

МОЖЛИВІ ВАРІАНТИ УДОСКОНАЛЕННЯ ІСНУЮЧИХ ПІДХОДІВ ДО ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН НА ОСНОВІ ВИКОРИСТАННЯ ПОТЕНЦІАЛУ АЛГОРИТМІВ МАШИННОГО ТА ГЛИБОКОГО НАВЧАННЯ, СЕНТИМЕНТУ НОВИННОГО КОНТЕНТУ Й ЕМОЦІЙ У КОМЕНТАРЯХ КОРИСТУВАЧІВ

На даний час серед джерел інформації Інтернет займає перше місце. В останній період суттєво зросла роль онлайн соціальних мереж (ОСМ), що має як позитивні, так і негативні наслідки. Негативна роль ОСМ пов'язується із поширенням фейкових новин, які впливають на повсякденне життя людей, маніпулюють їхніми думками та почуттями, змінюють їхні переконання і можуть призвести до прийняття неправильних рішень. Проблема розповсюдження фейкових новин в ОСМ на даний час є глобальною, а формування механізмів протидії – актуальним завданням сьогодення.

На сьогодні існують різні апробовані підходи до виявлення фейкових новин. Зокрема, один з підходів базується на використанні різних алгоритмів машинного (ML) та глибокого (DL) навчання. Інший – на використанні результатів аналізу настрою новинного контенту та аналізу емоцій у коментарях користувачів. Проведене авторами дослідження інших підходів виявлення фейкових новин, які відрізняються від наведених, дозволило зробити висновок про те, що зазначені підходи є ефективними і перспективними в частині використання їх потенціалу для розробки нових моделей з високими показниками ефективності на різних наборах даних.

У статті сформовано та формалізовано авторські ідеї щодо удосконалення існуючих підходів виявлення фейкових новин на основі використання потенціалу зазначених підходів. Перша ідея базується на реалізації механізму комбінування методів машинного (ML) та глибокого (DL) навчання, а також результатів аналізу настрою новинного контенту й емоцій у коментарях користувачів, при якому враховується можливість забезпечення достатнього рівня ефективності виявлення фейкових новин, певного рівня значень показників обраної метрики, а також певного рівня функціональних характеристик авторського методу. Друга ідея базується на реалізації механізму комбінування функціоналу двох методів з числа зазначених двох груп, який би забезпечував оптимальні параметри виявлення фейкових новин за визначеними критеріями та показниками.

Обґрунтування ідей передбачало попереднє здійснення: постановки досліджуваної задачі; функціонального аналізу алгоритмів машинного (ML) та глибокого (DL) навчання, а також алгоритмів виявлення фейкових новин, що базуються на використанні результатів аналізу настрою новинного контенту й емоцій у коментарях користувачів; опису метрики для оцінки ефективності методів виявлення фейкових новин. За результатами обґрунтування перспективності ідей проведено формалізацію задач виявлення фейкових новин в авторській постановці.

Ключові слова: онлайн соціальні мережі; фейкові новини; метод; модель; алгоритм; формалізація.

Вступ та постановка проблеми у загальному вигляді. До недавнього часу люди отримували новини та інформацію переважно з газет і телевізійних каналів. Проте з появою Інтернету ситуація змінилася кардинально. Інтернет зайняв перше місце серед джерел інформації [1]. Суттєво зросла роль онлайн соціальних мереж (ОСМ) [2 - 4]. Стрімкий технологічний розвиток, що спостерігається в останні роки, особливо в галузі мобільної промисловості, зробив соціальні медіа, такі як Facebook, Twitter, Instagram та Sina Weibo, доступними і невід'ємною частиною нашого повсякденного життя [2, 5]. У результаті дослідження, проведеного в 2017 році, встановлено, що 67% осіб у США отримують новини в основному з соціальних мереж [6]. Згідно з Global Overview Report (<https://datareportal.com/reports/digital-2021-global-overview-report>, доступно 11 листопада 2021

року) кількість людей у світі з обліковими записами в ОСМ у 2021 році досягла рівня приблизно 4,20 мільярди, що складає понад 53% загального населення світу [7]. ОСМ стали не лише засобом комунікації між людьми, а й інструментом обміну інформацією та формування громадської думки [8]. Соціальні медіа надали віртуальне середовище для публікацій [9], обговорення, обміну поглядами та глобальної взаємодії між користувачами [10], без обмежень за місцем, часом або обсягом контенту [11].

Основна причина використання соціальних медіа користувачами у всьому світі - отримання новин та оперативне слідкування за поточними подіями.

Багато людей використовують соціальні медіа для публікації новин та інформації через свої облікові записи або сторінки, оскільки публікація новин на цих платформах відрізняється від публікації у традиційних ЗМІ тим, що не займає багато часу, не потребує витрат і не підлягає обмеженням аудиту [5, 12]. Структура соціальних медіа дозволяє поширювати новини в режимі реального часу та швидко, незалежно від достовірності цих новин [3]. Наприклад, у США в 2012 році 49% користувачів ділилися новинами в соціальних мережах. Згідно з доповіддю Pew Research Center (<https://www.pewresearch.org/>, доступно 23 серпня 2021 року), у 2016 році понад 62% користувачів щоденно отримували новини з ОСМ [12], а у 2018 році дві третини дорослого населення США отримували новини з платформ соціальних медіа [13].

Однак, в останній період ОСМ відіграють не лише позитивну, а й негативну роль. Це пов'язано із забезпеченням ними поширення фейкових новин. Дослідження, проведене в 2016 році, показало, що 23% осіб у США ділилися фейковими новинами, як навмисно, так і ненавмисно [14]. У Китаї більше третини інформаційних повідомлень, які стають трендами на мікроблогах, є фейковою інформацією [15]. Деякі популярні джерела інформації, що вважаються надійними (наприклад, Вікіпедія), також поширюють невірну інформацію або фейкові новини [16].

Фейкові новини можна визначити як опубліковані новинні статті, що містять неправдиву інформацію з метою свідомого введення читачів в оману [2] або здійснення зловживань [17]. Зазвичай у фейкових новинах містяться високопровокаційні повідомлення, створення яких переслідує фінансову чи політичну цілі. У 2017 році «фейкові новини» за версією словника [5] були оголошені офіційним словом року.

Згідно з доповіддю GDATA, 59% користувачів ОСМ стикалися з неправдивою інформацією [18]. Понад 57% користувачів ОСМ очікували, що опубліковані новини будуть неточними [13]. Сайт Statista представив статистику (<https://www.statista.com/statistics/649221/fake-news-expose-responsible-usa/>, доступно 4 березня 2021 року) за 27 серпня 2019 року, що базується на опитуванні, проведеному у США в 2018 році, щодо того, як ОСМ відповідають за поширення фейкових новин. Згідно наведених даних, 29% учасників вважають, що за поширення фейкових новин в основному відповідають соціальні медіа, а 60% відмітили, що ці платформи за поширення фейкових новин відповідають частково. Фейкові новини мають на 70% більше шансів поширюватися, ніж новини реальні [12]. Згідно з дослідженнями щодо швидкості поширення фейкових новин, твіти, які містять фальсифіковану інформацію в Twitter, досягають користувачів у шість разів швидше, ніж надійні твіти [19]. Швидкість поширення фейкових новин в Інтернеті більша, ніж реальних [12], оскільки люди цікавляться новою інформацією або новинами [20] та схильні ділитися останньою інформацією [21], особливо розповідати про важливі новини, не перевіряючи їх достовірність [22]. Повторний перегляд фейкових новин робить їх одержувачам знайомими, підвищує їх вірогідність і призводить до їх поширення як реальних новин [23].

Фейкові новини впливають на повсякденне життя людей [1, 8], маніпулюють їхніми думками та почуттями [24, 25], змінюють їхні переконання [26] і можуть призвести до прийняття неправильних рішень [1]. Розповсюдження фейкових новин у соціальних мережах негативно впливає на суспільство [27, 28] у багатьох сферах, зокрема таких, як політична, економічна, соціальна, медична, технологічна та спортивна [15, 26]. Основними намірами поширення фейкових новин є фінансовий зиск, поширення ненависті на підставі

екстремістських мотивів, маніпулювання свідомістю людей з політичних міркувань або створення упереджених думок з виборчих міркувань [29, 30] тощо. Результатом поширення фейкових новин, як правило, є страх, паніка та фінансові втрати [28].

Все це вказує на негативні наслідки поширення фейкових новин і їх негативний соціальний вплив.

Про негативну роль фейкових новин можна судити, зокрема з таких прикладів. Одна з новин на Reddit спричинила реальну стрілянину (<https://www.rollingstone.com/politics/politics-news/anatomy-of-a-fake-news-scandal-125877/>, доступно на 15 березня 2022 року). Під час виборчої кампанії до виборів президента США у 2016 році було виявлено понад мільйон постів, пов'язаних із фейковою новиною, відомою як PIZZAGATE (<https://tinyurl.com/z38z5zh>, доступно на 15 червня 2022 року). У фейковій новині про мусульманську спільноту в Індії (<https://thelogicalindian.com/fact-check/muslim-spirit-restaurant-covid-19-coronavirus-20457>, доступно 16 липня 2021 року) на підставі доповіді BBC (<https://www.bbc.com/news/world-asia-india-53165436>, доступно 16 липня 2021 року) стверджувалось, що мусульмани намірено поширюють COVID-19. Це призвело до збільшення ворожнечі до мусульман і закликів до економічного бойкоту. Крім того, 20 найпопулярніших фейкових новин були більш обговорюваними, ніж 20 найпопулярніших реальних історій (<https://tinyurl.com/y8dckwhr>, доступно на 15 червня 2022 року).

