

УДК 316.774

DOI <https://doi.org/10.32782/2707-9147.2025.106.16>

І. В. ПИГОЛЕНКО

кандидат філософських наук, доцент,
доцент кафедри соціології

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

В. А. ГАСАНОВ

магістрант факультету соціології і права

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

ВИКОРИСТАННЯ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ СОЦІАЛЬНИХ МЕДІА

Стаття присвячена розгляду використання машинного навчання, для аналізу сучасних соціальних медіа, що дає можливість кластеризувати теми обговорень, виявляти фейковий контент та візуалізувати взаємозв'язки між користувачами. Соціальні мережі, як цифрові структури взаємодії користувачів, надають унікальні можливості для дослідження соціальних процесів, обміну інформацією, формування спільнот, а також аналізу впливу та зв'язків між учасниками.

У статті обґрунтована актуальність та необхідність використання алгоритмів машинного навчання у сфері аналізу соціальних медіа, оскільки саме ці алгоритми дозволяють виявляти приховані зв'язки, теми, патерни поведінки та фейковий контент. Обґрунтовано важливість вдосконалення систем аналізу соціальних медіа на основі машинного навчання як для наукових досліджень, так і для практичного застосування у сферах безпеки, журналістики, маркетингу та громадянського суспільства.

В мовах розвитку цифрового суспільства зростає потреба в інтелектуальних автоматизованих системах, здатних ефективно аналізувати великі обсяги даних у режимі реального часу, що дає можливість покращити продуктивність та оперативність реагування в критичних ситуаціях.

В статті здійснено аналіз використання алгоритмів машинного навчання для аналізу текстових даних, зокрема його підгалузі як-от обробка природної мови (NLP) та глибоке навчання, що забезпечує широкий інструментарій для реалізації цих задач.

В статті звернуто увагу на необхідності модернізації класичних соціологічних методів роботи з інформацією, у зв'язку з постійним зростанням величезних масивів неструктурованих даних.

Попри значну кількість наукових досліджень у сфері новітніх соціальних медіа, ця тема залишається актуальною, у зв'язку з постійним розвитком засобів комунікації, соціальних медіа та соціальних мереж, що

впливає на трансформацію соціального простору у сучасному цифровому суспільстві.

Ключові слова: соціальні медіа, соціальні мережі, комунікація, машинне навчання, аналіз даних, дослідження.

Постановка проблеми. У сучасному цифровому суспільстві соціальні медіа перетворилися на головну платформу для комунікації, обміну інформацією, формування громадської думки та впливу на соціальні процеси. Обсяги даних, які генеруються користувачами щоденно, перевищують мільярди одиниць контенту, що створює безпрецедентні можливості для аналізу соціальних взаємодій, але одночасно породжує нові виклики щодо обробки, інтерпретації та верифікації інформації. Саме тому зростає потреба в інтелектуальних автоматизованих системах, здатних ефективно аналізувати великі обсяги даних у режимі реального часу.

Особливої актуальності набуває використання алгоритмів машинного навчання у сфері аналізу соціальних медіа, оскільки саме ці алгоритми дозволяють виявляти приховані зв'язки, теми, патерни поведінки та фейковий контент. Розробка і оптимізація систем аналізу соціальних медіа на основі машинного навчання є надзвичайно важливою як для наукового дослідження, так і для практичного застосування у сферах безпеки, журналістики, маркетингу та громадянського суспільства.

Попри значні досягнення в галузі автоматизації аналізу соціальних медіа, існують численні проблеми, пов'язані з якістю класифікації даних, виявленням дезінформації, нестабільністю алгоритмів при зміні контексту та мовного середовища. Система аналізу повинна не лише бути здатною швидко обробляти вхідні потоки даних, а й гнучко адаптуватися до нових тенденцій – від нових форматів повідомлень до зміни інформаційних патернів поширення новин. Також, є недостатня адаптованість існуючих алгоритмів до специфіки українського простору соціальних медіа.

Аналіз останніх досліджень і публікацій. Питання аналізу соціальних медіа активно досліджується в роботах українських вчених Веселяк В.В., Грицюк Ю.І., Дереторська Т.І., Дорошенко І.В., Кнігніцька Т.В., Лобовікова О.О., Мельников А.С. та ін. та іноземних Fatima M., Mitchell T., Quinlan J. R., Witten I. H., Frank E., Pasha M., Sarker I.H. та ін.

Метою даної статті є розгляд використання машинного навчання, для аналізу сучасних соціальних медіа в умовах розвитку цифрового суспільства.

