

## ЗАСТОСУВАННЯ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ВИЯВЛЕННЯ МАНІПУЛЮВАННЯ ГРОМАДСЬКОЮ ДУМКОЮ

Гуменюк М. М.

У статті досліджується застосування технологій штучного інтелекту (ШІ) для виявлення маніпуляцій громадською думкою. В умовах стрімкого розвитку цифрових медіа та зростання обсягів інформації, проблема маніпуляцій і дезінформації стала актуальною як для науковців, так і для практиків у галузі інформаційної безпеки.

Основна увага приділяється аналізу сучасних методів ШІ, таких як машинне навчання, обробка природної мови (NLP) та нейронні мережі, які використовуються для виявлення і аналізу закономірностей маніпуляцій у текстах, соціальних мережах та інших джерелах інформації. Розглядаються конкретні приклади застосування цих технологій для моніторингу та аналізу інформаційних потоків, а також ефективність різних підходів у контексті боротьби з дезінформацією.

Дослідження включає аналіз можливостей і обмежень використання ШІ для виявлення несанкціонованих інформаційних впливів, розглядаючи як методи виявлення так і способи протидії маніпуляціям. Вивчено, як алгоритми можуть допомогти у виявленні фейкових новин, бот-мереж та інших форм дезінформації, а також виявлено проблеми, які потребують подальшого вдосконалення технологій, такі як неповна точність розпізнавання контексту і можливість маніпуляції самими системами ШІ.

Результати дослідження підкреслюють важливість інтеграції інноваційних технологій для забезпечення прозорості та об'єктивності інформаційного середовища. Окреслено перспективи подальших досліджень, які можуть сприяти розвитку більш ефективних інструментів для виявлення і запобігання маніпуляціям.

У статті також розглядається вплив інтеграції ШІ на існуючі системи інформаційної безпеки, зокрема, як ці технології можуть бути впроваджені в існуючі механізми моніторингу та аналізу медіа-контенту. Аналізуються приклади успішних реалізацій ШІ в боротьбі з дезінформацією, а також проблеми, що виникають при їх застосуванні на практиці.

Особливу увагу приділено етичним питанням, пов'язаним з використанням ШІ для аналізу інформаційних потоків. Обговорюються потенційні ризики,

такі як вторгнення в особисте життя, зловживання даними та питання конфіденційності. Запропоновано рекомендації для забезпечення етичного використання технологій ШІ в контексті інформаційної безпеки.

Висвітлено необхідність міждисциплінарного підходу в дослідженні та впровадженні технологій ШІ для виявлення маніпуляцій громадською думкою, що включає співпрацю між фахівцями в галузі інформаційних технологій, соціальних наук та права. Обґрунтовано, що ефективне вирішення проблеми маніпуляцій вимагатиме комплексного підходу, що охоплює технічні, етичні та правові аспекти.

**Ключові слова:** штучний інтелект, маніпуляція громадською думкою, дезінформація, обробка природної мови, машинне навчання, нейронні мережі, соціальні мережі, інформаційна безпека, фейкові новини, бот-мережі, детекція маніпуляцій, алгоритми виявлення, інформаційний моніторинг, контекстуальний аналіз, етичні питання ШІ.

**Humeniuk M. M. Application artificial intelligence technologies to detect manipulation of public opinion**

The article explores the use of artificial intelligence (AI) technologies to detect manipulation of public opinion. In the context of the rapid development of digital media and the growth of information, the problem of manipulation and disinformation has become relevant for both scientists and practitioners in the field of information security.

The focus is on the analysis of modern AI methods, such as machine learning, natural language processing (NLP), and neural networks, which are used to identify and analyse patterns of manipulation in texts, social media, and other sources of information. Specific examples of the use of these technologies for monitoring and analysing information flows, as well as the effectiveness of various approaches in the context of combating disinformation, are considered.

The study includes an analysis of the possibilities and limitations of using AI to detect unauthorised information influences, considering both detection methods and ways to counteract manipulations. The study examines how algorithms can help detect fake news, botnets and other forms of disinformation, and identifies issues that require further technology

development, such as incomplete contextual recognition accuracy and the possibility of manipulation of AI systems themselves.

The results of the study highlight the importance of integrating innovative technologies to ensure transparency and objectivity of the information environment. Prospects for further research are outlined, which can contribute to the development of more effective tools for detecting and preventing manipulation.