У 2021 році Facebook оголосив, що було закрито близько 1,3 мільярда фейкових облікових записів та видалено понад 12 мільйонів публікацій з хибною інформацією про COVID-19 та вакцини [31].

У багатьох людей виникають труднощі у відрізненні фейкових новин від реальних, незалежно від статі, віку чи рівня освіти [16]. Відрізнити фейкові новини від реальних складно, оскільки, як свідчать наукові дослідження, людська здатність розрізняти правдиву та неправдиву інформацію відносно незначна і становить близько 54% [32].

Оскільки фейкові новини стали глобальним викликом і великою загрозою для демократії, економіки та мирного співіснування [33], різні суб'єкти (громадські організації, журналісти, політики, дослідники) працюють над зменшенням ризику [12].

Отже, проблема розповсюдження фейкових новин в ОСМ на даний час є глобальною, а формування механізмів протидії – актуальним завданням сьогодення. Її вирішення пов'язується з формуванням моделей, які виявляють фейкові новини та обмежують можливість їх поширення [34].

Аналіз останніх досліджень і публікацій, в яких започатковано розв'язання даної проблеми. На даний час існують різні апробовані підходи до виявлення фейкових новин. Зокрема, один з підходів базується на використанні різних алгоритмів машинного (ML) та глибокого (DL) навчання. Інший – на використанні результатів аналізу настрою новинного контенту та аналізу емоцій у коментарях користувачів. Існує і ряд інших підходів, які заслуговують на увагу, подальший аналіз і дослідження. При цьому кожен із них характеризується певним рівнем ефективності на різних масивах даних.

Слід зауважити, що для перевірки ефективності підходів часто застосовуються такі набори даних, як LIAR, FakeNewsNet-PolitiFact, FakeNewsNet-GossipCop і COVID-19.

LIAR – це великий, доступний для широкого загалу набір даних фейкових новин [35]. Набір даних містить близько 12800 записів і дві основні складові: профілі користувачів і короткі політичні заяви. Характеристики профілю користувача включають ім'я спікера, роботу, партійну належність, штат, кредитну історію та контекст. Заяви (зроблені в період з 2007 по 2016 рік) були позначені редакцією Politifact.com за допомогою шести деталізованих категорій: правда, переважно правда, наполовину правда, майже неправда, неправда і шар зігрівається. Ці шість міток є відносно збалансованими за розміром. Загалом кожна заява має свою пов'язану мітку та інформацію про автора цієї заяви.

FakeNewsNet - це комплексний набір даних (<https://github.com/KaiDMML/FakeNewsNet>, доступно на 20 березня 2022 року), який складається з повних текстових новинних статей, зібраних з веб-сайтів politifact.com і gossipcop.com (доступно на 18 березня 2022 року). Кожен з них містить новинні статті та інформацію про соціальний контекст.

COVID-19 - це збірник соціальних медіа-повідомлень, коментарів і новин, пов'язаних із COVID-19, класифікованих як реальні або фейкові на основі їхньої правдивості. Набір даних [36] був зібраний з різних соціальних медіа-платформ, таких як Twitter і YouTube. Крім того, організатори збірника зібрали 10700 соціальних медіа-повідомлень та новинних статей про COVID-19 у формі анотованого набору даних англійською мовою.

Також слід зауважити, що критеріями оцінки, які широко використовуються в завданнях класифікації тексту, часто є точність (A), точність (P), чутливість (R) та $F1$ -показник. Також при роботі з незбалансованим набором даних для оцінки продуктивності моделей в якості міри продуктивності часто використовується площа під кривою ROC (AUC), яка є мірою для порівняння алгоритмів навчання та побудови оптимальних моделей навчання [37-39].

До переліку класичних алгоритмів ML прийнято відносити логістичну регресію (LR), метод опорних векторів (SVM), дерево рішень (DT), наївний байєсівський класифікатор (NB), випадковий ліс (RF), $XGBoost$ (XGB), а також комбінацію цих алгоритмів. До високорівневих алгоритмів ML відносять згорткові нейронні мережі (CNN), двонаправлені рекурентні мережі з короткотривалою пам'яттю ($BiLSTM$), двонаправлені рекурентні мережі з вентиляними блоками ($BiGRU$), комбінації $CNN-BiLSTM$ і $CNN-BiGRU$, а також гібридний підхід на основі цих технік. Моделями на основі глибокого навчання є моделі $BERTbase$ та $RoBERTabase$.

У роботі [40] представлено огляд підходів до виявлення фейкових новин на основі використання алгоритмів машинного навчання ML з двома сценаріями методів представлення слів (статистичними та контекст-незалежними). Крім того, у роботі [40] проведено порівняльну оцінку восьми передових моделей машинного навчання, а саме CNN , $BiLSTM$, $BiGRU$, $CNN-BiLSTM$, $CNN-BiGRU$, різних гібридних моделей з двома типами моделей текстового представлення (контекст-незалежними і контекст-свідомими моделями вбудови), $BERTbase$, $RoBERTabase$.

Авторами роботи [40] встановлено, що комбінація класичних методів машинного навчання з ознаками $TF-IDF$ краща за інші методи на наборі даних $Liar$, включаючи передові моделі машинного навчання. Метод $BERTbase$ забезпечив точність на рівні з найкращими моделями. $RoBERTabase$ показав найкращі результати на наборі даних $PolitiFact$ з $F1$ -оцінкою 93.17%. Метод SVM з ознаками $TF-IDF$ продемонстрував результати кращі за результати моделей глибокого навчання на наборі даних $GossipCop$. На наборі даних $COVID-19$ найкращою моделлю був $BERTbase$. Однак експериментальні результати засвідчили, що жодна окрема техніка не змогла забезпечити найкращі показники продуктивності на всіх наборах даних.

Багато досліджень щодо виявлення фейкових новин в OCM залежать від однієї чи декількох ознак, таких як зміст, мережеве поширення або користувач [41 - 43]. Аналіз коментарів користувачів для визначення їхнього ставлення до новин може відігравати важливу роль у виявленні фейкових новин [44 - 46] та надавати уявлення про достовірність опублікованих новин [5, 34]. У роботі [47] стверджується, що коментарі користувачів мають велику дискримінантну цінність при виявленні фейкових новин, де вираження сентименту [38] або емоцій [48] має вирішальне значення. У праці [20] зазначено, що реакція користувачів на фейкові новини виражає емоції страху, огиди та здивування, тоді як на реальні новини - емоції очікування, суму, радості та довіри. Однак автори цієї праці не досліджували, наскільки добре емоції можуть ідентифікувати фейкові новини. Згідно [49], новизна може бути важливою складовою фейкових новин і значно підвищувати можливості їх поширення та прийняття в суспільстві. Більшість існуючих досліджень, які використовують аналіз сентименту, зосереджуються на сигналах сентименту вмісту фейкових новин [50]. Часто користувачі використовують емодзі замість текстових коментарів, щоб висловити свої думки про певні новини в OCM [51 - 53]. У цьому контексті техніки глибокого навчання (DL) суттєво сприяють класифікації, прогнозуванню та аналізу текстового контенту [54, 55]. Це пов'язано з їхньою здатністю до ефективного навчання [2, 12], виявлення ознак і складних патернів [56, 57].

У працях [3, 58] продемонстровано, що додавання ознак на основі аналізу настроїв та аналізу емоцій збільшує точність виявлення фейкових новин для більшості моделей глибокого навчання порівняно з використанням лише текстових ознак. Також авторами зазначених праць встановлено, що ознаки на основі аналізу настроїв новин та аналізу емоцій коментарів користувачів цих новин можуть бути використані соціальними медіа-платформами для боротьби з поширенням фейкових новин. Проте, застосування зазначеного підходу пов'язане з труднощами при роботі з незбалансованим набором даних.

Проведене авторами дослідження інших підходів виявлення фейкових новин дозволило зробити висновок про те, що проаналізовані вище підходи є ефективними і перспективними в частині використання їх потенціалу для розробки нових моделей з високими показниками ефективності виявлення фейкових новин на різних наборах даних.

Тому метою даної статті є формування та формалізація авторської ідеї щодо удосконалення існуючих підходів виявлення фейкових новин на основі використання потенціалу алгоритмів машинного (ML) та глибокого (DL) навчання, а також результатів аналізу настроїв новинного контенту й аналізу емоцій у коментарях користувачів.

Виклад основного матеріалу

Для досягнення визначеної мети вбачається за доцільне:

здійснити постановку досліджуваної задачі;

здійснити функціональний аналіз алгоритмів машинного (ML) та глибокого (DL) навчання, а також алгоритмів виявлення фейкових новин, що базуються на використанні результатів аналізу настроїв новинного контенту й емоцій у коментарях користувачів;

описати метрику для оцінки ефективності методів виявлення фейкових новин;

описати авторську ідею комплексування потенціалу алгоритмів машинного (ML) та глибокого (DL) навчання, а також результатів аналізу настроїв новинного контенту й емоцій у коментарях користувачів;

здійснити формалізацію задачі виявлення фейкових новин, що реалізує авторську ідею її вирішення.

Постановка досліджуваної задачі.

Задача виявлення фейкових новин розглядатиметься, як бінарна класифікація на основі заданих ознак і моделі для прогнозування вірогідності того, що певна новина є фейковою чи реальною.

Формально задача виглядає так. Нехай задано множину новинних статей, висловлювань або твітів $A = \{a_1, a_2, \dots, a_k\}$, де k – це довжина вхідного тексту, із заданими початковими ознаками. Завдання полягає у встановленні фейковості чи реальності їх контенту на основі застосування до заданих ознак кожного елемента множини A авторської моделі.