Виклад основного матеріалу. Соціальні медіа (англ. Social media) – вид мас-медіа, ряд онлайн-ових технологій на принципах Веб 2.0, завдяки яким споживачі контенту через свої дописи стають його співавторами і можуть взаємодіяти, співпрацювати, спілкуватися, ділитися інформацією або брати участь у будь-якій іншій

соціальній активності із теоретично усіма іншими користувачами певного сервісу [7].

Соціальні медіа, як об'єкт дослідження, надають величезний масив неструктурованих даних, для роботи з якими класичні соціологічні методи потребують модернізації.

Соціальні медіа є частиною повсякденного життя. Це надає користувачам простий спосіб взаємодії та спілкування з іншими, не зустрічаючись віч-на-віч. Ця форма спілкування надає багато можливостей компаніям і окремим особам для охоплення великої аудиторії. Переваги соціальних медіа впливають з активної участі та створення нового привабливого контенту.

В Україні соціальні медіа набули значної популярності. Серед найбільш використовуваних платформ – Facebook, Instagram та YouTube. Вони стали важливими каналами для спілкування, отримання новин та ведення бізнесу. Значна частина українців використовує соціальні медіа для отримання інформації та взаємодії.

Також, варто відзначити використання цифрової етнографії (нетнографії), що дозволяє дослідити поведінку онлайн-спільнот у природному середовищі. У поєднанні з методами машинного навчання ці підходи дозволяють формувати складні моделі аналізу громадської думки, динаміки конфліктів, впливу фейків на громадську думку тощо.

Сучасні дослідження соціальних медіа доводять, що ці платформи не просто відображають громадську думку, а й активно формують її. У соціальних медіа відбувається обговорення важливих політичних, соціальних та культурних тем. Користувачі не лише споживають інформацію, а й стають її творцями. Це змінює структуру суспільної комунікації – від ієрархічної (ЗМІ → споживач) до мережевої (учасники ↔ учасники). Саме в соціальних медіа відбувається політична мобілізація, організація протестів, кампанії за чи проти соціальних ініціатив, що свідчить про трансформацію ролі громадян у суспільному житті.

Особливої ваги це набуває в періоди політичних криз або воєнних конфліктів, коли соціальні медіа стають не тільки джерелом інформації, а й інструментом впливу.

Система аналізу соціальних медіа – це цілісний технологічний комплекс, побудований на принципах гнучкої модульності, масштабованості та безпеки. Її структура передбачає логічне розмежування функцій, що дозволяє ефективно розв'язувати завдання від збору та підготовки даних до глибокого аналізу та прийняття рішень. Якісне проектування архітектури системи забезпечує її стійкість до навантажень, адаптивність до змін середовища та здатність до інтеграції з новими джерелами та аналітичними інструментами. Такі системи є незамінними інструментами у сфері інформаційного аналізу, здатними надати глибоке розуміння поведінки онлайн-аудиторії та прогнозувати її реакції на ті чи інші події. Їх використання сприяє підвищенню ефективності

комунікацій, оптимізації маркетингових стратегій і прийняттю обґрунтованих управлінських рішень у цифрову епоху.

Вплив соціальних медіа на різноманітні соціальні процеси та практики все більше стає у центрі уваги соціальних та гуманітарних наук.

Соціальні та гуманітарні науки більше цікавить роль соціальних медіа в контексті загальної соціальної реальності, а не їх технічна функціональність. Вивчення цієї ролі може базуватися на загальному уявленні про соціальні медіа як про категорію інтернет-сайтів, які дозволяють користувачам самостійно створювати інформаційний контент. Очевидною проблемою сучасних теоретичних досліджень соціальних медіа є відносний еkleктизм поточних наукових досліджень.

Сучасні соціальні і гуманітарні науки базуються на принципі поліпарадигмальності, тому традиції аналізу одного і того ж явища в ній, як правило, дуже різноманітні. У даному випадку ситуація ускладнюється надзвичайно високою динамікою розвитку об'єкта дослідження – соціальних медіа, які регулярно набувають нових властивостей та інструментів.