The article also examines the impact of AI integration on existing information security systems, in particular, how these technologies can be implemented in existing mechanisms for monitoring and analysing media content. Examples of successful implementations of AI in the fight against disinformation are analysed, as well as the problems that arise when applying them in practice.

Special attention is paid to ethical issues related to the use of AI to analyse information flows. Potential risks such as invasion of privacy, data misuse, and privacy issues are discussed. Recommendations are offered to ensure the ethical use of AI technologies in the context of information security.

The article highlights the need for an interdisciplinary approach to the research and implementation of AI technologies to detect manipulation of public opinion, including cooperation between experts in the field of information technology, social sciences and law. It is substantiated that an effective solution to the problem of manipulation will require an integrated approach covering technical, ethical and legal aspects.

**Key words:** Artificial Intelligence, Public Opinion Manipulation, Disinformation, Natural Language Processing, Machine Learning, Neural Networks, Social Media, Information Security, Fake News, Bot Networks, Manipulation Detection, Detection Algorithms, Information Monitoring, Contextual Analysis, Ethical Issues in AI.

**Постановка проблеми.** У сучасному інформаційному суспільстві проблема маніпуляцій громадською думкою та дезінформації набуває дедалі більшого значення. З розвитком цифрових медіа та соціальних платформ, обсяг інформації, що циркулює у суспільстві, зростає експоненціально, створюючи сприятливе середовище для поширення маніпулятивних і дезінформаційних повідомлень. Традиційні методи виявлення і боротьби з такими явищами виявляються недостатніми через обмежену здатність аналізувати великі обсяги даних у реальному часі та виявляти складні взаємозв'язки маніпуляцій.

Технології штучного інтелекту (ШІ) пропонують нові можливості для вирішення цієї проблеми завдяки здатності автоматизувати процеси аналізу інформації, виявлення аномалій, а також прогнозування потенційних загроз. Зокрема, методи

машинного навчання, обробки природної мови (NLP) і нейронні мережі демонструють значний потенціал у виявленні маніпуляцій і дезінформації в текстах і соціальних медіа. Однак, впровадження ШІ в цю сферу супроводжується рядом складнощів і викликів. Перш за все, необхідно розв'язати проблеми точності та ефективності алгоритмів у виявленні контекстуальних маніпуляцій і фейкових новин. Також важливо враховувати етичні аспекти використання ШІ, такі як захист приватності користувачів і запобігання зловживанням з боку самих систем ШІ. Таким чином, проблема полягає в тому, як ефективно застосовувати технології ШІ для виявлення і протидії маніпуляціям громадською думкою, забезпечуючи високу точність, етичність і відповідність сучасним вимогам інформаційної безпеки. Дослідження цих аспектів має на меті розробити ефективні інструменти та методи для виявлення і запобігання дезінформації, сприяючи таким чином збереженню інформаційної прозорості та об'єктивності.

**Аналіз останніх досліджень і публікацій.** Окремі аспекти досліджуваної проблематики були об'єктом дослідження таких вітчизняних вчених, як: Аніщенко Г., Денисюк Ж., Дзюба М., Замаруєва І., Зотов В., Жарков Я., Григоров О. В., Григорова Т., Марченко О., Мартиненко І., Мельников Є., Москаленко О., Никоненко А., Окунь А., Пономарьов О., Резанова Н., Россада Т., Стрижак В., Таран В., Турчин О., Василенко М., Візниця Ю., Шевченко Т. та інших. Крім того, досліджувана проблематика розглядалася у працях таких іноземних вчених як: Bain P., Bastian B., Beltrametti M., Binns R., Chatila R., Chazerand P., Cheng J., Cows J., Dignum V., Floridi L., Greene D., Hoffman A., Ienca M., Jobin A., Katz E., Kartsaklis D., Lazarsfeld P., Stark L., Vaes J., Vayena E., Veale M. та інших.

**Метою статті** є дослідження і аналіз застосування технологій штучного інтелекту (ШІ) для виявлення та протидії маніпуляціям громадською думкою.