Функціональний аналіз алгоритмів машинного (ML) та глибокого (DL) навчання, а також алгоритмів виявлення фейкових новин, що базуються на використанні результатів аналізу настроїв новинного контенту й емоцій у коментарях користувачів.

Авторами ряду наукових праць для виявлення фейкових новин було досліджено і випробувано велику кількість ML-алгоритмів. Зокрема:

Логістична регресія (LR): Логістична регресія - це статистична модель, яка застосовується як базова модель для широкого спектру завдань текстової класифікації.

Support Vector Machine (SVM): Класифікатор SVM - це потужний класифікатор, який є перспективним для вирішення низки завдань обробки природної мови.

Мультиноміальний наївний байєсівський (MNB): MNB - це популярний вид ймовірнісного алгоритму, який дає суттєві результати в різних завданнях обробки природної мови.

Дерево рішень (DT): DT - це алгоритм на основі дерева, в якому кінцеві вузли представляють високорівневі ознаки. Кожен вузол відображає вихід, а листок - клас мітки. У вузлах приймається рішення на підставі навчання з нагляду, яке перетворює відображення ознак і значень у бажані результати.

Випадковий ліс (RF): RF – це алгоритм, що складається з набору дерев рішень, кожне з яких навчено на випадковому наборі ознак.

XGBoost (XGB): XGB - це комбінований алгоритм машинного навчання, в основі якого знаходиться градієнтний бустінг і який базується на деревах рішень. За допомогою бустінгу дерева будуються послідовно і кожне з них спрямоване на зменшення помилок попереднього.

Ensemble: Ensemble - це метод навчання, який об'єднує наведені вище алгоритми машинного навчання для покращення продуктивності.

Також було досліджено такі алгоритми:

CNN: CNN – це одновимірний згортковий нейронний мережа, що є потужним методом машинного навчання для автоматичного виділення ознак з текстових вхідних даних. CNN може автоматично видобувати локальні ознаки, вона менш обчислювально витратна, ніж інші алгоритми машинного навчання. Її архітектура включає один шар CNN з 128 фільтрами розміром ядра 5, які активуються функцією активації ReLU. Згенеровану карту ознак потім вдосконалюють і зменшують за допомогою шару максимального зведення, що призводить до отримання найбільш відповідної інформації. Після цього вихід розгортають і передають на вихідний шар з одиницею, яка активується сигмоїдною функцією активації.

LSTM: Модель LSTM [59] виправляє недоліки RNN, додаючи адитивну та мультиплікативну взаємодії до формули рекурентності та окремих станів пам'яті. Складність моделі можна збільшити, стекаючи LSTM-шари. За допомогою трьох гейтів - входу, забуття і виходу - моделі LSTM усувають проблеми ванішення та вибуху градієнта, які властиві RNN. Важливою характеристикою моделей LSTM є їхня здатність захоплювати віддалені залежності. У них використовується один шар BiLSTM з 128 одиницями для кодування вхідного тексту.

GRU: У варіанті GRU існують лише два гейти: гейт оновлення та гейт скидання. Гейт оновлення об'єднує гейти забуття та входу і вирішує, яка інформація буде передана поточному стану. Гейти скидання визначають, коли ігнорувати попередній прихований стан [60].

Проста RNN розглядає контекст минулого, але не може враховувати контекст майбутнього. Тому для врахування майбутнього та попереднього контексту використовуються бідирекційні LSTM (BiLSTM) і бідирекційні GRU (BiGRU) завдяки їхнім конструкціям. Для забезпечення їх переваги об'єднуються шари прихованих обчислень вперед і назад, що контролює потік інформації в обох напрямках і призводить до кращого навчання.

Незважаючи на те, що BiLSTM та BiGRU виявляють переваги при вирішенні ряду проблем NLP, вони не позбавлені двох недоліків: зі збільшенням високорозмірного вхідного простору зростає їх складність, що призводить до ще більшої складності при їх оптимізації; ці моделі можуть захоплювати інформацію про послідовний і наступний контекст (концепція бідирекційності), вони не можуть фокусуватися на найважливіших частинах контекстуальної інформації тексту. Щоб усунути перший недолік, можна використовувати CNN для зменшення розмірності простору ознак, зберігаючи при цьому інформативні ознаки з тексту. Крім того, CNN може захоплювати та виділяти локальні патерни.

CNN-BiLSTM: Гібридизація моделей на основі рекурентних мереж із CNN допомагає виділяти суттєві ознаки, захоплювати локальні контекстуалізовані патерни та покращувати точність моделі. Спочатку використовується один шар CNN з 128 фільтрами та розміром ядра 5 для обробки вхідних векторів і виділення локальних ознак. Результати карт ознак шару CNN подаються на вхід одному шару BiLSTM з 128 одиницями, щоб вивчити довгострокові залежності локальних ознак новинних статей. Далі йде вихідний шар з однією одиницею, активованою сигмоїдною функцією. За допомогою RNN можна вивчити і захопити часові та контекстуальні ознаки та довгострокові залежності вхідного тексту, а важливі локальні ознаки можна виявити, використовуючи потужність CNN у роботі з просторовими відношеннями [61, 62].

CNN-BiGRU: Схожа до моделі CNN-BiLSTM. Відмінність полягає в тому, що архітектуру із шаром BiLSTM замінено шаром BiGRU.

Гібрид (Hybrid): Це гібридна модель, яка об'єднує три моделі: один шар CNN із 128 нейронами та розміром ядра 5, за яким слідує максимальне згортання, потім шар BiLSTM з 128 одиницями, а далі шар BiGRU з 128 одиницями.

Ще один підхід з виявлення фейкових новин базується на застосуванні трансформерів. Однак їх використання на даний час поки що обмежене.

Для вирішення досліджуваної проблеми дослідники пропонували різні методи інтерпретації значення слова за допомогою векторів вбудовування. Методи на основі нейронних мереж, такі як Word2Vec і GloVe, загалом використовувалися для навчання векторів вбудовування слова з великих корпусів слів. Однак ці моделі вбудовування мають недолік у тому, що вони не враховують контекст і статичні вбудовування слів генеруються незалежно від їх контексту. Тому для досягнення більш точних результатів модель повинна бути здатною вловлювати семантичні та контекстуальні патерни. Більше того, вдосконалені моделі машинного навчання можуть автоматично видобувати семантичну інформацію з вхідних даних для виявлення фейкового контенту, але вони не можуть точно розпізнавати фейковий контент без глибокого розуміння тексту. Таким чином, в останні роки зросло зацікавлення до парадигми уваги. У галузі обробки природної мови спостерігається загальний зсув парадигми, спрямований на розробку набору моделей, які не лише підвищують точність, але й вирішують проблему відсутності маркованих даних.

Крім того, існує нагальна потреба в автоматичному виявленні фейкових новин. Це складна задача, оскільки існуючі моделі машинного навчання (до появи моделей на основі трансформерів) не забезпечували глибокого семантичного розуміння текстових вхідних даних. Це призвело до впровадження попередньо навчених мовних моделей на основі трансформерів. Використання таких моделей, навчених на великих немаркованих даних, для завдань класифікації тексту стає все більш популярним. Для адаптації до завдання наступного рівня над новими шарами нейронних мереж розташовуються попередньо навчені шари у мовних моделях [63]. У цьому випадку на верхньому рівні попередньо навчених мовних моделей (PLMs) додається повністю з'єднаний (FC) шар для цілей класифікації.

Ключовими моделями з потужним мовним представленням, навченим на величезних обсягах текстового корпусу, є моделі BERT [64] і RoBERTa [65].

BERT: BERT – модель, що використовує багаторівневий бідирекціональний трансформерний кодер, який одночасно моделює лівий і правий контексти [64]. У результаті BERT генерує вектори, які враховують контекст. BERT у подальшому надає можливість подолати одностороннє обмеження, проводячи передпочаткове навчання з використанням невдосконаленого завдання передбачення, яке включає в себе маскову мовну модель (Masked Language Model, MLM), яка розуміє контекст і передбачає слова. Таким чином, модель може генерувати векторне представлення, яке захоплює загальну інформацію вихідного тексту. Семантичне представлення кожного слова у вхідному тексті можна покращити за допомогою механізму уваги, підсилюючи семантичне представлення на основі контексту слова. Механізм уваги відіграє важливу роль в архітектурі трансформера, оскільки надає різні ваги різним частинам тексту в залежності від їхнього внеску в результат. Функція уваги відображає запити та слідує за парами ключ-значення та векторами-результатами. Механізм уваги використовує функцію активації Softmax, яка нормалізує вхідні дані до значення між 0 і 1.

Незважаючи на те, що BERT містить мільйони параметрів (наприклад, BERTbase має 110 мільйонів параметрів, а BERTlarge - 340 мільйонів параметрів) [64], BERT є відносно ефективним для застосування у завданнях наступного рівня з використанням спільно налаштованих параметрів на основі попередньо навченої моделі.

RoBERTa [65]: RoBERTa - оптимізована версія підходу BERT. У цьому методі BERT перенавчається за допомогою: вдосконаленої методології навчання, включаючи видалення завдання передбачення наступного речення з попереднього навчання; використання у 10 разів більшої кількості даних, ніж BERT; введення динамічного маскуванню з більшими

пакетами даних, так що токени, які маскуються, змінюються під час навчання, на відміну від статичного маскування, використовуваного в BERT. Таким чином, RoBERTa відрізняється від BERT в способі підходу до попереднього навчання.