Аналіз текстових даних за допомогою алгоритмів машинного навчання є однією з найдинамічніших і водночас найскладніших сфер сучасної обробки інформації. У цифрову епоху обсяг текстового контенту, що генерується користувачами щодня, стрімко зростає, охоплюючи соціальні мережі, форуми, відгуки, електронні листи, новини, блоги тощо. Це зумовлює необхідність ефективних методів аналізу, які здатні автоматично обробляти великі обсяги тексту, виділяти з нього значущу інформацію, здійснювати класифікацію, оцінку тональності, виявляти приховані теми, передбачати поведінку користувачів та інші аналітичні завдання. Машинне навчання, зокрема його підгалузі як-от обробка природної мови (NLP) та глибоке навчання, забезпечує широкий інструментарій для реалізації цих задач [9].

Однією з найважливіших переваг машинного навчання у роботі з текстами є здатність алгоритмів виявляти закономірності в неструктурованих даних без явних інструкцій від аналітика. Це відкриває можливість адаптивного аналізу, що змінюється відповідно до тематики тексту, лексичних особливостей, контексту та стилістики. Для цього застосовуються різні типи алгоритмів: від класичних моделей, як-от наївний баєсівський класифікатор чи логістична регресія, до сучасних архітектур нейронних мереж, таких як LSTM, Transformer або BERT. Залежно від мети дослідження, алгоритми підбираються таким чином, щоб оптимально відображати логіку обробки тексту: класифікація, тематичне моделювання, виявлення аномалій або прогнозування змін у поведінці.

Підготовка тексту до машинного аналізу є критично важливим етапом. Текстові дані, зазвичай, є зашумленими, містять граматичні помилки, нетипові конструкції, жаргон, а також можуть бути різними

мовами. Стандартизований процес попередньої обробки включає токенізацію (розбиття на слова або фрази), лематизацію (зведення слів до базової форми), видалення стоп-слів (часток, прийменників, займенників), нормалізацію (перетворення тексту до нижнього регістру), а також кодування (наприклад, через TF-IDF або word embeddings). Вибір методів трансформації тексту безпосередньо впливає на якість побудованих моделей, а, отже, на точність прогнозів і висновків [17].

Успішність роботи моделі визначається не лише архітектурою, а й правильністю навчання. Дослідник має підібрати відповідну кількість навчальних прикладів, забезпечити баланс класів, уникати переобучення (overfitting), використовувати крос-валідацію, а також обирати метрики, що відповідають специфіці задачі: точність (accuracy), повноту (recall), специфічність, F1-міру тощо. Важливо також вчасно оновлювати моделі відповідно до змін контенту в соціальних мережах, оскільки мовні тренди швидко змінюються [14].

У рамках прикладних задач, алгоритми машинного навчання для тексту все частіше застосовуються у маркетингових дослідженнях. Компанії використовують класифікацію відгуків для визначення задоволеності клієнтів, прогнозування намірів купівлі, моніторингу конкурентів. У сфері HR – для автоматичного аналізу резюме, відбору кандидатів. У юридичних системах – для класифікації документів, виявлення ключових фрагментів у договорах. Журналістика використовує тематичне групування новин, пошук фейкових матеріалів та прогнозування резонансу подій у мережі. У державному управлінні – для аналізу громадської думки, виявлення загроз, формування антикризової політики.

Реалізація таких систем передбачає не лише побудову моделі, а й створення інтерфейсів, що дозволяють користувачам взаємодіяти з результатами. Для цього створюються панелі керування, де можна фільтрувати текстові дані за часом, категоріями, джерелами, використовувати пошук за ключовими словами, переглядати найпопулярніші теми або найбільш згадуваних осіб. Часто аналітика поєднується з візуалізацією: хмари слів, графи, інтерактивні часові діаграми [12].

У застосуванні машинного навчання до текстових даних виникають і певні виклики. Насамперед – це мовна неоднорідність даних, наявність синонімів, полісемії, контекстуальних значень, метафор і іронії, що ускладнює точне розпізнавання змісту. Алгоритми часто виявляють упередження, оскільки навчаються на нерівномірно представлених даних. Наприклад, недостатнє представлення жіночого голосу в тренувальній вибірці може призвести до гендерного викривлення результатів. Також, складною є задача інтерпретації результатів – деякі моделі, особливо глибокі нейронні мережі, діють як «чорні скриньки», і аналітик не завжди може пояснити, чому модель зробила певне передбачення [10].

Не менш критичною проблемою є упередженість у даних, яка може призвести до дискримінаційних висновків і несправедливих рішень. Вона виникає через нерівномірне представлення груп у тренувальних наборах, і вирішується шляхом ретельного тестування моделей на збалансованих вибірках. Брак інтерпретованості глибоких моделей часто спричиняє недовіру до результатів, що особливо важливо в сферах, де потрібна прозорість – у таких випадках застосовують інструменти Explainable AI, зокрема, LIME або SHAP, які пояснюють логіку моделі.