**Виклад основного матеріалу.** Технології штучного інтелекту (ШІ) пропонують нові можливості для виявлення і протидії таким маніпуляціям. Штучний інтелект, завдяки своїм потужним аналітичним можливостям, здатен обробляти величезні обсяги даних та виявляти текстові та смислові аномалії, що можуть свідчити про спроби маніпуляцій. Зокрема, методи машинного навчання, обробка природної мови (NLP) та нейронні мережі демонструють значний потенціал у сфері виявлення дезінформації, фейкових новин та

бот-мереж. Однак, впровадження таких технологій супроводжується рядом викликів. Проблеми точності розпізнавання контексту, етичні питання, пов'язані з конфіденційністю та захистом особистих даних, а також ризики зловживання самими системами ШІ потребують серйозного аналізу і вирішення.

З огляду на те, що маніпулювання громадською думкою може мати різні прояви та мету, доцільно підкреслити, що у будь-якому випадку зазначене має ознаки психологічного впливу з метою впливу на поведінку певного числа представників суспільства, що може бути зокрема і в рамках інформаційних воєн, з метою зміни психологічних характеристик у бажаному для впливу напрямку (поглядів, думок, світосприйняття, ціннісних орієнтирів, умонастроїв, стереотипів поведінки) [2, с. 40]. Очевидно, у такому контексті, громадська думка транслює позицію значної частини представників суспільства, що у свою чергу впливає на суспільну свідомість в цілому, яка є сукупністю ідей, теорій, поглядів, уявлень, почуттів, вірувань, емоцій людей, настроїв, в яких відображається природа, матеріальне життя суспільства та вся система суспільних відносин [3, с. 619]. Очевидно, що громадська думка або суспільна думка, як консолідоване судження суспільства або значної його частини, може бути піддана впливу та маніпуляції.

Маніпулювання суспільною думкою - один з ключових способів керування масами. Маніпуляція здійснюється таємно і ставить своїм основним завданням встановлення контролю за поведінкою людини, що позбавляє свободи вибору об'єкта маніпуляції змінюючи його уявлення, особисту думку та цілі у потрібному напрямку. За багатьма ознаками маніпуляція суспільною думкою нагадує війну невеликою, але добре організованою та озброєною армією чужинців проти величезного мирного населення, яке до цієї війни не готове. Панівна меншість всіляко заважає роботі з викриття «гіпнотизерів», намагається не допустити маси до знання доктрин та технологій маніпуляції їхньою свідомістю. Здебільшого це досягається щедрою винагородою «тих, хто з нами» та блокуванням «тих, хто не з нами» [1].

З огляду на те, що маніпулювання громадською думкою може мати різні прояви та мети, важливо підкреслити, що цей процес має чіткі ознаки психологічного впливу, спрямованого на зміну поведінки певних груп суспільства. Такі маніпуляції можуть бути частиною інформаційних воєн, що переслідують мету корегування психологічних

характеристик населення в бажаному напрямку - будь то зміна поглядів, світосприйняття, ціннісних орієнтирів, умонастроїв чи стереотипів поведінки. Важливу роль у цьому контексті відіграють сучасні технології штучного інтелекту.

Штучний інтелект, зокрема, може суттєво підвищити ефективність маніпуляцій завдяки своїй здатності аналізувати великі обсяги даних, виявляти закономірності в поведінці користувачів і передбачати їхні реакції. Застосування алгоритмів машинного навчання дозволяє детально досліджувати інтереси та уподобання різних соціальних груп, що дозволяє адаптувати інформаційний контент до індивідуальних або групових особливостей, підвищуючи його впливовість.

Водночас, в умовах такої технологічної інтервенції, громадська думка, яка відображає позицію значної частини суспільства, може стати об'єктом підвищеного впливу і маніпуляцій. ШІ здатен модифікувати інформаційний ландшафт, створюючи специфічні наративи та формулюючи нові стереотипи, що у свою чергу вплине на суспільну свідомість. Таким чином, суспільна свідомість, яка є сукупністю ідей, теорій, поглядів, уявлень та емоцій, відображає матеріальне життя суспільства та всю систему суспільних відносин. В умовах активного використання ШІ для маніпулювання інформацією, громадська думка стає не тільки відображенням існуючих соціальних реалій, але і цілеспрямованою конструкцією, що формується під впливом технологічних інструментів.

