

ДИРІВ АНДРІЙ

Національний університет «Львівська політехніка»

<https://orcid.org/0009-0009-0615-1161>e-mail: andrii.i.dyriv@lpnu.ua**ЛОЗИНСЬКА ОЛЬГА**

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-5079-0544>e-mail: olha.v.lozynska@lpnu.ua

АНАЛІЗ ВПЛИВУ ЗНАКІВ МАТЕМАТИЧНИХ ОПЕРАЦІЙ У ПЛАГІАТІ МАТЕМАТИЧНИХ ФОРМУЛ ТА ЇХ ПЕРЕТВОРЕННЯХ

Дослідження присвячено вирішенню актуальної проблеми плагіату математичних формул у наукових публікаціях, яка набуває особливого значення у контекстах, де синтаксичні зміни можуть приховувати порушення авторських прав. Особливу увагу приділено комплексному аналізу впливу математичних операцій та їх символічного представлення на процеси виявлення плагіату. Основною метою роботи стало розроблення ефективного підходу до ідентифікації синтаксично модифікованих формул, структуру яких було змінено з метою маскування запозичень математичних концепцій без зміни їхнього семантичного змісту.

Особливості математичної мови, такі як жорстка семантика і структурованість виразів, ускладнюють адаптацію традиційних інструментів виявлення текстового плагіату до специфіки математичних текстів. Аналіз продемонстрував, що стандартні системи пропускають до 62% випадків плагіату у текстах, що включають математичні вирази. В умовах зростання обсягу цифрових публікацій дослідники виділили необхідність розробки міждисциплінарних підходів, які поєднують методи математичного аналізу, комп'ютерної лінгвістики та теорії обчислювань.

Дослідження охоплює кілька важливих компонентів. Методи синтаксичного аналізу — розбиття формул на окремі складові (токени), побудова дерев логічних операцій задля визначення їхньої структури. Семантичний аналіз — алгебраїчне спрощення та канонічне представлення формул, що дозволяє порівнювати математично еквівалентні вирази навіть за різних синтаксичних варіацій. Статистичний підхід — формування операційної сигнатури формул на основі частотного розподілу операторів та символів для швидкого виявлення подібностей. Векторні моделі — векторизація математичних виразів, що відкриває можливість застосування алгоритмів машинного навчання для аналізу формул.

Особлива увага приділяється впливу математичних операцій на ефективність виявлення плагіату. Наприклад, арифметичні операції демонструють високу частоту використання та є відносно легкими для аналізу завдяки семантичній стабільності. Натомість степеневі, функціональні та інтегрально-диференціальні операції, через їхній складний семантичний зміст і численні правила трансформації, створюють значно більші виклики для автоматизованих систем виявлення.

У ході експериментів, аналіз впливу операцій проведено на кількох рівнях. Нормалізація представлення формул, що дозволяє зменшити синтаксичну варіативність. Оцінка структурної подібності за деревами операцій, що дозволяє виявити еквівалентність навіть за прихованих змін. Порівняння коефіцієнта семантичної подібності, що базується на алгебраїчних перетвореннях.

Важливим результатом дослідження стало створення таблиці й оцінки впливу різних типів операцій. Вона демонструє, що функціональні та інтегрально-диференціальні оператори є критичними для виявлення плагіату через їхню складність і варіативність у записі. Наприклад, формули з логарифмами чи тригонометричними функціями можуть бути представлені у багатьох рівнозначних формах із суттєвими візуальними відмінностями.

Дослідження також підкреслює необхідність удосконалення існуючих алгоритмів токенизації, нормалізації та семантичного аналізу. У висновках зазначається, що інтеграція машинного навчання та формування гібридних підходів стане ключем до створення потужних систем перевірки математичного контенту. Робота пропонує перспективи подальшого розвитку інструментів академічного моніторингу, спрямованих на забезпечення доброчесності в науці та підвищення загального рівня прозорості дослідницької діяльності.