Нарешті, порушення конфіденційності є потенційним джерелом юридичних ризиків, особливо при роботі з публічними текстами із соціальних мереж. Щоб уникнути таких наслідків, слід застосовувати анонімізацію даних та узгоджувати обробку з відповідними політиками конфіденційності.

Отже, усвідомлення цих ризиків та активне використання сучасних технологічних і етичних рішень дозволяє зробити аналіз текстових даних більш надійним, точним і відповідальним.

Окрему увагу слід приділяти етичним аспектам. Обробка текстових даних, особливо у випадках із соціальними мережами, вимагає дотримання конфіденційності, узгодження з політикою GDPR та іншими нормативами. Збір та аналіз публічної інформації має здійснюватися з урахуванням прав користувачів та уникнення дискримінаційного використання даних. Важливо також забезпечувати прозорість алгоритмів і можливість зворотного зв'язку з боку кінцевих користувачів.

Соціальні мережі, як цифрові структури взаємодії користувачів, надають унікальні можливості для дослідження соціальних процесів, обміну інформацією, формування спільнот, а також аналізу впливу та зв'язків між учасниками. Дослідження структур соціальних мереж є важливою частиною соціальної інформатики, обробки великих даних, прикладної соціології, а також цифрового маркетингу. Такий аналіз передбачає використання спеціальних методів і моделей, які дають змогу розглядати мережу як систему взаємопов'язаних елементів, що взаємодіють за певними законами та утворюють складну ієрархію зв'язків [16].

Початковим етапом дослідження є побудова графа соціальної мережі. Користувачі або інші об'єкти (сторінки, групи, події) розглядаються як вузли, а взаємодії між ними – як ребра графа. Залежно від характеру взаємодії, ребра можуть бути напрямленими (наприклад, підписка в Twitter) або ненапрямленими (взаємне додавання у друзі на Facebook). Крім того, зв'язки можуть бути зваженими – із зазначенням сили або частоти взаємодії. Отриманий граф дозволяє провести подальший кількісний та якісний аналіз структури, виявити закономірності в розподілі зв'язків, рівень згуртованості мережі, наявність кластерів або ключових акторів.

Одним з основних напрямів є аналіз центральності, тобто виявлення найбільш важливих вузлів у мережі. Існує кілька типів центральності: ступенева (degree centrality), що враховує кількість зв'язків; проміжкова (betweenness centrality), яка вимірює, скільки разів вузол потрапляє на найкоротший шлях між іншими вузлами; та близькість (closeness centrality), що оцінює середню відстань від вузла до всіх інших у мережі. Такі метрики дозволяють виявити лідерів думок, інфлюенсерів, координаторів спільнот або, навпаки, периферійних користувачів. Для візуалізації використовується розташування вузлів за центральністю, що дозволяє інтуїтивно зрозуміти топологію мережі [15].

Ще одним важливим аспектом є виявлення спільнот у мережі. Користувачі часто формують групи або кластери на основі спільних інтересів, тем обговорень, географічної приналежності або типу контенту. Для автоматичного виявлення таких спільнот використовуються алгоритми кластеризації графів: модульність Ньюмана, алгоритм Louvain, spectral clustering та інші. Виявлення спільнот дозволяє ефективніше таргетувати рекламні кампанії, виявляти локальних лідерів думок та прогнозувати поширення інформації. Також, це корисно у виявленні дезінформаційних мереж, бот-мереж або організованих кампаній впливу.

Значне поширення отримав аналіз інформаційної дифузії – як поширюється інформація в межах соціальної мережі. Дослідження включає моделювання процесів репостів, цитування, лайків, коментарів, а також аналіз затримки в часі поширення повідомлень. У цьому контексті важливим є поняття каскаду – ланцюжка поширень одного повідомлення через різних користувачів. Структура каскадів може бути вузькою, якщо повідомлення поширюється лише в межах однієї спільноти або розгалуженою, коли воно «вірусно» переходить між групами. Моделі дифузії можуть бути як емпіричними (на основі історичних даних), так і симуляційними – з використанням SIR, SIS моделей, або моделі порогу. Розуміння того, як і чому певний контент стає вірусним, є стратегічно важливим для медіа, брендів та державних структур.