У сучасних умовах поширеним є використання систем ШІ на базі алгоритмів машинного навчання та лінгвістичних нейромереж для захисту даних з одного боку, й для маніпуляції ними з іншого. Наприклад, О. Москаленко та Т. Григорова визначають це як галузь розвитку штучного інтелекту, де застосовуються алгоритми для отримання даних і навчання на їхній основі, тобто алгоритм вчиться з власного досвіду [4]. Інші дослідники підкреслюють, що машинне навчання є підмножиною технологій ШІ, яку можна описати як набір алгоритмів, що дозволяють комп'ютерам навчатися виконувати завдання, на які вони спочатку не були запрограмовані. Іншими словами, це алгоритмічна програма, яка сама може створювати програми для вирішення різних завдань [5].

Це підкреслює значущість машинного навчання як одного з найпотужніших інструментів у сфері штучного інтелекту. Машинне навчання дозволяє комп'ютерам адаптуватися до нових завдань і середовищ, що робить його надзвичайно цінним у різних галузях, з метою виявлення маніпулювань

громадською думкою [7]. Одним із ключових аспектів машинного навчання є його здатність до самонавчання, що відкриває нові можливості для автоматизації процесів, аналізу великих обсягів даних і прогнозування. Важливо також зазначити, що застосування машинного навчання не обмежується лише теоретичними дослідженнями. Воно активно вже використовується у реальному житті для покращення різних сервісів і продуктів. Від рекомендованих систем на платформах стрімінгу до автономних транспортних засобів, машинне навчання є рушійною силою багатьох інновацій, що змінюють наше повсякденне життя, а тому використання штучного інтелекту на базі алгоритмів машинного навчання може бути ефективним підходом до виявлення маніпуляцій громадською думкою [6].

Машинне навчання та штучний інтелект повинні враховувати позитивні етичні принципи для забезпечення безпеки та захисту прав людини. Це включає прозорість, коли алгоритми ШІ мають бути зрозумілими та пояснюваними, щоб користувачі могли зрозуміти, як приймаються рішення. Конфіденційність також є критично важливою: дані користувачів повинні оброблятися з повагою до їхньої приватності та захищатися від несанкціонованого доступу. Справедливість вимагає уникнення упереджень і дискримінації в алгоритмах, забезпечуючи рівне ставлення до всіх груп населення [8].

Разом із тим, нині технології штучного інтелекту можуть бути використані для виявлення маніпулювання громадською думкою через кілька ключових способів. Так, штучний інтелект здатний аналізувати великі обсяги даних з соціальних мереж і новинних сайтів, виявляючи аномалії та тенденції, що свідчать про маніпуляції. Крім того, алгоритми обробки природної мови можуть розпізнавати фальшиві новини та дезінформацію, перевіряючи факти та порівнюючи їх із достовірними джерелами. Ще один спосіб використання технології полягає в аналізі поведінки користувачів, відстеженні та виявленні координованих кампаній з маніпулювання громадською думкою. Разом із тим, штучний інтелект на базі машинного навчання може ідентифікувати ботів і фейкові акаунти за допомогою алгоритмів машинного навчання, що зменшує їхній вплив, що можливо шляхом постійного моніторингу соціальних мереж, виявляючи маніпулятивний контент. Ще один спосіб передбачає використання технології передбачає здатність до розпізнавання візуального контенту, такого як зображення та відео, на предмет

маніпуляцій за допомогою алгоритмів комп'ютерного зору, так ШІ може аналізувати емоційне забарвлення повідомлень, виявляючи вплив маніпуляцій на настрої громадськості. Крім того, ШІ ідентифікує ключові джерела та осіб, які впливають на громадську думку, що дозволяє ефективніше протидіяти маніпуляціям. Таким чином, на нашу думку, ці та інші підходи до впровадження технології ШІ до вирішення проблематики виявлення маніпуляцій громадською думкою мають значний потенціал та можливі до впровадження.

**Висновки.** Технології штучного інтелекту (ШІ) пропонують нові можливості для виявлення та протидії маніпуляціям громадською думкою завдяки потужним аналітичним можливостям, здатності обробляти величезні обсяги даних і виявляти текстові та смислові аномалії. Методи машинного навчання, обробка природної мови (NLP) та нейронні мережі демонструють значний потенціал у виявленні дезінформації, фейкових новин та бот-мереж. Однак впровадження таких технологій супроводжується викликами, такими як проблеми точності розпізнавання контексту, етичні питання щодо конфіденційності та захисту особистих даних, а також ризики зловживання самими системами ШІ. Маніпулювання громадською думкою має різні прояви та мету, завжди пов'язані з психологічним впливом на поведінку певних груп суспільства. Це може бути частиною інформаційних воєн, спрямованих на зміну психологічних характеристик у бажаному напрямку, включаючи погляди, світосприйняття, ціннісні орієнтири та стереотипи поведінки. Громадська думка, як консолідоване судження суспільства, може бути піддана впливу та маніпуляціям.