Ключові слова: плагіат формул, знаки математичних операцій, аналіз формул, інтелектуальна власність, методи виявлення плагіату, математичний контент

DYRIV ANDRII, LOZYNKA OLGA

Lviv Polytechnic National University

ANALYSIS OF THE IMPACT OF MATHEMATICAL OPERATION SIGNS IN MATHEMATICAL FORMULA PLAGIARISM AND THEIR TRANSFORMATIONS

The research is devoted to addressing the pressing issue of plagiarism in mathematical formulas within scientific publications, which holds particular importance in contexts where syntactic changes can obscure copyright violations. Special attention is given to the comprehensive analysis of the impact of mathematical operations and their symbolic representations on plagiarism detection processes. The primary objective of the study is to develop an effective approach for identifying syntactically modified formulas, whose structure has been altered to conceal the borrowing of mathematical concepts without changing their semantic meaning.

The specific features of mathematical language, such as rigid semantics and structured expressions, complicate the adaptation of traditional text plagiarism detection tools to the peculiarities of mathematical texts. The analysis revealed that standard systems fail to detect up to 62% of plagiarism cases in texts that include mathematical expressions. With the increasing volume of digital publications, researchers emphasized the necessity of developing interdisciplinary approaches that integrate methods of mathematical analysis, computational linguistics, and theoretical computation.

The research encompasses several key components. Syntactic analysis: Decomposition of formulas into distinct components (tokens) and the construction of logical operation trees to determine their structure. Semantic analysis: Algebraic simplification and canonical representation of formulas that allow the comparison of mathematically equivalent expressions, even with varying syntactic forms. Statistical approach: Formation of operational signatures of formulas based on frequency distributions of operators and symbols to quickly detect similarities. Vector models: Vectorization of mathematical expressions, enabling the application of machine learning algorithms to analyze formulas.

Particular attention is paid to the effects of mathematical operations on the efficiency of plagiarism detection. For example, arithmetic operations exhibit high usage frequency and are relatively straightforward to analyze due to their semantic stability. In contrast, exponential, functional, and integral-differential operations pose significant challenges to automated systems due to their intricate semantics and numerous transformation rules. During the experiments, the impact of operations was analyzed on several levels. Normalization of formula representations to reduce syntactic variability. Assessment of structural similarity using operational trees to identify equivalences despite concealed modifications. Comparison of semantic similarity coefficients, based on algebraic transformations.

A significant result of the study was the creation of a table and evaluation of the influence of different types of operations. It demonstrated that functional and integral-differential operators are critical for plagiarism detection due to their complexity and variability in representation. For example, formulas involving logarithmic or trigonometric functions can be expressed in multiple equivalent forms with substantial visual differences.

The research also highlights the necessity of improving existing algorithms for tokenization, normalization, and semantic analysis. The conclusions underscore that the integration of machine learning and the development of hybrid approaches will be pivotal in creating robust systems for verifying mathematical content. The study offers prospects for further development of academic monitoring tools aimed at ensuring integrity in science and enhancing the overall level of transparency in research activities.

Keywords: formula plagiarism, mathematical operation signs, formula analysis, intellectual property, plagiarism detection methods, mathematical content.

Стаття надійшла до редакції / Received 26.05.2025

Прийнята до друку / Accepted 25.06.2025

Вступ

Актуальність проблеми. Проблема плагіату в науковій літературі набуває особливого значення в математичному контексті, де незначні зміни в позначеннях операцій можуть маскувати неоригінальний зміст. Математична спільнота стикається з численними випадками неправомірного запозичення формул без належного цитування, що ускладнюється багатоманітністю представлень ідентичних математичних концепцій [1][2]. Традиційні інструменти виявлення плагіату, орієнтовані переважно на текстовий аналіз, виявляються малоефективними для математичних нотацій, де семантична еквівалентність часто прихована за синтаксичними відмінностями [3]. Зростаюча доступність наукових публікацій у цифровому форматі та поширення електронних баз даних підвищують актуальність розробки спеціалізованих методів виявлення математичного плагіату для захисту інтелектуальної власності та підтримки академічної доброчесності.

