніторингу

Категорія	Інструмент	Функціональне призначення
Видобування даних	yt-dlp + RYD API	Завантаження метаданих та аудіо
Обробка сигналів	MoviePy	Обробка та конвертація аудіофайлів
Розпізнавання мовлення	OpenAI Whisper	Генерація текстових транскриптів
Аналіз природної мови (NLP)	BERT Classifier	Класифікація та аналіз настроїв
Керування даними	Pandas / TinyDB	Збереження та фільтрація результатів.

Процес ідентифікації маніпулятивного контенту починається з пошуку відео за ключовими словами (наприклад, "експорт зерна", "аграрна реформа") та вилучення метаданих через yt-dlp. Потім проводиться фільтрація відео за статистичними ознаками: виключення коротких роликів (Shorts) або відео з аномальними показниками залученості. Далі відбувається транскрибація відібраного контенту за допомогою Whisper. Отримання тексту дозволяє

порівняти його семантику із заголовком. Після цього обчислюється косинусна подібність між ембедінгами заголовка та тексту. Низька подібність є маркером клікбейту. На завершальному етапі текст класифікується моделлю BERT для визначення наявності дезінформаційних нарративів. Візуалізація результатів здійснюється через дашборди (наприклад, на базі Streamlit) для зручного моніторингу фахівцями аграрної галузі. Такий підхід дозволяє перетворити складне завдання виявлення маніпуляцій у чіткий технологічний процес, інтегрований у розбудову цифрової аграрної сфери.

Впровадження інтелектуального аналізу метаданих дозволяє ефективно виявляти маніпуляції навіть в умовах обмеженої публічної статистики YouTube. Поєднання API RYD із глибоким семантичним аналізом транскриптів моделями BERT забезпечує точність класифікації понад 85%. Це створює надійну базу для захисту інформаційно-комунікаційних процесів у розбудові аграрної сфери України.

Подальший розвиток дослідження передбачає розробку повноцінної архітектури нейронної мережі для семантичного аналізу відеоконтенту та її інтеграцію у вигляді веб-сервісу для фахівців з комунікацій в агросекторі

Список використаних джерел:

1. Vosoughi S., Roy D., Aral S. The spread of true and false news online. Science. 2018. Vol. 359, no. 6380. P. 1146–1151. URL: <https://doi.org/10.1126/science.aap9559>
2. Jurafsky D., Martin J. H. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. – 3rd ed. – Draft. – Stanford : Stanford University, 2023. 560 p.

Abstract: *The paper investigates the problem of spreading manipulative content concerning the agricultural sector on the YouTube video platform. An automated approach for detecting manipulations based on video text metadata analysis is proposed. Data collection is carried out using the yt-dlp tool (integrated with the "Return YouTube Dislike" API), the Whisper model is utilized for audio content transcription, and text evaluation is performed via natural language processing and machine learning methods, specifically BERT models for semantic analysis. The proposed approach enables automated monitoring of the information space and contributes to enhancing the information security of the agricultural sector.*

Keywords: *manipulative content, YouTube, agricultural sector, metadata, yt-dlp, OpenAI Whisper, natural language processing, machine learning, information security.*

Науковий керівник:

Пархоменко О. Ю.,

*канд. фіз.-мат. наук, доцент кафедри економічної кібернетики, комп'ютерних наук та інформаційних технологій
Миколаївський національний аграрний університет*