У дослідженні структур також актуальним є аналіз щільності мережі, який характеризує загальний рівень взаємопов'язаності. Чим вища щільність, тим більше зв'язків між учасниками. Щільні мережі мають високий рівень довіри, але можуть бути схильними до інформаційної замкнутості. Натомість розріджені мережі мають кращу здатність до розповсюдження нової інформації, але нижчий рівень соціального капіталу. Щільність залежить від типу платформи, цілей взаємодії та загальної активності користувачів. Аналіз щільності застосовується для оцінки стабільності мережі, виявлення вузьких місць, а також для планування втручань (наприклад, інформаційних кампаній).

Значну роль у структурі соціальних мереж відіграють мости – зв'язки між віддаленими спільнотами. Їхня роль полягає в перенесенні

інформації з одного сегмента мережі до іншого. Виявлення мостів має стратегічне значення: користувачі, які утворюють такі зв'язки, можуть не бути найактивнішими, але саме вони забезпечують циркуляцію новин між сегментами аудиторії. Мости важливі також у боротьбі з дезінформацією: через них можуть поширюватися як шкідливі, так і корисні повідомлення.

Інструментальне забезпечення для дослідження структур соціальних мереж доволі розвинене. Серед найпоширеніших платформ – Gephi, що дозволяє візуалізувати великі графи з численними фільтрами та аналітичними функціями. NodeXL інтегрується з Excel і підходить для початкового аналізу невеликих мереж. NetworkX – бібліотека Python для обчислення метрик та побудови графів. Для великих мереж використовуються Graph-tool, Pajek, Cytoscape. Із сучасних хмарних платформ популярні Neo4j для зберігання та запитів до графових баз даних, а також Cloud Analytics платформи зі спеціалізованими модулями аналізу соціальних мереж [11].

Дослідження структур соціальних мереж активно застосовується в прикладних сферах. У політиці – для виявлення координації бот-мереж або аналізу передвиборчих кампаній. У маркетингу – для вивчення споживацьких спільнот та побудови інфлюенс-кампаній. У сфері охорони здоров'я – для відстеження поширення інформації про вакцинацію або епідемії. У безпековій сфері – для виявлення радикальних угруповань чи каналів вербування. Такі дослідження мають важливе значення для розробки ефективних стратегій комунікації, боротьби з фейковими новинами та підвищення інформаційної стійкості суспільства.

Таким чином, аналіз структури соціальних мереж є складним, але надзвичайно важливим інструментом для розуміння цифрової взаємодії. Він дає змогу не лише описати поточний стан мережі, а й передбачити її еволюцію, виявити ключові вузли, оцінити ефективність поширення інформації та підготувати ґрунт для впливу на соціальні процеси. У поєднанні з машинним навчанням та інструментами візуалізації такі дослідження дозволяють реалізовувати аналітику нового покоління, здатну працювати з великими обсягами динамічних даних у режимі реального часу.

Висновки. З огляду на динаміку розвитку цифрових технологій, дослідження соціальних медіа залишатиметься актуальним у найближчі десятиліття. Основними перспективними напрямками вважаються: автоматизований аналіз великих даних (big data), застосування алгоритмів машинного навчання, вивчення візуального контенту (меми, фото, відео), дослідження цифрової ідентичності та онлайн-активізму. Усе більшої ваги набувають міждисциплінарні підходи, що поєднують соціогуманітарні та технічні науки, а саме соціологію, лінгвістику, психологію, інформатику, право.

Узагальнюючи, можна стверджувати, що машинне навчання відкриває широкі перспективи для аналізу текстових даних, забезпечуючи як

високу точність, такі адаптивність до нових завдань. Водночас, успішна реалізація таких систем вимагає не лише технічних знань, а й уважності до якості даних, контексту застосування, етичних стандартів. Інтеграція алгоритмів машинного навчання з сучасними інтерфейсами, автоматизацією та візуалізацією створює потужні інструменти, які можуть бути використані в освіті, науці, бізнесі та державному секторі, формуючи нову культуру обробки тексту в цифрову епоху.