Аналіз функціонального аспекту застосування наведених методів передбачає проведення аналізу попередньої обробки вхідного тексту. Попередня обробка вхідного тексту включала в себе видалення стоп-слів і пунктуації. Текст проходив процес токенизації на токени і побудови словника на основі вивченої лексики в даному текстовому корпусі. Такий словник використовувався для відображення кожного токена в єдиний цілочисельний ідентифікатор. Отримані послідовності доповнювалися або обрізалися до фіксованої кількості записів, оскільки моделям необхідно було подавати вектори однакової довжини. При цьому недоліком є те, що при обрізанні послідовностей втрачалася деяка (корисна) інформація. Ці послідовності потім перетворювалися в фіксовані вектори вбудовування слова. Потім нейронні мережі ініціалізувалися контекстно-незалежними попередньо навченими моделями вбудовування слова, такими як Word2Vec і GloVe, та контекстно-інформованими попередньо навченими мовними моделями BERTbase. У класичних алгоритмах машинного навчання для перетворення токенизованих текстів в ознаки використовувалися CountVectorizer (CV) та TF-IDF, як статистичні здобувачі ознак для моделей машинного навчання.

Також для аналізу функціонального аспекту застосування наведених методів доцільно оцінити результати застосування описаних вище моделей при виявленні фейкових новин на різних масивах даних.

Аналіз результатів на наборі даних LIAR.

Як впливає з [40], найкращий результат точності, досягнутий на наборі даних LIAR, становив 63,9% для комбінації з ознаками TF-IDF. Типовий результат при цьому становить 62%. Найгірший результат склав 51,78% для DT з ембедінгами GloVe. Класифікатор не міг виявити деякі закономірності у вхідному тексті.

При проведенні порівняльної оцінки результатів різних моделей встановлено, що більшість моделей досягли точності не більше 62% з BERTbase, а модель CNN-BiLSTM з використанням ембедінгів BERTbase точності 63,06%. Найгірший результат було досягнуто для моделей на основі рекурентних мереж з ембедінгами Word2Vec.

Оскільки набір даних LIAR містить короткі політичні заяви, складно отримати корисні вказівки, які могли б допомогти розрізнити фейкові новини від реальних новин, зокрема при використанні складних моделей, таких як LSTM, оскільки це може збільшити ризик перенавчання.

Як зазначено в роботі [66], інформація новинної статті та розмір набору даних - це важливі фактори, що впливають на продуктивність моделей на основі рекурентних мереж. Такі моделі більш схильні долати перенавчання, коли в новинній статті надається достатньо інформації. Отримані результати показують, що класичні методи машинного навчання випереджають високо вдосконалені методи глибокого навчання на наборі даних LIAR, включаючи найновіші передові мовні моделі, такі як BERTbase та RoBERTabase.

Аналіз результатів на наборі даних FakeNewsNet.

Результати, які отримані на наборах даних PolitiFact і GossipCop, свідчать про те, що надійність має постійно високу точність, випереджаючи базові SAF та CNN навіть з комбінаціями соціальних вказівок та контенту новин на обох наборах даних. Найбільше значення точності (89,93%) на наборі даних PolitiFact було досягнуто за допомогою комбінації із CV як методу вилучення ознак. Ця модель (Ensemble+CV) виявилася другою за результатами на наборі даних GossipCop, з SVM+TF-IDF ($F1$ -оцінка 91,55%) з числа інших моделей на наборі даних GossipCop. Методи глибокого навчання, зокрема найновіші передові мовні моделі, забезпечили високу продуктивність у багатьох завданнях обробки природної мови. RoBERTabase показав найкращий результат на наборі даних PolitiFact з $F1$ -оцінкою 93,17%, тоді як результати точності (92%) були на рівні з передовою моделлю машинного навчання за допомогою представлень BERTbase. Для набору даних GossipCop CNN-BERTbase виявився найкращою моделлю (значення $F1$ 91,45%). Це свідчить про те,

що контекстно-орієнтовані моделі на основі трансформатора допомагають виявляти корисні закономірності для розрізнення фейкових новин від реальних. Оскільки набір даних GossipCop досить незбалансований, різні методи збалансування, такі як вибіркове збільшення та зменшення, можуть допомогти збалансувати набір даних і підвищити продуктивність виявлення.

Аналіз результатів на наборі даних COVID-19.

BERT ефективно може виявляти фейковий контент, оскільки він має можливість кодувати глибоку семантичну контекстуальну інформацію. За результатами досліджень можна зробити висновок, що модель BERTbase, до якої додано лінійний шар для класифікації, показала найкращі результати порівняно з аналогами та іншими класичними моделями машинного та глибокого навчання на наборі даних COVID-19. Другою за ефективністю на наборі даних COVID-19 була модель CNN-BERTbase.

У зв'язку з тим, що методи векторних просторів, такі як CV та TF-IDF, не враховують контекст, використання цих представлень з моделями машинного навчання ґрунтується на зовнішньому вигляді токенів при прийнятті остаточних рішень, незалежно від їхнього контексту. Результати дослідження свідчать про те, що моделі векторного простору неефективні у виявленні глибокої семантики та контекстуальних закономірностей, що містяться в користувацькому змісті, створеному на Twitter. Однією з головних переваг BERT (та його модифікацій) у випадку Twitter (де користувацький зміст часто містить орфографічні помилки, шум та скорочення), є використання підтокенів замість фіксованих токенів. Це ідеально підходить для використання з такими даними [67], оскільки воно оперує на рівні підтокенів замість стандартних контекст-незалежних векторних вбудов слів на рівні слова.

Оцінимо тепер функціональний аспект алгоритмів виявлення фейкових новин, що базуються на використанні результатів аналізу настрою новинного контенту й емоцій у коментарях користувачів.

Насамперед відмітимо, що перевірку моделей, що відповідають зазначеним алгоритмам, як правило, здійснюють на наборі даних Fakeddit.

Набір даних Fakeddit (<https://github.com/entitize/Fakeddit>, доступний на 22 лютого 2022 року) є великомасштабним та багатомодальним набором даних (текст та зображення), зібраним з соціальної медіа-платформи Reddit за період з 19 березня 2008 року по 24 жовтня 2019 року [68]. Цей набір даних складається з понад мільйона повідомлень з різних галузей. До цих повідомлень додаються різні ознаки, такі як зображення, коментарі, користувачі, домени та інші метадані. У цьому наборі даних міститься багато шуму та нульових значень. Одне повідомлення може містити кілька коментарів або не мати коментарів. Для кожного повідомлення дослідники надали три мітки, класифіковані двома, трьома та шістьма способами.

У роботі [69] відзначається, що існує зв'язок між настроєм опублікованих новин та правдивістю новин. На основі цього автори роботи використовували ознаку, засновану на настрої (відношення кількості позитивних слів до кількості негативних слів), для розвитку власної моделі виявлення фейкових новин. Для покращення точності визначення фейкових новин у [70] запропоновано нову стратегію, яка включає збагачення об'єднаного набору даних настроєм як ключовою ознакою. Ефективність запропонованої стратегії перевірялася на основі трьох різних наборів даних. Результати показали ефективність запропонованого рішення. У роботі [50] підтверджено, що різницю між фейковими та справжніми новинами можна відрізнити за допомогою подвійної емоції (емоція видавця та суспільна емоція). Автори цієї роботи запропонували «Ознаки подвійної емоції», щоб виразити подвійну емоцію та взаємодію між ними. Вони також показали, що запропоновані ними ознаки легко інтегруються в існуючі моделі виявлення фейкових новин.

У роботі [71] запропоновано стратегію EmoCred, модель на основі довгострокової пам'яті (LSTM), яка враховує емоційні сигнали для виявлення різниці між справжніми та фейковими твердженнями. Головним кроком EmoCred є вилучення емоційних сигналів з тверджень. Дослідники вивчали три різні методи визначення емоційних сигналів у

твердженнях: метод, який використовує сучасні лексикони емоцій та базується на лексоні (emoLexi); метод, який визначає інтенсивність емоцій у твердженнях (emoInt); метод нейронної мережі (NN), який вказує на рівень інтенсивності тверджень і представляє кількість емоційних реакцій, які це твердження може викликати у читачів (emoReact). Дослідження авторів зазначеної праці, в якому використані реальні набори даних, продемонструвало цінність емоційних сигналів при оцінці вірогідності.

У роботі [72] проведено класифікацію новин на реальні та фальшиві на основі аналізу емоцій у новинах за допомогою комбінації моделей, що включають у себе згорткові нейронні мережі (CNN) і мережі BI-LSTM з механізмом уваги. Результати застосування моделі виявилися кращими, ніж результати застосування інших моделей.

Нижче наведені результати застосування описаних вище моделей при виявленні фейкових новин на різних масивах даних.

У роботі [68] наведено багатомодальний набір даних Fakeddit, застосовано модель BERT для виявлення фейкових новин на основі тексту і техніки ResNet для виявлення фейкових новин на основі зображень. Дослідники виявили, що точність результатів моделі, побудованої на основі BERT і ResNet з використанням ознак як тексту, так і зображень, була кращою, ніж у моделі, яка використовує лише текст для виявлення фейкових новин.

У роботі [73] запропоновано модель виявлення фейкових новин, яка ґрунтується на DeepNet та реальних даних наборів даних BuzzFeed і Fakeddit. З використанням факторизації тензорів, яка інтегрує контент новин та дані, що базуються на соціальному контексті, DeepNet демонстрував кращі параметри, ніж поточні моделі виявлення фейкових новин. Результати показали, що використання комбінації ознак на основі соціального контексту і ознак контенту новин призводило до більш точних результатів DeepNet.