ШІ здатний суттєво підвищити ефективність маніпуляцій завдяки своїй здатності аналізувати великі обсяги даних, виявляти закономірності в поведінці користувачів і передбачати їхні реакції. Застосування алгоритмів машинного навчання дозволяє детально досліджувати інтереси та уподобання різних соціальних груп, адаптуючи інформаційний контент до індивідуальних або групових особливостей. Водночас громадська думка, яка відображає позицію значної частини суспільства, може стати об'єктом підвищеного впливу та маніпуляцій. ШІ здатен модифікувати інформаційний ландшафт, створюючи специфічні наративи та формуючи нові стереотипи, що впливають на суспільну свідомість. Таким чином, суспільна свідомість, яка є сукупністю ідей, теорій, поглядів та емоцій, відображає матеріальне життя суспільства та всю систему суспільних відносин.

Машинне навчання дозволяє комп'ютерам адаптуватися до нових завдань і середовищ, що робить його цінним інструментом у виявленні маніпуляцій громадською думкою. Важливо також зазначити, що застосування машинного навчання активно використовується в реальному житті для покращення різних сервісів і продуктів, що змінюють наше повсякденне життя. Використання штучного інтелекту на базі алгоритмів машинного навчання може бути ефективним підходом до виявлення маніпуляцій громадською думкою. Для забезпечення безпеки та захисту прав людини машинне навчання та штучний інтелект повинні враховувати позитивні етичні принципи. Це включає прозорість алгоритмів, конфіденційність даних користувачів, справедливість, уникнення упереджень та дискримінації, а також відповідальність розробників і користувачів ШІ. Врахування цих принципів сприяє створенню етичних та надійних систем ШІ, які підтримують права людини та суспільні цінності. Таким чином, технології ШІ можуть бути використані для виявлення маніпулювання громадською думкою шляхом аналізу великих обсягів даних, розпізнавання фальшивих новин, дезінформації, ідентифікації ботів і фейкових акаунтів, а також постійного моніторингу соціальних мереж. Ці підходи мають значний потенціал і можуть бути впроваджені для вирішення проблематики маніпуляцій громадською думкою.

#### Література

1. Шустенко О. Маніпулятивні технології як феномен цифрової дійсності. *Вісник Львівського університету. Серія філос.-політолог. студії*. 2023. Вип. 49. С. 200-204.

2. Жарков Я. М., Дзюба М. Т., Замаруєва І. В., ін. Інформаційна безпека особистості, суспільства, держави, Київ.: Видавничо-поліграфічний центр «Київський університет», 2008. 274 с.

3. Суспільне буття і суспільна свідомість // Філософський енциклопедичний словник / В. І. Шинкарук (гол. ред.) та ін. Київ : Інститут філософії імені Григорія Сковороди НАН України : Абрис, 2002. С. 619. 742 с.

4. Москаленко О., Григорова Т. Алгоритми штучного інтелекту для пошуку інформації в системах дистанційного навчання. *Прикладні питання математичного моделювання*. 2020. Т. 3, № 1. С.133.

5. Василенко М., Шевченко Т. Застосування штучного інтелекту в публічному управлінні, судочинстві та правоохоронній діяльності: міждисциплінарне дослідження. *Право і суспільство*. 2021. № 5, с. 151.

6. Floridi L., Cows J., Beltrametti M., Chatila R., Chazerand P., Dignum V., et al. AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*. 2018. 28(4). P. 689-707.

7. Григоров О. В., Аніщенко Г. О., Стрижак В. В., Петренко Н. О.І., Турчин О. В., Окунь А. О., Пономарьов О. Е. Штучний інтелект. Машинне навчання. Автомобіль і електроніка. *Сучасні технології*. 2019. 15. С. 20.

8. Jobin A., Ienca M., Vayena E. The global landscape of AI ethics guidelines. *Nature Machine Intelligence*. 2019. 1(9). P. 389-399.

*Гуменюк М. М.,  
orcid.org/0009-0006-3118-721X  
співробітник  
Служби безпеки України*