Виявлення плагіату в математичних формулах становить особливий виклик через унікальну природу математичної мови. На відміну від звичайного тексту, математичні вирази мають чітко визначену структуру та семантику, яка може зберігатися при значних синтаксичних змінах. Дослідження показують, що стандартні системи виявлення плагіату демонструють низьку ефективність при роботі з математичними текстами, пропускаючи до 62% випадків плагіату у формулах [4]. Це створює серйозну проблему для академічної спільноти, особливо у галузях з високою концентрацією математичного контенту, таких як фізика, інженерія, комп'ютерні науки та економіка.

Мета та завдання дослідження

Дослідження спрямоване на всебічний аналіз впливу математичних операцій та їх символічних представлень на процеси виявлення плагіату в математичних формулах. Основною метою є розробка теоретичного підґрунтя та практичних підходів для ефективного ідентифікації трансформованих математичних виразів, які зберігають семантичний зміст при зміні формальної структури. Особлива увага приділяється встановленню об'єктивних критеріїв розмежування допустимих математичних перетворень та неправомірних запозичень.

Враховуючи складність проблеми, дослідження має міждисциплінарний характер, інтегруючи методи математичного аналізу, теоретичної інформатики, комп'ютерної лінгвістики та теорії академічної доброчесності. Такий підхід дозволяє розглянути проблему з різних перспектив та розробити комплексні рішення, що відповідають потребам наукової спільноти.

Об'єкт та предмет дослідження

Об'єктом дослідження виступають математичні формули та вирази в науковій літературі, їх символічні представлення та різноманітні трансформації. Ці формули можуть належати до різних галузей математики та її застосувань, включаючи алгебру, аналіз, геометрію, теорію ймовірностей, статистику та прикладну математику. Особливий інтерес становлять формули, що зазнали перетворень, які потенційно могли бути застосовані для маскування плагіату.

Предметом дослідження є методи та алгоритми виявлення подібності між математичними формулами, вплив математичних операцій на процеси ідентифікації плагіату, а також критерії розмежування легітимних перетворень та неправомірних запозичень математичних виразів. Дослідження охоплює як синтаксичні, так і семантичні аспекти порівняння математичних виразів, а також алгоритмічні підходи до автоматизованого виявлення плагіату в математичних текстах.

Теоретичні основи дослідження

Математичні операції та їх властивості. Математичні операції становлять фундаментальні елементи формальної мови математики, забезпечуючи інструментарій для обчислень та перетворень. Базові математичні операції охоплюють арифметичні дії (додавання, віднімання, множення, ділення), степеневі операції (піднесення до степеня, добування кореня), інтегрально-диференціальні операції (інтегрування, диференціювання) та агрегаційні операції (сумування, множення серій). Кожна з цих операцій може бути представлена різними символами: множення, наприклад, позначається як $*$, $\cdot$, або $\times$, що створює первинну варіативність математичних записів [5].

Інтегрально-диференціальні операції представляють особливий інтерес у контексті виявлення плагіату через їхню складну структуру та різноманітність представлень. Наприклад, похідна функції $f(x)$ може

бути записана як $f(x)$, df/dx , або $Df(x)$, що створює додаткові можливості для варіативного представлення однакових математичних концепцій. Аналогічно, інтеграли можуть мати різні форми запису, включаючи різноманітні позначення меж інтегрування та диференціала [6].

Специфіка плагіату в математичних формулах. Плагіат у сфері математичних формул має низку специфічних характеристик, що відрізняють його від текстового плагіату [7]. Універсальність математичної нотації зумовлює природну подібність незалежно створених формул, а обмеженість альтернативних представлень для деяких математичних концепцій ускладнює встановлення факту неправомірного запозичення. Семантична еквівалентність при синтаксичній варіативності дозволяє створювати видимість оригінальності шляхом незначних модифікацій, що не змінюють математичного змісту.