Список використаної літератури

1. Белл Д. Прихід постіндустріального суспільства. *Сучасна зарубіжна соціальна філософія. Хрестоматія* / Упоряд. В. Лях. Київ, 1996. С. 194–250.
2. Веселяк В. В., Грицюк Ю. І. Методи машинного навчання в епідеміологічних дослідженнях. *Scientific Bulletin of UNFU*. 2024. №34. С. 59–67.
3. Дорошенко І.В., Книгніцька Т.В., Дереторська Т.І. Порівняння алгоритмів машинного навчання для прогнозування смертності від вірусу COVID-19. *VIRUS. SWorldJourn*. 2022. № 2, С. 72–77.
4. Кутуєв П.В. Трансформації модерну: інституції, ідеї, ідеології : монографія. Херсон : Видавничий дім «Гельветика», 2016. 516 с.
5. Лобовікова О. О., Мельніков А. С. Соціальні мережі як феномен інформаційного суспільства. *Вісник Львівського університету*. 2011. №. 5. С. 154–160.
6. Маклюєн Г. Галактика Гутенберга. Становлення людини друкованої книги / пер. з англ. А. Галушка, В. Самойленко. 2024. 384 с.
7. Соціальні медіа. *Вікіпедія: вільна енциклопедія*. URL: Режим доступу: https://uk.wikipedia.org/wiki/Соціальні_медіа (дата звернення: 21.04.2025)
8. Яковлев М.В., Дубас М.О., Купка О.В. Цифрова соціологія та постінформаційне суспільство: методологічні виклики й дослідницькі перспективи. *Науковий журнал «Габітус»*. 2023. № 55. С. 33–38.
9. Ardabili S. F., Mosavi A., Ghamisi P., Ferdinand F., Varkonyi-Koczy A. R., Reuter U., Rabczuk T., Atkinson P. M.. COVID-19 Outbreak Prediction with Machine Learning. *SSRN Electronic Journal*. 2020, №13.
10. Artificial Intelligence and COVID-19: Deep Learning Approaches for Diagnosis and Treatment / Jamshidi M. et al. *IEEE Access*. 2020. №8. P. 109581-109595. DOI: 10.1109/ACCESS.2020.3001973
11. Essien A, Petrounias I, Sampaio P, Sampaio S. Improving Urban Traffic Speed Prediction Using Data Source Fusion and Deep Learning. *IEEE International Conference on Big Data and Intelligent Computing (BigComp)*. IEEE. 2019. P. 1–8. DOI:10.1109/BIGCOMP.2019.8679231
12. Fatima M., Pasha M. Survey of Machine Learning Algorithms for Disease Diagnostic. *Journal of Intelligent Learning Systems and Applications*, 2017. № 9. P. 1–16.
13. Mitchell T. Machine Learning. McGraw Hill. 1997.
14. Quinlan J. R. C4.5: Programs for Machine Learning. Mach L.J. Morgan Kaufmann Publishers, San Mateo, California. 1993. 299 p.
15. Sarker I.H, Alkahtani H., Alsolami F., Khan A., Abusharq Y., Siddiqui M.K. Context pre-modeling: an empirical analysis for classification based user-centric context-aware predictive modeling. *Journal Big Data*. 2020. №7. P. 1–23.

16. Weka: Practical machine learning tools and techniques with Java implementations / Witten I. H. et al. New Zealand. Department of Computer Science, The University of Waikato. 1993.

17. Witten I. H., Frank E. Data Mining: Practical Machine Learning Tools and Techniques. San Francisco: Morgan Kaufmann Publishers. 2nd edition. 2005. 560 p.

Pygolenko I. V., Gasanov V. A. Using machine learning to analyze social media

The article is devoted to the consideration of the use of machine learning for the analysis of modern social media, which makes it possible to cluster discussion topics, detect fake content and visualize relationships between users. Social networks, as digital structures of user interaction, provide unique opportunities for the study of social processes, information exchange, community formation, as well as the analysis of influence and connections between participants.

The article substantiates the relevance and necessity of using machine learning algorithms in the field of social media analysis, since it is these algorithms that allow to detect hidden connections, topics, patterns of behavior and fake content. The importance of improving social media analysis systems based on machine learning is substantiated both for scientific research and for practical application in the fields of security, journalism, marketing and civil society.

In the languages of development of digital society, there is a growing need for intelligent automated systems capable of effectively analyzing large amounts of data in real time, which makes it possible to improve productivity and responsiveness in critical situations.

The article analyzes the use of machine learning algorithms for text data analysis, in particular its subfields such as natural language processing (NLP) and deep learning, which provide a wide range of tools for implementing these tasks.

The article draws attention to the need to modernize classical sociological methods of working with information in connection with the constant growth of huge arrays of unstructured data.

Despite a significant amount of scientific research in the field of new social media, this topic remains relevant due to the constant development of communication tools, social media and social networks, which affects the transformation of social space in modern digital society.

Key words: social media, social networks, communication, machine learning, data analysis, research.