У роботі [74] також використано набір даних Fakeddit і представлено багатомодальну мережеву архітектуру, яка дозволяє об'єднувати різні рівні та види об'єднання інформації, включаючи не тільки текст заголовків, але і метадані та інший контент, що стосується заголовків новин. Їхня методологія для виявлення неправдивої інформації залежить від чотирьох модальностей вводу: основний текстовий контент новини або запису; додаткова інформація або реакція на основний контент (наприклад, коментарі); візуальний контент запису; будь-які інші доступні метадані. Для врахування унікальності фундаментальної структури модальностей, автори роботи вони об'єднали інформацію на різних рівнях. Результати запропонованої ними моделі були покращені завдяки додатковим модальностям, що свідчить про те, що вони надають цінну інформацію.

Унікальна мережа вилучення та мережа мислення, відома як SERN, а також механізм визначення уваги на основі речення, були наведені в роботі [75]. Автори роботи пояснили, що відповіді, надіслані різними читачами, містять як факти, так і деякі висновки. Вони використали графічну мережу мислення. Дослідники у своїй роботі використали набори даних Fakeddit і RHEME та вказали, що запропонована ними модель продемонструвала перспективні результати порівняно з іншими сучасними моделями.

У роботі [76] розроблено підхід, який використовує ознаки контенту новин та ознаки соціального контексту для виявлення фейкових новин. Підхід авторів цієї праці базується на архітектурі трансформатора, яка складається з двох блоків: блоку кодування для вилучення значущих представлень з даних фейкових новин та блоку декодування для передбачення майбутньої поведінки на основі минулих даних. Вони провели обширні випробування на даних Fakeddit та NELA-GT-2019, щоб оцінити ефективність запропонованого ними підходу. Вони використовували стратегію недорогого семпсування, яка включає видалення записів з більшості класу з метою наближення більшості класу до меншості класу, щоб вирішити проблему нерівноваги даних в обох наборах даних. Автори роботи [76] використовували ознаки соціального контексту набору даних Fakeddit. За результатами застосування їхньої моделі можна зробити висновок, що для оптимальної роботи моделі потрібно враховувати як контент новин, так і соціальний контекст.

За результатами порівняльного оцінювання наведених моделей встановлено, що найефективніша модель виявлення, яка використовується на валідаційному наборі даних, це

Bi-LSTM зі значенням AUC 96,77% та 97,81% за $F1$ -оцінкою. Проте, навіть для неї мають місце труднощі при роботі з незбалансованим набором даних.

Таким чином, проведений функціональний аналіз алгоритмів машинного (ML) та глибокого (DL) навчання, а також алгоритмів виявлення фейкових новин, що базуються на використанні результатів аналізу сентименту новинного контенту й емоцій у коментарях користувачів, дозволяє зробити такі висновки: на даний час жодна з проаналізованих існуючих моделей виявлення фейкових новин не є універсальною для різних наборів даних; різні моделі та їх комбінації мають як переваги, так і недоліки, які не дозволяють ефективно застосовувати моделі; зважаючи на переваги та недоліки проаналізованих моделей, за доцільне вбачається дослідити питання про можливість поєднання описаних у даній роботі підходів до виявлення фейкових новин; поєднання підходів може розглядатися як на рівні комбінування методів з різних підходів, так і на рівні комбінування їх функціоналу на базі окремого методу; при проведенні дослідження щодо удосконалення моделей доцільно детермінувати метрику продуктивності методів і встановити масиви даних, на базі яких проводитиметься перевірка ефективності пропонуваного методу.

Метрика для оцінки ефективності методів виявлення фейкових новин.

За результатами проведеного аналізу в якості показників продуктивності методів виявлення фейкових новин вбачається за доцільне обрати точність (A), точність (P), чутливість (R), $F1$ -показник, а також площу під кривою ROC (AUC).

Точність (A) - це міра здатності класифікатора правильно класифікувати інформацію як фейкову або реальну. Точність (A) можна оцінити так

$$A_{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Точність (P) - це міра точності класифікатора, згідно якої низьке значення вказує на велику кількість помилкових позитивних результатів. Точність (P) представляє собою кількість позитивних прогнозів, поділену на загальну кількість позитивних значень класу, передбачених, і обчислюється за формулою

$$P_{recision} = \frac{TP}{TP + FP} \quad (2)$$

Чутливість (R) вважається мірою повноти класифікатора (наприклад, низьке значення чутливості вказує на багато помилкових негативних результатів), де кількість справжніх позитивів ділиться на суму справжніх позитивів і кількості помилкових негативів:

$$R_{ecall} = \frac{TP}{TP + FN} \quad (3)$$

$F1$ -показник ($F1$) обчислюється як зважена гармонічна середня точності і чутливості класифікатора:

$$F1 = \frac{2 \cdot P_{recision} \cdot R_{ecall}}{P_{recision} + R_{ecall}} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (4)$$

У формулах (1)-(4) TP , TN , FP , FN - це відповідно справжні позитиви, справжні негативи, помилкові позитиви і помилкові негативи.

Площа під ROC -кривою, відома як AUC , є мірою, яка використовується для порівняння алгоритмів навчання та побудови оптимальних моделей навчання. AUC , що близька до 1, вказує на перевагу системи, яка може точно відрізнити реальні та фейкові новини, тоді як AUC , що близька до 0, вказує на слабку систему (тобто система вважатиме всі фейкові новини реальними і навпаки) [38]. Наприклад, якщо AUC дорівнює 0,9, це означає, що модель на 90% може відрізнити негативні і позитивні класи. AUC зазвичай використовується в задачах незбалансованої класифікації [37], таких як виявлення фейкових новин.

Міра AUC може бути знайдена з виразу

$$AUC = \frac{1 - FPR + TPR}{2} \quad (5)$$

Справжній позитивний рівень або TPR - це скорочення, яке вказує на відсоток позитивних прикладів, які успішно класифіковані. На відміну від цього, FPR , який вказує на хибно позитивний рівень, це відношення випадків, які були неправильно класифіковані як негативні, до всіх інших випадків [39].

Авторські ідеї щодо комплексування потенціалу алгоритмів машинного (ML) та глибокого (DL) навчання, а також результатів аналізу настрою новинного контенту й емоцій у коментарях користувачів.

З урахуванням результатів проведеного аналізу авторам видається, що подальше дослідження механізмів виявлення фейкових новин можна здійснювати за двома напрямками (ідеями).

Ідея 1.

Ідея 1 базується на реалізації механізму комбінування методів машинного (ML) та глибокого (DL) навчання, а також результатів аналізу настрою новинного контенту й емоцій у коментарях користувачів. Однак доцільне комбінування має враховувати можливість забезпечення певної ефективності виявлення фейкових новин, певного рівня значень показників обраної метрики, а також певний рівень функціональних характеристик авторського методу.

Ідея 2.

Ідея 2 базується на реалізації механізму комбінування функціоналу ефективних існуючих методів на базі окремо обраного методу. При цьому важливим завданням є коректне обрання базового методу та вибір того функціоналу існуючих методів з числа аналізованих, який би забезпечив достатній рівень ефективності реалізації авторської ідеї на різних наборах даних.

Формалізація задач виявлення фейкових новин, що реалізує авторські ідеї.

Для перевірки гіпотези щодо ефективності авторських ідей 1, 2 доцільною є формалізація постановки задач. Здійснимо її.

Формалізація задачі згідно ідеї 1.

Нехай задано множину новинних статей, висловлювань або твітів $A = \{a_1, a_2, \dots, a_k\}$, де k – це довжина вхідного тексту, із невідомими початковими ознаками. Завдання полягає у встановленні фейковості чи реальності контенту елементів множини A на основі застосування до ознак кожного елемента множини A авторської моделі.

Авторська модель являє собою комбінацію алгоритмів машинного (ML) та глибокого (DL) навчання, а також алгоритмів виявлення фейкових новин на основі аналізу настрою новинного контенту й емоцій у коментарях користувачів. Умовне позначення цих алгоритмів таке: M_1, M_2, \dots, M_n . Їх кількість рівна n . Серед елементів $M_i, (i = \overline{1, n})$ можуть бути «чисті»

алгоритми машинного (ML) та глибокого (DL) навчання, або їх комбінації, що досліджені на даний час, а також «чисті» алгоритми виявлення фейкових новин на основі аналізу настрою новинного контенту й емоцій у коментарях користувачів, що проаналізовані вище, або їх можливі комбінації.

Нехай кожен з алгоритмів $M_i, (i = \overline{1, n})$ на наборі даних $j, (j = \overline{1, s})$ має певну ефективність $E_i^{(j)}, (i = \overline{1, n}, j = \overline{1, s})$.

Зважаючи на те, що для кожного алгоритму $M_i, (i = \overline{1, n})$ на окремих наборах даних $j, (j = \overline{1, s})$ відомими є величини $A_{accuracy.i}^{(j)}, P_{recision.i}^{(j)}, R_{ecall.i}^{(j)}, F1_i^{(j)}, AUC_i^{(j)}$ (визначаються із використанням формул (1)-(5)), то ефективність алгоритму $M_i, (i = \overline{1, n})$ на масиві

даних $j, \left(j = \overline{1, s} \right)$ можна представити у вигляді

$$E_i^{(j)} = g_i^{(j)} \left(A_{accuracy.i}^{(j)}, P_{recision.i}^{(j)}, R_{ecall.i}^{(j)}, F1_i^{(j)}, AUC_i^{(j)} \right). \quad (6)$$

Тоді ефективність кожного алгоритму $M_i, \left(i = \overline{1, n} \right)$ на довільному наборі даних може являти собою функцію $E_i = f \left(E_i^{(1)}, E_i^{(2)}, \dots, E_i^{(s)} \right)$.

А отже, ефективність авторського алгоритму може бути представлена у вигляді

$$E = \bar{f} \left(E_1, \dots, E_n \right) = \bar{f} \left(E_1^{(1)}, \dots, E_n^{(1)}, \dots, E_1^{(s)}, \dots, E_n^{(s)} \right). \quad (7)$$

Зважаючи на фізичний зміст задачі виявлення фейкових новин, при опрацюванні авторської моделі необхідним є підбір таких моделей $M_i, \left(i = \overline{1, n} \right)$, застосування яких дозволило б у межах заданого часу розв'язати задачу, а також не перевищити задану обчислювальну складність алгоритму.

Нехай для кожної моделі $M_i, \left(i = \overline{1, n} \right)$ відомими є величини:

$T_i^{(j)}$ - максимальна тривалість роботи моделі $M_i, \left(i = \overline{1, n} \right)$ на масиві даних $j, \left(j = \overline{1, s} \right)$;

C_i - обчислювальна складність алгоритму $M_i, \left(i = \overline{1, n} \right)$.

Тоді загальний час роботи $M_i, \left(i = \overline{1, n} \right)$ на всіх наборах даних $j, \left(j = \overline{1, s} \right)$ можна представити у вигляді

$$T_i = \tilde{f} \left(T_i^{(1)}, \dots, T_i^{(s)} \right), \quad (8)$$

а загальний час роботи авторської моделі на всіх наборах даних може бути представлений так:

$$T = \hat{f} \left(T_1, \dots, T_n \right). \quad (9)$$

З урахуванням введених позначень, обчислювальну складність авторської моделі можна описати функціональною залежністю

$$C = \bar{f} \left(C_1, \dots, C_n \right). \quad (10)$$

А отже, задачу побудови авторської моделі можна сформулювати, як оптимізаційну задачу вибору з числа алгоритмів $M_i, \left(i = \overline{1, n} \right)$ тих з них, які б забезпечили

$$E \rightarrow \max, \quad (11)$$

де E визначається за формулою (7), при виконанні умов

$$T \leq T_{gp}, \quad (12)$$

$$C \leq C_{gp}, \quad (13)$$

де T і C знаходяться, відповідно, з формул (9) і (10).

У виразах (12), (13) T_{gp} , C_{gp} - це задані граничні значення тривалості роботи авторської моделі на масиві $A = \{a_1, a_2, \dots, a_k\}$ і обчислювальної складності моделі.

Слід зауважити, що в авторській моделі (1)-(13) виявлення фейкових новин невідомими є функціональні залежності $g_i^{(j)}, f, \bar{f}, \tilde{f}, \hat{f}, \bar{f}$. Їх встановлення є окремим завданням.

Як приклад, зважаючи на ймовірнісний зміст задачі виявлення фейкових новин, ефективність E в (7) може являти собою таку функціональну залежність

$$E = 1 - \prod_{i,j} \left(1 - \overline{E_i^{(j)}} \right), \left(i = \overline{1, n}, j = \overline{1, s} \right). \quad (14)$$

Застосування формули (14) для виявлення ефективності E не становить значних обчислювальних труднощів при наявності даних щодо величин (6).

Слід також зауважити, що суть задачі легко оцінити з даних, що наведені в табл. 1.

Таблиця 1

Існуючі моделі $M_i, \left(i = \overline{1, n} \right)$	Ефективність існуючих моделей на досліджуваних наборах даних $j, \left(j = \overline{1, s} \right)$				Максимальна тривалість роботи моделі $M_i, \left(i = \overline{1, n} \right)$ на масиві даних $j, \left(j = \overline{1, s} \right)$				Обчислювальна складність алгоритму $M_i, \left(i = \overline{1, n} \right)$
	$j = 1$	$j = 2$...	$j = s$	$j = 1$	$j = 2$...	$j = s$	
M_1	$E_1^{(1)}$	$E_1^{(2)}$...	$E_1^{(s)}$	$T_1^{(1)}$	$T_1^{(2)}$...	$T_1^{(s)}$	C_1
M_2	$E_2^{(1)}$	$E_2^{(2)}$...	$E_2^{(s)}$	$T_2^{(1)}$	$T_2^{(2)}$...	$T_2^{(s)}$	C_2
...
M_n	$E_n^{(1)}$	$E_n^{(2)}$...	$E_n^{(s)}$	$T_n^{(1)}$	$T_n^{(2)}$...	$T_n^{(s)}$	C_n
Граничні значення окремих величин					T_{zp}				C_{zp}

Формалізація задачі згідно ідеї 2.

Нехай задано множину новинних статей, висловлювань або твітів $A = \{a_1, a_2, \dots, a_k\}$, де k – це довжина вхідного тексту, із відомими початковими ознаками. Завдання полягає у встановленні фейковості чи реальності контенту елементів множини A на основі застосування до ознак кожного елемента множини A авторської моделі.

Авторська модель являє собою комбінацію алгоритмів машинного (ML) та глибокого (DL) навчання з алгоритмами виявлення фейкових новин на основі аналізу настрою новинного контенту й емоцій у коментарях користувачів.

Умовне позначення алгоритмів машинного (ML) та глибокого (DL) навчання (чи їх досліджених комбінацій) таке: $M_i^{(1)}, \left(i = \overline{1, l} \right)$. Умовне позначення алгоритмів виявлення фейкових новин на основі аналізу настрою новинного контенту й емоцій у коментарях користувачів (чи їх досліджених комбінацій) таке: $M_j^{(2)}, \left(j = \overline{1, r} \right)$.

Авторська модель являє собою поєднання однієї з моделей з масиву $M_i^{(1)}, \left(i = \overline{1, l} \right)$ із однією з моделей масиву $M_j^{(2)}, \left(j = \overline{1, r} \right)$. Позначимо можливі авторські моделі як $M_k = M_i^{(1)} \cup M_j^{(2)}, \left(k = \overline{1, l \cdot r} \right)$.

Нехай кожен з алгоритмів $M_i^{(1)}, \left(i = \overline{1, l} \right)$ на визначеному наборі даних має певну ефективність $E_i^{(1)}, \left(i = \overline{1, l} \right)$, а кожен з алгоритмів $M_j^{(2)}, \left(j = \overline{1, r} \right)$ ефективність $E_j^{(2)}, \left(j = \overline{1, r} \right)$.

Зазначені ефективності можуть бути знайдені з урахуванням формул (1)-(5), наприклад у вигляді

$$E_i^{(1)} = g_i^{(1)}(A_{accuracy.i}^{(1)}, P_{recision.i}^{(1)}, R_{ecall.i}^{(1)}, F1_i^{(1)}, AUC_i^{(1)}), \quad (15)$$

$$E_j^{(2)} = g_j^{(2)}(A_{accuracy.j}^{(2)}, P_{recision.j}^{(2)}, R_{ecall.j}^{(2)}, F1_j^{(2)}, AUC_j^{(2)}). \quad (16)$$

Із урахуванням (15)-(16) ефективність авторських моделей $M_k, \left(k = \overline{1, l \cdot r}\right)$ може бути подана, як

$$E_k = f(E_i^{(1)}, E_j^{(2)}). \quad (17)$$

Зважаючи на фізичний зміст задачі виявлення фейкових новин, при опрацюванні авторської моделі необхідним є підбір такої моделі з масиву $M_k = M_i^{(1)} \cup M_j^{(2)}, \left(k = \overline{1, l \cdot r}\right)$, застосування якої дозволило б у межах заданого часу розв'язати задачу, а також не перевищити задану обчислювальну складність алгоритму.

Нехай для кожної моделі $M_i^{(1)}, \left(i = \overline{1, l}\right)$ відомими є величини:

$T_i^{(1)}$ - максимальна тривалість роботи моделі $M_i^{(1)}, \left(i = \overline{1, l}\right)$ на визначеному досліджуваному масиві даних;

$C_i^{(1)}$ - обчислювальна складність алгоритму $M_i^{(1)}, \left(i = \overline{1, l}\right)$.

А для кожної моделі $M_j^{(2)}, \left(j = \overline{1, r}\right)$ ці величини такі:

$T_j^{(2)}$ - максимальна тривалість роботи моделі $M_j^{(2)}, \left(j = \overline{1, r}\right)$ на визначеному досліджуваному масиві даних;

$C_j^{(2)}$ - обчислювальна складність алгоритму $M_j^{(2)}, \left(j = \overline{1, r}\right)$.

Тоді загальний час роботи $M_k, \left(k = \overline{1, l \cdot r}\right)$ можна представити у вигляді

$$T_k = \tilde{f}(T_i^{(1)}, T_j^{(2)}). \quad (18)$$

З урахуванням введених позначень, обчислювальну складність авторської моделі можна описати функціональною залежністю

$$C_k = \bar{f}(C_i^{(1)}, C_j^{(2)}). \quad (19)$$

А отже, задачу побудови авторської моделі можна сформулювати, як оптимізаційну задачу вибору з числа алгоритмів $M_k, \left(k = \overline{1, l \cdot r}\right)$ того з них, який би забезпечував

$$E_k \rightarrow \max, \quad (20)$$

де E_k визначається за формулою (17), при виконанні умов

$$T_k \leq T_{k.ep}, \quad (21)$$

$$C_k \leq C_{k.ep}, \quad (22)$$

де T_k і C_k знаходяться, відповідно, з формул (18) і (19).

У виразах (21), (22) $T_{k.ep}$, $C_{k.ep}$ - це задані граничні значення тривалості роботи

авторської моделі $M_k, \left(k = \overline{1, l \cdot r}\right)$ на масиві $A = \{a_1, a_2, \dots, a_k\}$ і обчислювальної складності моделі.

Слід зауважити, що в авторській моделі (1)-(5), (15)-(22) виявлення фейкових новин невідомими є функціональні залежності $g_i^{(1)}, g_j^{(2)}, f, \tilde{f}, \bar{f}$. Їх встановлення є окремим завданням.

Висновки з даного дослідження і перспективи подальшого розвитку. Таким чином, у ході дослідження проаналізовано ряд положень, які дозволили сформулювати авторські ідеї щодо можливих підходів до формування нових моделей виявлення фейкових новин. З урахуванням проведеного аналізу можна зробити висновок, що ці ідеї можуть мати перспективу щодо можливого поліпшення існуючих результатів. Проведена формалізація авторських ідей дозволяє наблизитися до формування механізмів перевірки сформульованих гіпотез.

Перспективами подальших розвідок вбачається програмування та вирішення сформульованих авторських оптимізаційних задач, що стосуються формування удосконалених моделей виявлення фейкових новин, а також вирішення окремих часткових завдань стосовно визначення функціональних залежностей, які визначають окремі складові цих моделей.

ЛІТЕРАТУРА (REFERENCES):

1. Bondielli, A.; Marcelloni, F. A survey on fake news and rumour detection techniques. *Inf. Sci.* 2019, 497, 38–55. [CrossRef]
2. Islam, M.R.; Liu, S.; Wang, X.; Xu, G. Deep learning for misinformation detection on online social networks: A survey and new perspectives. *Soc. Netw. Anal. Min.* 2020, 10, 82. [CrossRef] [PubMed]
3. Bahad, P.; Saxena, P.; Kamal, R. Fake News Detection using Bi-directional LSTM-Recurrent Neural Network. *Procedia Comput. Sci.* 2019, 165, 74–82. [CrossRef]
4. Machová, K.; Mach, M.; Porezaný, M. Deep Learning in the Detection of Disinformation about COVID-19 in Online Space. *Sensors* 2022, 22, 9319. [CrossRef]
5. Liu, Y.; Wu, Y.-F.B. Fned: A deep network for fake news early detection on social media. *ACM Trans. Inf. Syst. (TOIS)* 2020, 38, 1–33. [CrossRef]
6. Zhou, X.; Jain, A.; Phoha, V.V.; Zafarani, R. Fake News Early Detection: An Interdisciplinary Study. *arXiv* 2019, arXiv:1904.11679.
7. DataReportal. Digital 2021 Global Digital Overview; DataReportal: Singapore, 2021.
8. Kwak, H.; Lee, C.; Park, H.; Moon, S. What is Twitter, a Social Network or a News Media? In *Proceedings of the 19th International Conference on World Wide Web; Association for Computing Machinery: New York, NY, USA, 2010; pp. 591–600.* [CrossRef]
9. Friggeri, A.; Adamic, L.; Eckles, D.; Cheng, J. Rumor cascades. In *Proceedings of the International AAAI Conference on Web and Social Media, Ann Arbor, MI, USA, 1–4 June 2014; Volume 8.*
10. Zhou, X.; Zafarani, R. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Comput. Surv.* 2020, 53, 1–40. [CrossRef]
11. Conroy, N.K.; Rubin, V.L.; Chen, Y. Automatic deception detection: Methods for finding fake news. *Proc. Assoc. Inf. Sci. Technol.* 2015, 52, 1–4. [CrossRef]
12. Habib, A.; Asghar, M.Z.; Khan, A.; Habib, A.; Khan, A. False information detection in online content and its role in decision making: A systematic literature review. *Soc. Netw. Anal. Min.* 2019, 9, 50. [CrossRef]
13. Sansonetti, G.; Gasparetti, F.; D'aniello, G.; Micarelli, A. Unreliable Users Detection in Social Media: Deep Learning Techniques for Automatic Detection. *IEEE Access* 2020, 8, 213154–213167. [CrossRef]

14. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. *Science* 2018, 359, 1146–1151. [CrossRef]
15. Zhao, Z.; Resnick, P.; Mei, Q. Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts. In *Proceedings of the 24th International Conference on World Wide Web; International World Wide Web Conferences Steering Committee: Geneva, Switzerland, 2015*; pp. 1395–1405. [CrossRef]
16. Kumar, S.; West, R.; Leskovec, J. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *Proceedings of the 25th International Conference on World Wide Web, Montreal, QC, Canada, 11–15 April 2016*; pp. 591–602.
17. Potthast, M.; Kiesel, J.; Reinartz, K.; Bevendorff, J.; Stein, B. A stylometric inquiry into hyperpartisan and fake news. *arXiv* 2017, arXiv:1702.05638.
18. Tchakounté, F.; Calvin, K.A.; Ari, A.A.A.; Mbogne, D.J.F.J.J.o.K.S.U.-C.; Sciences, I. A smart contract logic to reduce hoax propagation across social media. *J. King Saud Univ.-Comput. Inf. Sci.* 2020, 34, 3070–3078. [CrossRef]
19. Rath, B.; Gao, W.; Ma, J.; Srivastava, J. Utilizing computational trust to identify rumor spreaders on Twitter. *Soc. Netw. Anal. Min.* 2018, 8, 64. [CrossRef]
20. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. *Science* 2018, 359, 1146–1151. [CrossRef]
21. Al-Rakhami, M.S.; Al-Amri, A.M. Lies Kill, Facts Save: Detecting COVID-19 Misinformation in Twitter. *IEEE Access* 2020, 8, 155961–155970. [CrossRef]
22. Alkhodair, S.A.; Ding, S.H.; Fung, B.C.; Liu, J. Detecting breaking news rumors of emerging topics in social media. *Inf. Process. Manag.* 2020, 57, 102018. [CrossRef]
23. Zubair, T.; Raquib, A.; Qadir, J. Combating Fake News, Misinformation, and Machine Learning Generated Fakes: Insight’s from the Islamic Ethical Tradition. *ICR J.* 2019, 10, 189–212. [CrossRef]
24. Allcott, H.; Gentzkow, M. Social media and fake news in the 2016 election. *J. Econ. Perspect.* 2017, 31, 211–236. [CrossRef]
25. Shu, K.; Sliva, A.; Wang, S.; Tang, J.; Liu, H. Fake News Detection on Social Media: A Data Mining Perspective. *arXiv* 2017, arXiv:1708.01967.
26. Trends, G. “Fake News—Explore—Google Trends”. Available online: <https://trends.google.com/trends/explore?date=2010-01-01%202022-07-14&q=%22fake%20news%22> (accessed on 20 July 2022).
27. Langin, K. Fake news spreads faster than true news on Twitter—Thanks to people, not bots. *Sci. Mag.* 2018. Available online: <https://www.science.org/content/article/fake-news-spreads-faster-true-news-twitter-thanks-people-not-bots> (accessed on 20 February 2022).
28. Zubiaga, A.; Aker, A.; Bontcheva, K.; Liakata, M.; Procter, R. Detection and Resolution of Rumours in Social Media: A Survey. *ACM Comput. Surv.* 2018, 51, 1–36. [CrossRef]
29. Evolvi, G. Hate in a tweet: Exploring internet-based islamophobic discourses. *Religions* 2018, 9, 307. [CrossRef]
30. Al-Makhadmeh, Z.; Tolba, A. Automatic hate speech detection using killer natural language processing optimizing ensemble deep learning approach. *Computing* 2020, 102, 501–522. [CrossRef]
31. Feng, S.; Banerjee, R.; Choi, Y. Syntactic Stylometry for Deception Detection. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*; Association for Computational Linguistics: Jeju Island, Republic of Korea, 2012; pp. 171–175.
32. de Oliveira, N.R.; Medeiros, D.S.; Mattos, D.M. A sensitive stylistic approach to identify fake news on social networking. *IEEE Signal Process. Lett.* 2020, 27, 1250–1254. [CrossRef]
33. Zhou, X.; Jain, A.; Phoha, V.V.; Zafarani, R. Fake news early detection: A theory-driven model. *Digit. Threat. Res. Pract.* 2020, 1, 1–25. [CrossRef]
34. Lin, L.; Chen, Z. Social rumor detection based on multilayer transformer encoding blocks. *Concurr. Comput. Pract. Exp.* 2021, 33, e6083. [CrossRef]

35. Wang, W.Y. “Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection. arXiv 2017, arXiv:1705.00648.
36. Patwa, P.; Sharma, S.; Pykl, S.; Guptha, V.; Kumari, G.; Akhtar, M.S.; Ekbal, A.; Das, A.; Chakraborty, T. Fighting an infodemic: COVID-19 fake news dataset. In Proceedings of the International Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation; Springer: Berlin/Heidelberg, Germany, 2021; pp. 21–29.
37. Eke, C.I.; Norman, A.A.; Shuib, L.; Nweke, H.F. Sarcasm identification in textual data: Systematic review, research challenges and open directions. *Artif. Intell. Rev.* 2020, 53, 4215–4258. [CrossRef]
38. Alonso, M.A.; Vilares, D.; Gómez-Rodríguez, C.; Vilares, J. Sentiment analysis for fake news detection. *Electronics* 2021, 10, 1348. [CrossRef]
39. Elhadad, M.K.; Li, K.F.; Gebali, F. Detecting Misleading Information on COVID-19. *IEEE Access* 2020, 8, 165201–165215. [CrossRef]
40. Alghamdi, O.; Lin, Y.; Luo, S.: A comparative study of machine learning and deep learning techniques for fake news detection. *Information* 2022, 13, 576, 28 p.
41. Guo, M.; Xu, Z.; Liu, L.; Guo, M.; Zhang, Y. An Adaptive Deep Transfer Learning Model for Rumor Detection without Sufficient Identified Rumors. *Math. Probl. Eng.* 2020, 2020, 7562567. [CrossRef]
42. Varshney, D.; Vishwakarma, D.K. Vishwakarma, Hoax news-inspector: A real-time prediction of fake news using content resemblance over web search results for authenticating the credibility of news articles. *J. Ambient Intell. Humaniz. Comput.* 2020, 12, 8961–8974. [CrossRef]
43. Kim, Y.; Kim, H.K.; Kim, H.; Hong, J.B. Do Many Models Make Light Work? Evaluating Ensemble Solutions for Improved Rumor Detection. *IEEE Access* 2020, 8, 150709–150724. [CrossRef]
44. Yaakub, M.R.; Latiffi, M.I.A.; Zaabar, L.S. A review on sentiment analysis techniques and applications. *IOP Conf. Ser. Mater. Sci. Eng.* 2019, 551, 012070. [CrossRef]
45. Santhoshkumar, S.; Babu, L.D. Earlier detection of rumors in online social networks using certainty-factor-based convolutional neural networks. *Soc. Netw. Anal. Min.* 2020, 10, 20. [CrossRef]
46. Tian, L.; Zhang, X.; Wang, Y.; Liu, H. Early detection of rumours on twitter via stance transfer learning. In *Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, 14–17 April 2020, Proceedings, Part I 42*; Springer: Cham, Switzerland, 2020; Volume 12035, p. 575.
47. Albahar, M. A hybrid model for fake news detection: Leveraging news content and user comments in fake news. *IET Inf. Secur.* 2021, 15, 169–177. [CrossRef]
48. Ghanem, B.; Rosso, P.; Rangel, F. An emotional analysis of false information in social media and news articles. *ACM Trans. Internet Technol. (TOIT)* 2020, 20, 1–18. [CrossRef]
49. Kumari, R.; Ashok, N.; Ghosal, T.; Ekbal, A. What the fake? Probing misinformation detection standing on the shoulder of novelty and emotion. *Inf. Process. Manag.* 2022, 59, 102740. [CrossRef]
50. Zhang, X.; Cao, J.; Li, X.; Sheng, Q.; Zhong, L.; Shu, K. Mining dual emotion for fake news detection. In *Proceedings of the WWW '21: The Web Conference 2021, Ljubljana, Slovenia, 19–23 April 2021*; pp. 3465–3476.
51. Zimbra, D.; Abbasi, A.; Zeng, D.; Chen, H. The state-of-the-art in Twitter sentiment analysis: A review and benchmark evaluation. *ACM Trans. Manag. Inf. Syst. (TMIS)* 2018, 9, 1–29. [CrossRef]
52. Feng, Z. Hot news mining and public opinion guidance analysis based on sentiment computing in network social media. *Pers. Ubiquitous Comput.* 2019, 23, 373–381. [CrossRef]
53. Imran, A.S.; Daudpota, S.M.; Kastrati, Z.; Batra, R. Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets. *IEEE Access* 2020, 8, 181074–181090. [CrossRef]
54. Pota, M.; Ventura, M.; Catelli, R.; Esposito, M. An effective BERT-based pipeline for Twitter sentiment analysis: A case study in Italian. *Sensors* 2020, 21, 133. [CrossRef]

55. Dang, C.N.; Moreno-García, M.N.; Prieta, F.D.L. An approach to integrating sentiment analysis into recommender systems. *Sensors* 2021, 21, 5666. [CrossRef]
56. Li, Q.; Hu, Q.; Lu, Y.; Yang, Y.; Cheng, J. Multi-level word features based on CNN for fake news detection in cultural communication. *Pers. Ubiquitous Comput.* 2020, 24, 259–272.
57. Correia, F.; Madureira, A.M.; Bernardino, J. Deep Neural Networks Applied to Stock Market Sentiment Analysis. *Sensors* 2022, 22, 4409. [CrossRef] [PubMed]
58. Subramani, S.; Wang, H.; Vu, H.Q.; Li, G. Domestic violence crisis identification from facebook posts based on deep learning. *IEEE Access* 2018, 6, 54075–54085. [CrossRef]
59. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* 1997, 9, 1735–1780. [CrossRef]
60. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv 2014, arXiv:1406.1078.
61. Zhang, X.; Chen, F.; Huang, R. A combination of RNN and CNN for attention-based relation classification. *Procedia Comput. Sci.* 2018, 131, 911–917. [CrossRef]
62. Zhou, C.; Sun, C.; Liu, Z.; Lau, F. A C-LSTM neural network for text classification. arXiv 2015, arXiv:1511.08630.
63. Gururangan, S.; Marasovi'c, A.; Swayamdipta, S.; Lo, K.; Beltagy, I.; Downey, D.; Smith, N.A. Don't stop pretraining: Adapt language models to domains and tasks. arXiv 2020, arXiv:2004.10964.
64. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv 2018, arXiv:1810.04805.
65. Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. Roberta: A robustly optimized bert pretraining approach. arXiv 2019, arXiv:1907.11692.
66. Khan, J.Y.; Khondaker, M.T.I.; Afroz, S.; Uddin, G.; Iqbal, A. A benchmark study of machine learning models for online fake news detection. *Mach. Learn. Appl.* 2021, 4, 100032. [CrossRef]
67. Horne, L.; Matti, M.; Pourjafar, P.; Wang, Z. GRUBERT: A GRU-Based Method to Fuse BERT Hidden Layers for Twitter Sentiment Analysis. In Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing: Student Research Workshop; Association for Computational Linguistics: Suzhou, China, 2020; pp. 130–138
68. Nakamura, K.; Levy, S.; Wang, W.Y. r/fakeddit: A new multimodal benchmark dataset for fine-grained fake news detection. arXiv 2019, arXiv:1911.03854.
69. Ajao, O.; Bhowmik, D.; Zargari, S. Sentiment aware fake news detection on online social networks. In Proceedings of the ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 2507–2511.
70. Bhutani, B.; Rastogi, N.; Sehgal, P.; Purwar, A. Fake news detection using sentiment analysis. In Proceedings of the 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India, 8–10 August 2019; pp. 1–5.
71. Giachanou, A.; Rosso, P.; Crestani, F. Leveraging emotional signals for credibility detection. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Paris, France, 21–25 July 2019; pp. 877–880.
72. Kumar, S.; Asthana, R.; Upadhyay, S.; Upreti, N.; Akbar, M. Fake news detection using deep learning models: A novel approach. *Trans. Emerg. Telecommun. Technol.* 2020, 31, e3767. [CrossRef]
73. Kaliyar, R.K.; Kumar, P.; Kumar, M.; Narkhede, M.; Namboodiri, S.; Mishra, S. DeepNet: An efficient neural network for fake news detection using news-user engagements. In Proceedings of the 2020 5th International Conference on Computing, Communication and Security (ICCCS), Patna, India, 14–16 October 2020; pp. 1–6.

74. Kirchknopf, A.; Slijepčević, D.; Zeppelzauer, M. Multimodal Detection of Information Disorder from Social Media. In Proceedings of the 2021 International Conference on Content-Based Multimedia Indexing (CBMI), Lille, France, 28–30 June 2021; pp. 1–4.
75. Xie, J.; Liu, S.; Liu, R.; Zhang, Y.; Zhu, Y. SeRN: Stance extraction and reasoning network for fake news detection. In Proceedings of the ICASSP 2021—2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 2520–2524.
76. Raza, S.; Ding, C. Fake news detection based on news content and social contexts: A transformer-based approach. *Int. J. Data Sci. Anal.* 2022, 13, 335–362. [CrossRef]

Barmak Oleksandr, Borovyk Oleh, Borovyk Dmytro, Skrypnyk Tetyana

POSSIBLE OPTIONS FOR IMPROVING EXISTING APPROACHES TO FAKE NEWS DETECTION BASED ON USING THE POTENTIAL OF MACHINE AND DEEP LEARNING ALGORITHMS, SENTIMENT OF NEWS CONTENT AND EMOTIONS IN USER COMMENTS

Currently, the Internet ranks first among sources of information. In the recent period, the role of online social networks (OSN) has significantly increased, which has both positive and negative consequences. The negative role of OSN is related to the spread of fake news that affects people's daily lives, manipulates their thoughts and feelings, changes their beliefs and can lead to wrong decisions. The problem of spreading fake news in OSN is currently global, and the formation of countermeasures is an urgent task today.

Today, there are various proven approaches to detecting fake news. In particular, one of the approaches is based on the use of different machine (ML) and deep (DL) learning algorithms. The other is based on the results of sentiment analysis of news content and analysis of emotions in user comments. The research conducted by the authors of other approaches to detecting fake news, which differ from the ones given, made it possible to conclude that the mentioned approaches are effective and promising in terms of using their potential for the development of new models with high performance indicators on various data sets.

In the article, the author's ideas regarding the improvement of existing approaches to detecting fake news based on the use of the potential of these approaches are formed and formalized. The first idea is based on the implementation of the mechanism of combining machine (ML) and deep (DL) learning methods, as well as the results of the analysis of the sentiment of news content and emotions in user comments, which takes into account the possibility of ensuring a sufficient level of effectiveness in detecting fake news, a certain level of the values of the selected metrics, as well as a certain level of functional characteristics of the author's method. The second idea is based on the implementation of a mechanism combining the functionality of two methods from among the specified two groups, which would provide optimal parameters for detecting fake news according to defined criteria and indicators.

The substantiation of the ideas involved the preliminary implementation of: setting the researched problem; functional analysis of machine (ML) and deep (DL) learning algorithms, as well as fake news detection algorithms based on the use of the results of sentiment analysis of news content and emotions in user comments; description of metrics for evaluating the effectiveness of methods for detecting fake news. According to the results of the substantiation of the perspective of the ideas, the tasks of detecting fake news in the author's production were formalized.

Keywords: online social networks; fake news; method; model; algorithm; formalization.