Серйозну проблему створює складність атрибуції математичних формул, особливо коли йдеться про загальновідомі концепції. Розмежування плагіату та незалежного відкриття або використання загальноприйнятих математичних положень потребує глибокого аналізу контексту та історії розвитку відповідної галузі. Як зазначає Сталнакер, у математиці часто виникають ситуації, коли схожі формули розробляються незалежно різними дослідниками, що працюють над подібними проблемами [2].

Ілюстративним прикладом складності виявлення плагіату в математичних формулах є відомий випадок з 2012 року, коли професор математики виявив, що його робота з комбінаторної геометрії була майже дослівно відтворена в іншій публікації [6]. Плагіатор замінив лише позначення змінних та дещо змінив порядок викладу, при цьому зберігши всі основні доведення та висновки. Цей випадок став показовим прикладом того, як незначні синтаксичні зміни використовуються для маскування факту запозичення математичних результатів.

Іншим показовим випадком є скандал 2007 року в одному з провідних математичних журналів, коли дві статті з теорії графів містили ідентичні формули з множинними косметичними змінами для маскування запозичення [3]. Аналіз цих публікацій виявив систематичну заміну операторів та перегруповання виразів, що зберігало математичну суть формул при значній зміні їхнього зовнішнього вигляду. Цей випадок підкреслив необхідність розробки спеціалізованих методів виявлення плагіату, здатних розпізнавати семантичну еквівалентність формул при наявності синтаксичних відмінностей.

Системний аналіз

Методи аналізу математичних виразів

Ефективний аналіз математичних виразів для виявлення плагіату потребує комплексного підходу, що поєднує різні методологічні стратегії [8]. Синтаксичний аналіз забезпечує розбиття формул на базові елементи з побудовою структурного дерева, що відображає ієрархію операцій та взаємозв'язки між компонентами. Такий підхід дозволяє визначити структурну відповідність між порівнюваними формулами навіть при зміні символічних позначень [9].

Семантичний аналіз спрямований на виявлення математичного змісту формул незалежно від їхньої синтаксичної форми, що досягається шляхом алгебраїчних перетворень виразів до канонічної форми [4]. Цей метод вимагає застосування систем символічних обчислень, здатних виконувати такі операції як розкриття дужок, зведення подібних членів, факторизація та інші алгебраїчні перетворення. Незважаючи на обчислювальну складність, семантичний аналіз є необхідним компонентом ефективної системи виявлення плагіату, особливо для формул, що зазнали складних трансформацій.

Статистичний аналіз операцій передбачає дослідження частотних характеристик математичних символів та операторів, формуючи сигнатуру формули, яка може слугувати своєрідним "відбитком пальців" виразу. Порівняння сигнатур дозволяє швидко ідентифікувати потенційно подібні формули, що є особливо корисним для аналізу великих обсягів математичних текстів.

Векторні представлення формул дозволяють трансформувати математичні вирази у числові вектори, що відкриває можливість застосування методів машинного навчання для класифікації та кластеризації формул за ступенем подібності. Цей підхід особливо ефективний для виявлення неочевидних зв'язків між формулами та ідентифікації шаблонів плагіату в великих корпусах математичних текстів.

Методи нормалізації математичних виразів доповнюють інструментарій аналізу, забезпечуючи можливість чисельного порівняння та приведення формул до стандартизованого вигляду [1]. Нормалізація включає такі процедури як стандартизація позначень операцій, впорядкування операндів, спрощення виразів та інші трансформації, що зберігають математичний зміст при уніфікації синтаксичної форми.

Комбінація цих підходів створює багаторівневу систему аналізу, здатну виявляти подібності між математичними виразами навіть при наявності трансформацій [5]. Важливо зазначити, що програмна реалізація методів аналізу математичних виразів потребує спеціалізованих алгоритмів для обробки математичної нотації.

Критерії оцінювання подібності формул

Об'єктивна оцінка подібності математичних формул вимагає застосування комплексу критеріїв, що враховують різні аспекти математичних виразів. Коефіцієнт структурної подібності оцінює відповідність синтаксичних структур порівнюваних формул, аналізуючи подібність дерев синтаксичного розбору, послідовність операторів та типи використаних математичних операцій. Цей коефіцієнт обчислюється на основі алгоритмів порівняння дерев, таких як алгоритм найбільшої спільної підструктури або методи обчислення відстані редактування для дерев [9].

Семантична еквівалентність встановлює ідентичність математичного значення формул після нормалізації та спрощення, що є найбільш вагомим, але й найскладнішим для автоматизованого обчислення критерієм. Визначення семантичної еквівалентності вимагає застосування методів символьних обчислень та алгебраїчних перетворень, здатних трансформувати вирази до канонічної форми для порівняння.

Додаткові показники, такі як коефіцієнт спільних підвиразів та трансформаційна відстань, доповнюють аналітичний інструментарій, дозволяючи виявляти часткові запозичення та оцінювати ступінь модифікації вихідної формули [5]. Коефіцієнт спільних підвиразів визначає частку ідентичних фрагментів у порівнюваних формулах, тоді як трансформаційна відстань вимірює кількість елементарних перетворень (заміна оператора, зміна порядку операндів тощо), необхідних для перетворення однієї формули в іншу.

У практичному застосуванні ці критерії часто інтегруються в комплексний показник подібності, що забезпечує більш надійну основу для прийняття рішень щодо наявності плагіату [3]. Цей показник може бути реалізований як зважена сума окремих коефіцієнтів, де ваги визначаються експериментально відповідно до специфіки предметної області та характеру аналізованих формул.

Алгоритми виявлення трансформацій

Для ефективного виявлення різноманітних трансформацій у математичних формулах розроблено спеціалізовані алгоритмічні підходи [1]. Алгоритми токенизації та нормалізації забезпечують перетворення математичних виразів у послідовність стандартизованих токенів, спрощуючи подальший аналіз. Токенизація математичних виразів вимагає врахування специфіки математичної нотації, включаючи обробку дробів, індексів, степенів та інших спеціальних конструкцій.

Алгоритми структурного порівняння застосовують методи аналізу дерев для виявлення ізоморфізму або часткової відповідності синтаксичних структур, що дозволяє ідентифікувати подібні формули навіть при зміні позначень змінних та операторів [9]. Ці алгоритми включають методи пошуку найбільшої спільної підструктури, обчислення відстані редагування для дерев та інші підходи, що дозволяють кількісно оцінити структурну подібність математичних виразів.

Векторно-просторові моделі представляють математичні вирази у вигляді числових векторів, уможливаючи застосування методів машинного навчання для класифікації та кластеризації формул за ступенем подібності. Вектори можуть формуватися на основі різних характеристик формул, включаючи частоти операторів, структурні особливості та семантичні ознаки.

Алгоритми семантичного аналізу використовують методи символьних обчислень для приведення виразів до канонічної форми, що дозволяє виявляти семантично еквівалентні формули незалежно від їхньої синтаксичної структури [4]. Ці алгоритми включають методи алгебраїчного спрощення, розкриття дужок, зведення подібних членів, факторизації та інших перетворень, що зберігають математичний зміст виразів.

Ключовими характеристиками ефективних алгоритмів виявлення трансформацій є стійкість до синтаксичних варіацій, чутливість до семантичних відмінностей, масштабованість та адаптивність до специфіки предметної області [3]. Розробка та вдосконалення таких алгоритмів становить активний напрям досліджень у сфері виявлення математичного плагіату, що постійно розвивається у відповідь на нові виклики та потреби наукової спільноти.

Проектування системи

На даному етапі хорошим варіантом буде зробити діаграму активностей, щоб на її основі продемонструвати які основні вузли активностей будуть в загальній системі та зокрема в поточному рішенні, яке детальніше досліджується в даній науковій роботі (Рис. 1).

Діаграма активностей ілюструє етапи роботи системи або дослідницького методу, призначеного для аналізу математичних формул та виявлення можливого плагіату. Процес розпочинається з завантаження даних, яке є вихідною точкою для всіх дій, пов'язаних із перевіркою текстів на математичний плагіат. Потім система розділяється на різні підсистеми, де основну увагу ми зосередимо на частині де перевіряється, чи містяться в текстах математичні формули. Якщо такі формули виявлено, процес переходить до наступних етапів їх аналізу, а якщо ні, дослідження завершується через відсутність релевантних даних.

Коли формули знайдено, вони піддаються процесу токенизації – розбивці формули на елементарні компоненти (токени), такі як операнди, змінні, знаки операцій тощо. Далі здійснюється аналіз операційної структури, що передбачає дослідження ієрархії операцій у формулі. Цей етап допомагає представити формулу у вигляді дерева операцій, наприклад, як дерево, де коренем є знак додавання, а гілки відповідають множенню та операндам.

На синтаксичному етапі система порівнює структури формул, аналізуючи схожість між деревами операцій. Основне завдання – виявити схожість між оригінальними та можливо видозміненими формулами. Після цього проходить семантичний аналіз, де визначається математична сутність формули через алгебраїчні перетворення, спрощення чи нормалізацію.

Додатково система виконує нормалізацію, що передбачає стандартизацію позначень, оператори уніфікуються. Іншим важливим етапом є векторизація формул – уявлення математичних виразів у вигляді векторів, що відображають частоту елементів. Ці векторні представлення дозволяють здійснювати автоматичне порівняння формул за ступенем подібності.

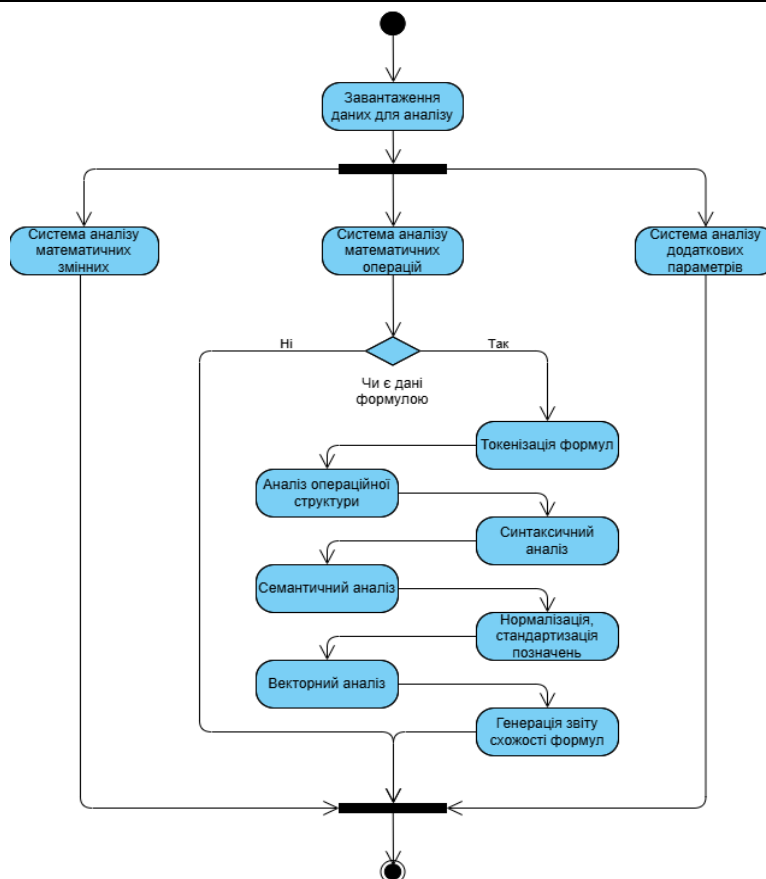


Рис. 1. Діаграма активностей

На основі виконаних розрахунків система визначає, чи є формули підозрілими. