

ОСОБЛИВОСТІ ЖАНРОВОЇ КЛАСИФІКАЦІЇ ЛІТЕРАТУРИ ЗА ДОПОМОГОЮ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ

Володимир Пасічник

*доктор технічних наук, професор,
Національний університет “Львівська політехніка”
ORCID ID 0000-0002-5231-6395
vpasichnyk@gmail.com*

Максим Яромич

*аспірант,
Національний університет “Львівська політехніка”
ORCID ID 0009-0005-3299-6695
yamax0312@gmail.com*

Ключові слова: жанрова класифікація, великі мовні моделі, параметризація жанрів, автоматизований аналіз тексту, жанрові онтології.

У статті розглянуто історичний контекст жанрової класифікації, параметризацію жанрів та їхню роль у структуризації літературних творів. Досліджено досвід застосування автоматизованої параметризації у сферах фільмів (Netflix) та музики (Spotify, Apple Music), що дозволяє персоналізувати користувацький досвід. Аналогічні методи використовуються в літературній сфері, зокрема в Amazon, де параметризація жанрів допомагає у продажах та рекомендаціях книг. У науковій літературі жанрова класифікація також відіграє ключову роль у систематизації знань, сприяючи ефективнішому пошуку та аналізу наукових робіт.

Одним із основних питань статті є роль великих мовних моделей у жанровій класифікації. Розглянуто, як ці моделі можуть ідентифікувати жанрові особливості текстів, виділяючи стилістичні прийоми та враховуючи лексичні й синтаксичні параметри. Окрему увагу приділено методам оцінки точності автоматизованої класифікації, а також аналізу матриць невідповідностей та оцінки впевненості моделей. Підкреслено, що інтеграція великих мовних моделей із жанровими онтологіями та вдосконалення адаптивних алгоритмів можуть значно покращити якість автоматичної класифікації.

У статті також порівнюються традиційні та автоматизовані підходи до жанрової класифікації. Традиційні методи базуються на експертному аналізі та враховують культурний і історичний контекст жанрів, що забезпечує глибину аналізу, але обмежує масштабованість. Автоматичні методи, навпаки, дозволяють швидко обробляти великі масиви текстів, але можуть нехтувати нюансами та гібридністю жанрів. Оптимальним підходом є поєднання цих методів, що забезпечує баланс між точністю та пояснюваністю.

Серед основних викликів жанрової класифікації великих мовних моделей виділено розмитість жанрових меж, вплив навчальних даних на упередження моделей та складність розпізнавання складних стилістичних прийомів. Запропоновані напрями вдосконалення включають розширення корпусів навчальних даних, інтеграцію моделей із літературознавчими базами знань, використання багаторівневої класифікації та покращення пояснюваності результатів.

Таким чином, стаття робить важливий внесок у дослідження автоматизованої жанрової класифікації, демонструючи переваги та обмеження LLM у цьому завданні.

PECULIARITIES OF LITERARY GENRE CLASSIFICATION USING LARGE LANGUAGE MODELS

Volodymyr Pasichnyk

*Doctor of Technical Sciences, Professor,
Lviv Polytechnic National University*

Maksym Yaromych

*Posgraduate Student,
Lviv Polytechnic National University*

Key words: *genre classification, large language models, genre parameterization, automated text analysis, genre ontologies.*

The article examines the historical context of genre classification, the parameterization of genres, and their role in structuring literary works. The experience of applying automated parameterization in the fields of films (Netflix) and music (Spotify, Apple Music) is analyzed, demonstrating how it enables personalized user experiences. Similar methods are used in the literary sector, particularly on Amazon, where genre parameterization aids in book sales and recommendations. In scientific literature, genre classification also plays a key role in knowledge systematization, contributing to more efficient search and analysis of academic works.

One of the central issues addressed in the article is the role of large language models in genre classification. The study explores how these models can identify genre-specific features of texts by highlighting stylistic techniques and considering lexical and syntactic parameters. Particular attention is given to methods for assessing the accuracy of automated classification, as well as analyzing confusion matrices and evaluating model confidence. It was emphasized that integrating large language models with genre ontologies and refining adaptive algorithms could significantly enhance the quality of automated classification.

The article also compares traditional and automated approaches to genre classification. Traditional methods rely on expert analysis and take into account the cultural and historical context of genres, providing deeper analytical insights but limiting scalability. Automated methods, on the other hand, allow for the rapid processing of large volumes of text but may overlook nuances and genre hybridity. The optimal approach is a combination of these methods, ensuring a balance between accuracy and interpretability.

Among the key challenges in genre classification using large language models, the study highlights the ambiguity of genre boundaries, the impact of training data biases on model performance, and the difficulty of recognizing complex stylistic techniques. Suggested improvements include expanding training data corpora, integrating models with literary knowledge bases, implementing multi-level classification strategies, and enhancing the explainability of classification results.

Thus, the article makes a significant contribution to research on automated genre classification, demonstrating both the advantages and limitations of large language models in this task.

Вступ. Жанрова класифікація літератури є однією з основних проблем як у традиційному літературознавстві, так і в сучасних дослідженнях у сфері штучного інтелекту та автоматизованої обробки текстів. Протягом століть жанр слугував важливим аналітичним інструментом, який допомагав дослідникам систематизувати літературні твори, визначати їхні особливості та

аналізувати еволюцію літературного процесу. Водночас, жанри є динамічними категоріями, що змінюються з розвитком суспільства, культури та технологій, що значно ускладнює їх формальну класифікацію.

З появою великих обсягів цифрової інформації та розширенням електронних бібліотек зростає потреба в автоматизованих підходах

до класифікації текстів за жанрами. Традиційні методи, що базуються на експертному аналізі, є малоефективними в умовах сучасного інформаційного середовища, оскільки вимагають значних витрат часу та людських ресурсів. Крім того, межі між жанрами стають дедалі розмитішими, що ускладнює процес їхнього визначення навіть для досвідчених літературознавців.

Сучасні методи штучного інтелекту, зокрема великі мовні моделі (LLM), пропонують нові можливості для автоматизованого визначення жанру тексту. Вони здатні аналізувати великі корпуси літературних творів, визначати патерни та особливості різних жанрів на основі статистичних закономірностей, синтаксичних структур і стилістичних характеристик. Проте, незважаючи на успіхи у використанні LLM для обробки природної мови, проблема жанрової класифікації залишається відкритою через низку викликів.

По-перше, великі мовні моделі можуть демонструвати упередження, оскільки вони навчаються на вже існуючих даних, що можуть бути нерепрезентативними або несистематизованими з точки зору жанрової диференціації. Це призводить до ситуацій, коли модель може некоректно визначати жанр твору або застосовувати жанрові категорії, що відображають домінуючі культурні традиції, ігноруючи менш популярні або експериментальні форми.

По-друге, великі мовні моделі мають труднощі з розрізненням гібридних жанрів або текстів, що поєднують риси кількох жанрових напрямів. Сучасна література все частіше відходить від традиційних жанрових схем, створюючи твори, що містять елементи, наприклад, наукової фантастики, детективу, фентезі та постмодерністської прози одночасно. Це створює серйозну проблему для автоматичних класифікаторів, оскільки вони часто намагаються віднести текст лише до одного жанру, ігноруючи його складну жанрову структуру.

По-третє, пояснюваність моделей жанрової класифікації залишається важливим питанням для дослідників. Великі мовні моделі працюють як «чорні скриньки», тобто вони можуть давати точні результати, проте процес прийняття рішень залишається незрозумілим. Це ускладнює інтерпретацію результатів, особливо для гуманітарних досліджень, де обґрунтованість класифікації має не менш важливе значення, ніж сама категоризація.

Ще однією важливою проблемою є відсутність універсальної системи параметризації жанрів, що могла б бути використана для навчання мовних моделей. Більшість сучасних підходів до жанрової класифікації базуються на аналізі окремих характеристик, таких як частотність певних слів, довжина речень або стиль викладу. Проте жанр

є багатовимірною категорією, що включає тематичні, структурні, стилістичні та функціональні аспекти. Без належної параметризації класифікація жанрів залишається фрагментарною і нестабільною.

З огляду на ці виклики, дослідження жанрової класифікації за допомогою LLM є актуальним та необхідним. Розробка нових методів, що дозволять підвищити точність та гнучкість автоматичної класифікації, сприятиме покращенню роботи цифрових бібліотек, літературних архівів, освітніх ресурсів і комерційних платформ з продажу книг. Більш досконала система класифікації жанрів також може бути корисною для письменників, редакторів і видавців, допомагаючи їм краще орієнтуватися в сучасному літературному просторі.

Таким чином, проблема жанрової класифікації літератури залишається відкритою як у літературознавстві, так і в комп'ютерній лінгвістиці. Незважаючи на значні досягнення у сфері штучного інтелекту, сучасні алгоритми все ще не можуть повністю відтворити складність жанрових категорій. Це робить подальші дослідження у цьому напрямку важливими та перспективними, оскільки вони можуть не лише вдосконалити технології автоматизованого аналізу тексту, але й сприяти кращому розумінню самої природи жанрів та їхньої ролі у літературному процесі.

Аналіз попередніх досліджень та публікацій. Упродовж останніх десятиліть автоматизація літературних досліджень зазнала значного розвитку завдяки застосуванню методів штучного інтелекту, зокрема великих мовних моделей. Від традиційного літературного аналізу, який базувався на експертному підході, дослідники поступово переходять до автоматизованих методів, що дозволяють обробляти великі корпуси текстів із високою точністю та швидкістю.

З появою методів машинного навчання та нейронних мереж автоматизована класифікація текстів за жанрами стала об'єктом численних досліджень. У дослідженні (Sobchuk, Sela, 2023) розглядається проблема класифікації літературних текстів за жанрами з використанням різних алгоритмів обробки тексту. Автори аналізують існуючі методи кластеризації та пропонують власний підхід, що поєднує різні етапи аналізу, такі як попередня обробка тексту, виділення текстових ознак та вимірювання відстаней між ними, для покращення точності класифікації.

Одним із ключових напрямів сучасних досліджень є розробка алгоритмів, здатних точно визначати жанр літературного твору на основі його лексичних, стилістичних та структурних характеристик. У роботі (Lepekhn, Sharoff, 2022) досліджується проблема оцінки впевненості прогнозів окремих класифікаторів та їх ансамблів у завданні

жанрової класифікації. Автори підкреслюють, що навіть при використанні ансамблевих методів результати можуть бути нестабільними, особливо при класифікації гібридних жанрів. Це вказує на необхідність подальшого вдосконалення моделей та врахування специфіки різних жанрів.

Ще одним важливим аспектом є розробка жанрових онтологій, які дозволяють систематизувати та формалізувати знання про жанри. У дослідженні (Schreiber, 2016) порівнюється ефективність двох підходів до створення жанрових онтологій: одного, що базується на ретельно відібраних і систематизованих даних, упорядкованих фахівцями, і іншого, який формується на основі позначень, доданих великою кількістю користувачів. Автор демонструє, що використання онтологій, створених на основі колективних даних, може бути ефективним для покращення жанрової класифікації, особливо в умовах швидко змінюваних жанрових категорій. Робота (Garbacz, 2006) демонструє формальну онтологію жанрів, яка пропонує ієрархічну систему категорій для представлення жанрової інформації.

Незважаючи на досягнення, залишається низка відкритих питань. Одне з них стосується пояснюваності результатів, отриманих за допомогою LLM. У роботі (Martin, Hood, 2024) розглянуто використання методів обробки природної мови (NLP) у літературних оглядах, наголошуючи на важливості розробки моделей, які не лише забезпечують високу точність класифікації, але й надають зрозумілі пояснення своїх рішень. Це особливо актуально для дослідників, які прагнуть інтерпретувати та довіряти результатам автоматизованого аналізу.

Отже, стан досліджень у галузі жанрової класифікації літератури за допомогою великих мовних моделей є досить активним. Дослідники приділяють особливу увагу розширенню можливостей моделей у розпізнаванні складних жанрових форм, вдосконаленню стилістичного аналізу, розробці жанрових онтологій та підвищенню пояснюваності класифікаційних моделей. Водночас, залишаються відкритими питання щодо впливу навчальних корпусів на упередженість результатів, пояснюваності жанрових класифікацій та необхідності поєднання експертного аналізу з методами машинного навчання. Подальший розвиток у цій сфері спрямований на створення гнучких, адаптивних і більш точних систем класифікації, що здатні відображати багатогранність сучасного літературного процесу.

Мета статті: проаналізувати можливості та обмеження жанрової класифікації літературних текстів за допомогою великих мовних моделей.

Матеріал і методи дослідження. Матеріалом для дослідження є наукові джерела із

літературознавства, теорії жанру та суміжних міждисциплінарних досліджень, які висвітлюють сучасні підходи до параметризації та класифікації текстів. Для аналізу використано теоретико-методологічний підхід, спрямований на дослідження жанрової класифікації літературних творів.

Для реалізації поставлених завдань застосовано метод порівняльного аналізу, що дозволив зіставити традиційні літературознавчі підходи до визначення жанрової приналежності творів із сучасними автоматизованими методиками, які базуються на алгоритмах машинного та глибокого навчання. Теоретичну основу становлять праці таких авторів, як С. Барабаш, О. Черниш, Н. Гоца, О. Собчук, А. Шела, О. Абрамчук та інших, що дозволяє забезпечити глибоке й всебічне осмислення проблематики.

Дослідження передбачає параметризацію жанрів літературних творів, що включає визначення низки кількісних та якісних характеристик, зокрема лексичних, синтаксичних, стилістичних і структурних метрик. Лексичні метрики охоплюють аналіз частоти використання характерних слів та термінів для конкретних жанрів, тоді як синтаксичні метрики аналізують складність та особливості побудови речень у різних типах текстів. Стилiстичні параметри включають аналіз використання специфічних художніх засобів (метафор, епітетів, символів тощо), а структурні – аналіз композиційних особливостей тексту (розділи, акти, сцени, сюжетна організація).

Додатковим елементом методології стало вивчення досвіду цифрових платформ, таких як Netflix, Spotify та Amazon, які успішно застосовують автоматичну параметризацію та класифікацію контенту для рекомендаційних систем. Аналіз цих прикладів дозволив виявити перспективні напрями перенесення цифрових методик у сферу літературознавства та оптимізувати запропоновану методику дослідження.

Таким чином, комплексна теоретико-методологічна база, використання передових алгоритмів машинного навчання та глибокий порівняльний аналіз джерел дозволяють сформувати цілісну систему методів, що забезпечують релевантність і точність отриманих результатів дослідження.

Результати та обговорення. *Поняття жанру та історичний контекст жанрової класифікації.* Термін «жанр» походить від французького «genre», що означає «рід» або «вид». У літературознавстві жанр визначається як тип художнього твору, який характеризується певними сталими, повторюваними формальними та змістовими ознаками. Ці ознаки дозволяють систематизувати та класифікувати літературні твори, розрізняючи їх за формою, змістом, стилем та призначенням. Як зазначається в роботі (Барабаш, 2019), жанр

є не лише класифікаційною категорією, а й засобом формування художньої традиції. Сучасні дослідження з теорії жанру вказують на складність і багатозначність цього поняття (Черниш, 2020). Жанри можуть еволюціонувати під впливом соціальних і культурних змін, а також змінюватися відповідно до комунікативних функцій та медійних платформ.

Жанрова класифікація має глибоке історичне коріння, що сягає античних часів. Уже в працях Арістотеля, зокрема в «Поетиці», було запропоновано першу систематизацію літературних жанрів, де виділялися епос, лірика та драма. Ця тріада стала основою для подальших класифікацій і залишається важливою до сьогодні (Жанри і жанрові процеси в історико-літературній перспективі, 2012). У період Середньовіччя жанрова система зазнала трансформацій під впливом християнської культури, що призвело до появи нових жанрів, таких як житія святих та релігійні гімни.

Розвиток друкарства в епоху Відродження сприяв поширенню світських жанрів, зокрема роману та есе. Ця епоха принесла відновила інтерес до античних форм, але з новим акцентом на гуманізм та індивідуалізм, що сприяло розвитку таких жанрів, як сонет, есе та новела. У XVII столітті, з приходом класицизму, встановлюється чітка ієрархія жанрів. Жанри поділялися на «високі» (трагедія, епопея, ода), призначені для зображення піднесених тем, та «низькі» (комедія, сатира, басня), що стосувалися повсякденного життя. Така ієрархія базувалася на принципах раціоналізму та відповідності строгим канонам. У XVIII–XIX століттях, з розвитком реалізму та романтизму, жанрова система продовжувала еволюціонувати, відображаючи зміни в суспільстві та культурі (Гоца, 2009).

У сучасному літературознавстві жанр розглядається як динамічна категорія, що постійно еволюціонує під впливом культурних, соціальних та історичних чинників. Дослідники підкреслюють, що жанр не є статичною формою, а радше процесом, який відображає зміну літературних практик та читачьких очікувань, постійно адаптуючись до нових історичних та культурних контекстів.

Параметризація жанрів. Вимірювання параметрів. Жанрова класифікація відіграє ключову роль у літературознавстві, оскільки дозволяє впорядкувати величезний масив літературних творів, роблячи їх аналіз більш структурованим і доступним. Визначення жанру допомагає не тільки в академічному середовищі, а й у книжковій індустрії, де правильне позиціонування книги впливає на її комерційний успіх. З розвитком цифрових технологій та штучного інтелекту параметризація жанрів стала важливою для алгоритмічних систем, які здійснюють автоматизований аналіз текстів.

Наприклад, платформи, що надають рекомендації читачам, як-от Amazon Kindle чи Google Books, значною мірою покладаються на автоматичну класифікацію жанрів для точного підбору творів відповідно до вподобань користувачів.

Параметризація жанрів передбачає визначення та опис характеристик, які дозволяють класифікувати літературні твори за жанровою приналежністю. Цей процес включає ідентифікацію формальних, змістових та стилістичних ознак, що є спільними для творів певного жанру. Варто враховувати, що жанр є не лише класифікаційною категорією, а й засобом формування художньої традиції.

Сучасні дослідження підтверджують, що жанри можуть еволюціонувати під впливом соціальних і культурних змін, а також змінюватися відповідно до комунікативних функцій та медійних платформ (Черниш, 2020). Наприклад, у сучасній художній літературі з'являються змішані жанри, такі як науково-фантастичний детектив або роман із елементами нон-фікшн, що ставить перед літературознавцями нові виклики в їх параметризації.

Для точного визначення жанру літературного твору використовують різні метрики. Лексичні метрики дозволяють аналізувати частоту використання певних слів або фраз, характерних для конкретного жанру. Наприклад, у науковій фантастиці часто зустрічаються терміни, пов'язані з технологіями та майбутнім, тоді як у детективній літературі домінують слова, що відображають розслідування, загадковість і кримінальні аспекти. Синтаксичні метрики допомагають оцінити складність речень, використання певних граматичних конструкцій, що особливо важливо для аналізу наукових і публіцистичних текстів. Дослідження показують, що наукові статті часто містять складні підрядні речення та термінологічно насичені конструкції, тоді як художня література допускає більшу варіативність у структурі речень (Sobchuk, Sela, 2023).

Окрім лексичних і синтаксичних особливостей, важливу роль відіграють стилістичні та структурні метрики. Аналіз художніх засобів, таких як метафори, епітети, символи, допомагає виокремити жанрові особливості тексту. Наприклад, лірична поезія відзначається високою концентрацією метафоричних конструкцій і емоційно насиченою лексикою, тоді як публіцистика тяжіє до раціональності та логічності викладу. Крім того, у різних жанрах спостерігається специфічна структура тексту: наприклад, драматичні твори чітко поділяються на акти та сцени, тоді як класичний роман містить розділи з послідовним розвитком сюжету (Бехта, Марчук, 2021).

Таким чином, параметризація жанрів є важливим інструментом для глибшого розуміння та

класифікації літературних творів. Вона дозволяє не лише впорядковувати літературний матеріал, а й автоматизувати процеси його аналізу, що має значний вплив як на літературознавство, так і на комерційну сферу книговидавництва. Сучасні дослідження підкреслюють, що жанр є динамічною категорією, яка постійно адаптується до нових історичних та культурних контекстів (Жанри і жанрові процеси в історико-літературній перспективі, 2012).

Досвід параметризації фільмів та музики. У сучасному цифровому світі параметризація контенту відіграє ключову роль у наданні користувачам персоналізованих рекомендацій та покращенні їхнього досвіду взаємодії з платформами. Особливо яскраво це проявляється у сферах стрімінгових сервісів, таких як Netflix для фільмів та серіалів, а також музичних платформ, які активно використовують параметризацію для класифікації та рекомендації контенту.

Netflix є одним із провідних прикладів успішного застосування параметризації для організації та рекомендації відеоконтенту. Платформа використовує складну систему класифікації, яка враховує різноманітні характеристики фільмів та серіалів, такі як жанр, акторський склад, рік випуску та інші метадані. Ця інформація дозволяє Netflix створювати детальні категорії та підкатегорії, що допомагають користувачам знаходити контент відповідно до їхніх уподобань. Крім того, система рекомендацій Netflix аналізує історію переглядів користувача, його оцінки та взаємодію з платформою, щоб пропонувати персоналізовані рекомендації. Цей підхід забезпечує більш точне та релевантне відображення контенту для кожного користувача, підвищуючи їхню задоволеність та залученість (Netflix Help Center, 2025).

Музичні стрімінгові платформи, такі як Spotify, Apple Music та інші, також активно використовують параметризацію для організації та рекомендації музичного контенту. Вони аналізують різні атрибути треків, включаючи жанр, темп, тональність, інструментальне оформлення та навіть емоційне забарвлення композицій. Ці дані, разом із поведінковими показниками користувачів, такими як історія прослуховувань, створені плейлисти та вподобання, дозволяють платформам створювати персоналізовані рекомендації та автоматично генерувати плейлисти, що відповідають настрою або активності користувача. Такий підхід не лише покращує користувацький досвід, але й сприяє відкриттю нових виконавців та жанрів, розширюючи музичні горизонти слухачів. У статті (Biswas, Dhabal, Venkateswaran, 2023) автори досліджують методи класифікації музичних жанрів, поєднуючи цифрову обробку сигналів та глибоке навчання. Вони пропонують новий алгоритм, який

використовує ці методи для виділення релевантних особливостей з аудіосигналів та класифікації їх за жанрами. Цей підхід може бути інтегрований у музичні додатки для покращення рекомендацій та організації контенту.

Успішний досвід параметризації фільмів та музики може бути адаптований і до сфери літератури. Онлайн-книгарні та електронні бібліотеки можуть впровадити системи класифікації, що враховують різноманітні параметри літературних творів, такі як жанр, стиль, тематика, історичний період, географічне розташування подій та інші метадані. Аналіз читацьких вподобань, історії переглядів та оцінок дозволить створювати персоналізовані рекомендації, допомагаючи читачам знаходити книги, що відповідають їхнім інтересам. Крім того, така параметризація може сприяти авторам та видавцям у кращому розумінні потреб аудиторії та адаптації контенту відповідно до її запитів.

Застосування параметризації у літературі може також включати аналіз стилістичних та структурних особливостей текстів, що дозволить більш точно визначати жанрову приналежність творів та їхню відповідність вподобанням конкретних читачів. Це, у свою чергу, сприятиме підвищенню задоволеності користувачів та стимулюватиме зростання продажів у книжковій індустрії.

Таким чином, досвід параметризації фільмів та музики демонструє ефективність такого підходу у покращенні користувацького досвіду та оптимізації процесу рекомендацій. Адаптація цих методів до літературної сфери має потенціал значно покращити взаємодію читачів з контентом та сприяти розвитку книжкового ринку в цілому.

Спроба параметризації за жанрами літератури в Amazon. Одним із найбільш показових прикладів такої адаптації є Amazon, який активно використовує параметризацію для впорядкування свого книжкового каталогу, покращення пошуку та підвищення рівня персоналізації рекомендацій.

Amazon, як одна з найбільших онлайн-платформ з продажу книг, приділяє значну увагу класифікації та параметризації літературних творів за жанрами. Цей підхід дозволяє не лише впорядкувати величезний асортимент, але й покращити користувацький досвід, забезпечуючи точні та релевантні рекомендації для читачів. Для ефективної параметризації книг Amazon використовує комбінацію метаданих та алгоритмічного аналізу, що дозволяє забезпечити найбільш точну класифікацію літературних творів і водночас підвищити їхню видимість у пошуковій системі платформи.

Метадані книги, такі як її жанр, піджанр, ключові слова та опис, відіграють центральну роль у процесі класифікації. Автори та видавці під час публікації можуть вказати основний жанр

книги, а також обрати додаткові категорії, що найточніше відображають її зміст. Цей процес допомагає Amazon позиціонувати книгу в каталозі та забезпечує її релевантність у рекомендаційних алгоритмах. Додатково важливими є ключові слова – система дозволяє авторам додавати до семи ключових слів або фраз, що описують сюжетні або стилістичні особливості книги. Ці ключові слова суттєво впливають на видимість книги в пошукових запитах та алгоритмах рекомендацій (Самвидав на Amazon, 2025).

Опис та анотація книги також впливають на її класифікацію. Детальний опис, який містить інформацію про сюжет, тематику та особливості стилю, допомагає алгоритмам краще розуміти зміст книги і відповідно коректніше визначати її жанрову приналежність. Крім текстових характеристик, на алгоритмічний аналіз впливає також історія продажів книги, рівень її популярності серед читачів та середній рейтинг, що формується на основі відгуків користувачів. Чим більше продажів і позитивних оцінок має книга в певному жанрі, тим частіше вона з'являється у відповідних підбірках та пропозиціях для інших користувачів.

Точна параметризація та класифікація книг безпосередньо впливає на їхню видимість на платформі, а отже, і на обсяги продажів. Книги, які правильно віднесені до відповідних жанрів та мають релевантні ключові слова, частіше з'являються в результатах пошуку та рекомендаціях, що збільшує ймовірність їхнього придбання. Крім того, участь у програмах на кшталт «Вибір редакції Amazon» може суттєво підвищити популярність книги та стимулювати продажі (Вибір редакції Amazon, 2025).

Наразі вже існують роботи, що аналізують загальні підходи до класифікації та рекомендацій у цифрових книжкових платформах. Наприклад, дослідження (Лозинська, 2018) розглядає фактори, що впливають на популярність книг та їхні продажі, включаючи роль класифікації та рекомендаційних систем.

Таким чином, параметризація літературних творів за жанрами на Amazon є критично важливою для оптимізації пошуку, покращення рекомендацій та збільшення продажів. Використання точних метрик та алгоритмів дозволяє платформі ефективно поєднувати книги з їхньою цільовою аудиторією, забезпечуючи задоволення потреб як читачів, так і авторів.

Жанрова класифікація у науковій літературі. Наукова література охоплює широкий спектр жанрів, кожен з яких має свої особливості та призначення. Розглянемо це на прикладі галузі штучного інтелекту. Основними жанрами є підручники, посібники, наукові статті, огляди, технічні звіти та конференційні матеріали. Кожен із цих жанрів

виконує специфічну функцію в науковій комунікації та сприяє поширенню знань у спільноті дослідників.

Підручники та посібники зазвичай призначені для освітніх цілей, надаючи систематизовану інформацію про основи та сучасні досягнення в галузі ІІІ. Вони слугують базою для студентів та фахівців, які прагнуть глибше зрозуміти предмет. Наукові статті та огляди, навпаки, фокусуються на представленні нових досліджень, методів та результатів, адресованих переважно науковій спільноті. Технічні звіти та конференційні матеріали часто містять попередні результати досліджень або детальний опис нових технологій та алгоритмів, що ще не були опубліковані в рецензованих журналах.

Параметризація, тобто визначення та використання специфічних характеристик для опису та класифікації об'єктів, відіграє важливу роль у систематизації наукової літератури. У контексті жанрової класифікації це означає ідентифікацію ключових параметрів, таких як структура документа, стиль викладу, цільова аудиторія, обсяг та глибина представленого матеріалу. Наприклад, наукові статті часто мають стандартну структуру IMRaD (Introduction, Methods, Results, and Discussion), тоді як підручники можуть бути організовані у вигляді розділів з теоретичним матеріалом та практичними завданнями.

Використання таких параметрів дозволяє автоматизованим системам точно визначати жанр конкретного документа, що, у свою чергу, сприяє ефективнішій організації та пошуку інформації. Це особливо важливо в умовах постійного зростання обсягу наукових публікацій, де ручна класифікація стає все менш практичною.

Сучасні технології ІІІ пропонують інструменти для автоматичної параметризації та класифікації наукових текстів. Алгоритми машинного навчання можуть аналізувати текстові особливості документів, такі як частота вживання певних термінів, складність синтаксичних конструкцій та загальна структура тексту, для визначення його жанру. Дослідження (Sobchuk, Sela, 2023) демонструє, як різні алгоритми можуть бути використані для кластеризації літературних жанрів на основі тематичної схожості текстів. Запропоновані методи можуть бути адаптовані для наукових текстів, зокрема в галузі ІІІ. У роботі (Гулик, 2023) продемонстровано, як ІІІ може бути застосований для аналізу лінгвістичних особливостей текстів, що сприяє їх класифікації та розумінню структурних характеристик.

Крім того, використання ІІІ може допомогти виявити нові теми та перспективні напрями для досліджень, аналізуючи великі обсяги наукових публікацій та визначаючи тенденції у розвитку

галузі. Це може бути корисним для дослідників, які прагнуть ідентифікувати прогалини в знаннях або нові області для вивчення.

Таким чином, параметризація та використання технологій штучного інтелекту у жанровій класифікації наукової літератури сприяють більш ефективному впорядкуванню знань, полегшуючи доступ до релевантної інформації та підтримуючи розвиток наукових досліджень у галузі штучного інтелекту.

Роль великих мовних моделей у сучасному аналізі тексту та автоматизація в гуманітарних науках. Невід’ємною частиною сучасних досліджень у галузі обробки природної мови та штучного інтелекту є великі мовні моделі. Їх здатність генерувати, аналізувати та інтерпретувати текст відкриває нові можливості для автоматизації та підвищення ефективності в гуманітарних науках.

Великі мовні моделі, такі як GPT-4, здатні розуміти контекст, структуру та семантику тексту, що дозволяє їм виконувати складні завдання аналізу. Вони можуть автоматично класифікувати тексти за жанрами, темами або емоційним забарвленням, що значно спрощує роботу дослідників. Наприклад, у статті (Mu, Dong, Bontcheva, 2024) дослідники демонструють, як LLM можуть замінити традиційні методи тематичного моделювання, забезпечуючи більш точні та релевантні результати. Крім того, вони здатні виявляти приховані зв’язки та патерни в текстах, що раніше вимагало значних людських ресурсів. Це відкриває нові перспективи для досліджень у галузі лінгвістики, літературознавства та історії.

Інтеграція великих мовних моделей у гуманітарні науки сприяє автоматизації рутинних та трудомістких завдань. Так, у статті (Karjus, 2024) розглядається, як LLM можуть допомогти в аналізі інтерв’ю, виявленні тематичних змін у літературі та аналізі історичних текстів. Це дозволяє дослідникам зосередитися на інтерпретації та критичному аналізі, залишаючи технічні аспекти моделі.

Також великі мовні моделі можуть бути використані для автоматичного створення анотацій, резюме та перекладів текстів, що підвищує доступність інформації та сприяє міждисциплінарним дослідженням.

Незважаючи на значні переваги, використання LLM у гуманітарних науках супроводжується певними викликами. Одним із них є необхідність забезпечення прозорості та інтерпретованості моделей, щоб дослідники могли розуміти та довіряти отриманим результатам. Крім того, важливо враховувати етичні аспекти, такі як упередженість моделей та конфіденційність даних. Очікується, що розвиток великих мовних моделей сприятиме глибшій інтеграції штучного інтелекту в гуманітарні науки, відкриваючи нові горизонти для досліджень та міждисциплінарної співпраці.

Параметризація та визначення жанрів за допомогою великих мовних моделей. Зручно використовувати великі мовні моделі і для параметризації та автоматичного визначення жанрів тексту. Вони здатні аналізувати складні мовні структури та контексти, що дозволяє ефективно класифікувати тексти за жанровими ознаками, що, у свою чергу, є важливим для організації великих обсягів інформації та покращення доступу до неї.

Параметризація тексту передбачає визначення та кількісну оцінку характеристик, які описують його зміст, стиль та структуру. Великі мовні моделі здатні автоматично витягувати такі параметри, аналізуючи лексичні, синтаксичні та семантичні особливості тексту. Це включає виявлення ключових тем, стилістичних прийомів, емоційного забарвлення та інших характеристик, що визначають жанр. Наприклад, дослідження (Doulaty, Saz, Raymond, 2016) демонструє, як моделі можуть використовувати акустичні та текстові особливості для класифікації медіа-контенту за жанрами.

Варто зазначити, що автоматичне визначення жанру тексту є складним завданням, оскільки жанри часто не мають чітких меж і можуть перекриватися. Великі мовні моделі навчаються на великих корпусах текстів різних жанрів, що дозволяє їм розпізнавати тонкі відмінності між ними. У дослідженні (Lepekhn, Sharoff, 2022) розглянуто використання ансамблів моделей для підвищення точності класифікації жанрів та оцінки впевненості в прогнозах.

Використання LLM для параметризації та визначення жанрів має широке практичне застосування. Це включає автоматичну індексацію бібліотечних фондів, рекомендаційні системи в медіа-сервісах, аналіз літературних творів та багато іншого. З розвитком технологій очікується, що великі мовні моделі стануть ще більш точними та ефективними в розпізнаванні та класифікації жанрів, що сприятиме покращенню доступу до інформації та її організації.

Таким чином, великі мовні моделі є потужним інструментом для параметризації та автоматичного визначення жанрів тексту, що відкриває нові можливості для дослідників та практиків у галузі обробки природної мови та інформаційних технологій.

Виділення стилістичних прийомів за допомогою великих мовних моделей. Великі мовні моделі значно вплинули на автоматизацію аналізу тексту, зокрема у виявленні та класифікації стилістичних прийомів. Їх здатність обробляти великі обсяги даних та розуміти контекст дозволяє ефективно ідентифікувати різноманітні стилістичні елементи, що є важливим кроком у визначенні жанру тексту.

Великі мовні моделі, такі як GPT-4 або BERT, навчаються на великих корпусах текстів, що дозволяє їм розпізнавати складні мовні структури та патерни. Для виділення стилістичних прийомів моделі аналізують різні рівні мовних особливостей. На лексичному рівні вони визначають частоту та вибір слів, специфічні терміни або діалектизми. На синтаксичному рівні досліджують структуру речень, наприклад, використання інверсій, повторів чи інших синтаксичних конструкцій. Семантичний рівень дозволяє виявляти метафори, іронію, сарказм та інші тропи. Для досягнення цього великі мовні моделі використовують механізми самоуваги (self-attention), які дозволяють моделі визначати важливість різних частин тексту в контексті всього документа. Це забезпечує глибоке розуміння тексту та його стилістичних особливостей.

Стилістичні прийоми є ключовими маркерами, що характеризують певний жанр. Наукові статті, наприклад, часто містять формальний стиль та специфічну термінологію, тоді як художня література може включати багатий набір метафор та емоційно забарвлених виразів. Автоматичне виявлення цих прийомів за допомогою великих мовних моделей дозволяє точно класифікувати тексти за жанрами. Дослідження (Doulaty, Saz, Raymond, 2016) демонструє, як аналіз стилістичних особливостей може бути використаний для ідентифікації жанру медіа-контенту.

Наукові дослідження підтверджують ефективність таких підходів. У роботі (Ворочек, Соловей, 2024) розглядається застосування великих мовних моделей для створення текстів, що відповідають певному стилю або жанру, підкреслюючи їх здатність адаптуватися до різних стилістичних вимог. Аналіз різних стилістичних прийомів та їх ролі у формуванні жанрової специфіки текстів також бути корисним для налаштування LLM на розпізнавання таких елементів.

Таким чином, великі мовні моделі є потужним інструментом для автоматичного виділення стилістичних прийомів, що безпосередньо впливає на точність визначення жанру тексту. Їх здатність аналізувати різні рівні мовних особливостей забезпечує глибоке розуміння стилістичних характеристик, що є ключовим для жанрової класифікації.

Порівняння традиційних і автоматичних підходів. У літературознавстві жанрова класифікація є ключовим інструментом для систематизації та аналізу літературних творів. Традиційні підходи до визначення жанру базуються на історично сформованих категоріях, що враховують змістові, структурні та стилістичні особливості творів. Традиційна класифікація жанрів часто спирається на аналіз тематичних елементів, композиційних

особливостей та авторського стилю. У дослідженні (Абрамчук, 2016) проаналізовано, як підвиди жанрів можуть визначатися за змістовою ознакою, наприклад історичний роман або роман жахів, а також за композиційними особливостями, такими як сонет або рондо. Такий підхід вимагає глибокого розуміння літературного контексту та еволюції жанрів, що дозволяє дослідникам точно класифікувати твори відповідно до їхніх характеристик.

З розвитком технологій та обробки великих обсягів даних з'являються автоматизовані підходи до жанрової класифікації. Ці методи використовують алгоритми машинного навчання та штучного інтелекту для аналізу текстів і визначення їхньої жанрової приналежності. Автоматична класифікація базується на кількісному аналізі лексичних, синтаксичних та стилістичних особливостей тексту, що дозволяє обробляти великі корпуси даних з високою швидкістю та точністю. У дослідженні (Sobchuk, Sela, 2023) наведено порівняння різних алгоритмів для автоматичного групування літературних творів за жанрами, демонструючи ефективність таких підходів у класифікації художньої літератури.

Як традиційні, так і автоматичні методи жанрової класифікації мають свої переваги та недоліки. Традиційні методи забезпечують глибокий та контекстуальний аналіз, враховуючи історичні та культурні аспекти розвитку жанрів. Вони дозволяють виявляти приховані жанрові переходи та інтертекстуальні зв'язки між творами, що складно реалізувати за допомогою автоматичних алгоритмів. Проте такі методи є трудомісткими, потребують великого часу та залежать від суб'єктивних інтерпретацій дослідників, що може впливати на узгодженість результатів.

Автоматичні підходи, навпаки, забезпечують швидкість та об'єктивність. У дослідженні (Aruk, Nuci, 2021) автори продемонстрували, що моделі глибокого навчання, такі як LSTM (Long Short-Term Memory), можуть досягати точності до 87,71% у класифікації педагогічних текстів, аналізуючи їхні мовні патерни та стилістичні особливості. Однак такі моделі часто не враховують нюанси, пов'язані з культурним контекстом або авторським задумом, і можуть бути менш ефективними при класифікації гібридних жанрів або експериментальної літератури.

У сучасних дослідженнях спостерігається тенденція до поєднання цих підходів. Використання автоматизованих інструментів для попередньої обробки та класифікації текстів, доповнене глибоким традиційним аналізом, дозволяє досягти більш точних та змістовних результатів у жанровій класифікації літературних творів. Це підтверджується у статті (Sobchuk, Sela, 2023), де

підкреслюється, що інтеграція автоматичних та традиційних методів дає змогу досягти найкращої точності класифікації, зберігаючи водночас контекстуальний аналіз. Таким чином, оптимальним шляхом є синергетичний підхід, що використовує переваги як машинного аналізу, так і експертного знання.

Потенційні проблеми та напрями для вдосконалення. Попри те, що жанрова класифікація літератури за допомогою великих мовних моделей відкриває широкі можливості для автоматизованого аналізу текстів та дозволяє значно пришвидшити процес категоризації та підвищити його масштабованість, цей підхід не позбавлений проблем, які обмежують його ефективність та точність. Важливим завданням стає не лише вдосконалення алгоритмів, але й адаптація класифікаційних методів до сучасних змін у літературному процесі.

Однією з головних проблем є розмитість жанрових меж та гібридність сучасних текстів. Літературні твори все частіше поєднують риси кількох жанрів, що ускладнює їхнє автоматичне розподілення в чітко окреслені категорії. Великі мовні моделі покладаються переважно на поверхневі лексичні та стилістичні маркери, що іноді призводить до хибних віднесення текстів до певного жанру навіть тоді, коли вони мають змішану жанрову природу. Контекстуальна залежність жанру створює ще одну серйозну перешкоду. Жанр може змінюватися залежно від культурного чи історичного середовища, що ускладнює класифікацію текстів, особливо якщо LLM навчалися на обмежених корпусах. Наприклад, один і той самий стилістичний набір може вважатися науковою фантастикою в одному контексті, а міфологічною прозою – в іншому.

Ще одним критичним аспектом є вплив навчальних даних. Якщо велика мовна модель була натренована на незбалансованому або нерепрезентативному корпусі, це може призводити до жанрових упереджень. Наприклад, модель, яка пройшла навчання переважно на англomовній літературі, може погано класифікувати тексти інших мов або не враховувати жанрові традиції літератури певних регіонів. Крім того, складнощі з обробкою стилістичних прийомів, таких як метафорика, іронія або міжтекстові зв'язки, можуть суттєво обмежувати можливості моделей у розпізнаванні складніших літературних жанрів.

Одним із перспективних напрямів вдосконалення жанрової класифікації є розробка багаторівневої класифікації, яка дозволить визначати не лише основний жанр, а й можливі піджанри або змішані жанрові структури. Це дозволить точніше відображати складну жанрову природу сучасної літератури. Важливим є також залучення міждисциплінарного підходу, коли великі мовні моделі

інтегруються з базами літературознавчих знань та жанровими онтологіями. Такий підхід може допомогти моделі не просто аналізувати тексти за окремими лексичними ознаками, а й враховувати загальноприйняті жанрові характеристики, що використовуються у філологічних дослідженнях.

Навчання на адаптивних корпусах, що охоплюють літературу різних мов, історичних періодів та жанрових напрямів, є ще одним способом підвищення точності класифікації. Крім того, контроль за упередженнями в даних може зменшити вплив популярних жанрів на результати класифікації. Використання додаткових інструментів для стилістичного аналізу також може значно покращити якість жанрового розпізнавання. Алгоритми, що аналізують ритм тексту, середню довжину речень, рівень складності синтаксису та частотність використання метафор, можуть підвищити точність жанрової ідентифікації.

Сприятливим фактором є також розвиток пояснюваних моделей штучного інтелекту, які не лише класифікують текст, але й можуть пояснити, чому він віднесений до того чи іншого жанру. Це відкриває можливість для перевірки результатів, що особливо важливо у літературознавстві та гуманітарних дослідженнях.

Таким чином, незважаючи на наявні труднощі, великі мовні моделі мають значний потенціал у жанровій класифікації літератури. Поєднання традиційних філологічних підходів із сучасними технологіями дозволить не лише покращити точність жанрової класифікації, а й створити нові можливості для дослідження літературних процесів. Подальші дослідження у цій галузі можуть бути зосереджені на вдосконаленні адаптивних моделей, інтеграції з літературними онтологіями та розробці методів для пояснення та інтерпретації результатів. Ці аспекти все частіше стають предметом наукових дискусій, що підтверджується дослідженнями, такими як (Sobchuk, Sela, 2023), де аналізуються методи кластеризації жанрів, або (Martin, Hood, 2024), що розглядає виклики застосування NLP у літературознавчих дослідженнях.

Висновки. Жанрова класифікація літератури є складним, багатовимірним процесом, який залишається актуальним як у традиційному літературознавстві, так і в сучасних дослідженнях у сфері автоматизованої обробки текстів.

Сучасні досягнення в галузі обробки природної мови, зокрема великі мовні моделі, відкривають нові можливості для жанрової класифікації, забезпечуючи автоматизацію цього процесу та підвищуючи його ефективність. Вони дозволяють аналізувати великі обсяги текстів, визначати їхню жанрову приналежність на основі лексичних, стилістичних та структурних характеристик, а також адаптувати класифікацію до нових літературних

тенденцій. Проте, як показує проведений аналіз, використання LLM у цьому контексті супроводжується рядом викликів.

Одним із головних питань, яке залишається відкритим, є проблема точності та інтерпретованості результатів.. Зокрема, великі мовні моделі інколи демонструють труднощі у визначенні змішаних жанрів, коли робота поєднувала риси оглядового дослідження та прикладного аналізу. Це свідчить про те, що сучасні методи автоматизованої класифікації потребують удосконалення для роботи з гібридними жанрами та перехідними жанровими формами. Крім цього, очевидною є необхідність контролю за балансом жанрового представлення в навчальних даних. Таким чином, подальший розвиток жанрової класифікації потребує формування збалансованих навчальних наборів, що містять достатньо різноманітні жанри для більш точної та коректної роботи моделей.

Ще один аспект, який вимагає уваги, — це пояснюваність роботи моделей. Недостатня пояснюваність може створювати проблеми для використання таких моделей у наукових та гуманітарних дослідженнях, де важливо не лише отримати результат, але й зрозуміти, як він був отриманий. Вирішення цієї проблеми можливе шляхом інтеграції методів Explainable AI (XAI), що дозволить створювати більш інтерпретовані системи жанрової класифікації.

Загалом, використання LLM для жанрової класифікації є перспективним напрямом, який має значний потенціал для автоматизованого аналізу літературних текстів. Проте, щоб досягти максимальної ефективності, необхідно вирішити низку проблем, зокрема покращити точність розпізнавання гібридних жанрів, забезпечити збалансованість навчальних даних, розробити методи пояснюваності класифікаційних рішень та застосувати багаторівневий підхід.

Подальші дослідження в цій сфері можуть бути зосереджені на інтеграції жанрових онтологій у роботу мовних моделей, використанні семантичних векторних просторів для визначення жанрової належності, а також на розробці нових алгоритмів, що враховують стилістичні та структурні особливості текстів. Важливим завданням залишається також адаптація методів жанрової класифікації для роботи з різними мовами та літературними традиціями, що дозволить створити універсальні системи класифікації для глобальних цифрових бібліотек та літературних архівів.

Таким чином, жанрова класифікація літератури за допомогою великих мовних моделей є складним, але перспективним напрямом досліджень. Вирішення цих питань дозволить значно покращити якість автоматизованої класифікації,

зробивши її більш точною, зрозумілою та корисною для різних сфер, від літературознавства до цифрових інформаційних систем.

ЛІТЕРАТУРА

1. Абрамчук, О.В. (23–24 березня 2016 р.). Жанри сучасної української літератури. *Матеріали XLV Науково-технічної конференції ВНТУ*. Вінниця, Україна. URL: <http://conferences.vntu.edu.ua/index.php/all-hum/all-hum-2016/paper/view/108/>.
2. Барабаш, С. (2011). Поняття “жанр” як теоретична категорія та чинник художньої специфіки творчості М. Козоріса. *Науковий часопис НПУ ім. М. П. Драгоманова. Серія 8. Філологічні науки (мовознавство та літературознавство)*, 11, 42–46.
3. Бехта, І. А., & Марчук, О. В. (2021). Структурно-типологічна параметризація художнього текстового простору англomовного фентезі. *Науковий вісник Міжнародного гуманітарного університету. Сер.: Філологія*, 47, 17–21.
4. Вибір редакції Amazon: добірка наших книжок. (2025). URL: <https://ridna-mova.com/vibir-redaktsii-amazon-dobirka-nashih-knigok.html>.
5. Ворочек, О. Г., & Соловей, І. В. (2024). Використання мовних моделей штучного інтелекту для генерації публікацій у соціальних мережах. *Технічна інженерія*, 93, 128–134. DOI: [https://doi.org/10.26642/ten-2024-1\(93\)-128-134](https://doi.org/10.26642/ten-2024-1(93)-128-134).
6. Гоца, Н. М. (2009). Проблема визначення поняття “стиль” жанру роману та методи його аналізу. *Філологічні трактати*, 3–4, 32–42.
7. Гулик, Ю. В. (2023). Використання штучного інтелекту для аналізу статистичних профілів текстів. *Науковий журнал Львівського державного університету безпеки життєдіяльності «Львівський філологічний часопис»*, 14, 29–35.
8. Жанри і жанрові процеси в історико-літературній перспективі. (2012). Кам’янець-Подільський : Аксіома.
9. Лозинська, Н. А. (2018). Феномен бестселера у сучасній видавничій справі: зарубіжний досвід та українські реалії. *Лідери книжкового ринку: матеріали Підсумкової науково-практичної конференції II туру Всеукраїнського конкурсу студентських наукових робіт із галузі знань «Журналістика»*. С. 1–53.
10. Самвидав на Amazon: Крок за кроком до публікації книги. (2025). URL: <https://mindthegraph.com/blog/uk/%D1%81%D0%B0%D0%BC%D0%B2%D0%B8%D0%B4%D0%B0%D0%B2-%D0%BD%D0%B0-%D0%B0%D0%BC%D0%B0%D0%B7%D0%BE%D0%BD%D1%96>.

11. Черниш, О. А. (2020). Теорія жанрів у сучасній лінгвістиці. *Вчені записки ТНУ імені В. І. Вернадського. Серія: Філологія. Соціальні комунікації*, 31, 70–75.
12. Apuk, V., & Nuci, K. P. (2021). Classification of Pedagogical Content Using Conventional Machine Learning and Deep Learning Models. <https://doi.org/10.48550/arXiv.2101.07321>.
13. Biswas, A., Dhabal, S., & Venkateswaran, P. (2023). Exploring Music Genre Classification: Algorithm Analysis and Deployment Architecture. <https://doi.org/10.48550/arXiv.2309.04861>.
14. Doulaty, M., Saz, O., & Raymond, W. M. (September 8–12, 2016). Automatic Genre and Show Identification of Broadcast Media. *17th Annual Conference of the International Speech Communication Association, Interspeech*. San Francisco, USA. P. 2115–2119. <https://doi.org/10.48550/arXiv.1606.03333>
15. Garbacz, P. (2006). An Outline of a Formal Ontology of Genres. (August 5–8, 2006). *Conference: Knowledge Science, Engineering and Management, First International Conference*. Guilin, China. P. 151–163. DOI: 10.1007/11811220_14.
16. Karjus, A. (2024). Machine-assisted quantizing designs: augmenting humanities and social sciences with artificial intelligence. <https://doi.org/10.48550/arXiv.2309.14379>.
17. Lepekhin, M., & Sharoff, S. (2022). Estimating Confidence of Predictions of Individual Classifiers and Their Ensembles for the Genre Classification Task. <https://doi.org/10.48550/arXiv.2206.07427>.
18. Martin, C., & Hood, D. (2024). The Use of Natural Language Processing in Literature Reviews. URL: <https://insights.axtria.com/hubfs/thought-leadership-whitepapers/Axtria-Insights-White-Paper-The-Use-of-Natural-Language-Processing-in-Literature-Reviews.pdf>
19. Mu, Y., Dong, C., Bontcheva, K., & Song, X. (2024). Large Language Models Offer an Alternative to the Traditional Approach of Topic Modelling. <https://doi.org/10.48550/arXiv.2403.16248>.
20. Netflix Help Center (2025). URL: <https://help.netflix.com/en>.
21. Schreiber, H. (2016). Genre Ontology Learning: Comparing Curated with Crowd-Sourced Folksonomies. (August 7–11, 2016). *Proceedings of the 17th International Society for Music Information Retrieval Conference*. New York City, USA. P. 400–406. URL: <https://archives.ismir.net/ismir2016/paper/000074.pdf>.
22. Sobchuk, O., & Sela, A. (2023). Computational thematics: Comparing algorithms for clustering the genres of literary fiction. <https://doi.org/10.48550/arXiv.2305.11251>.

Отримано: 10.03.2025
Прийнято: 09.04.2025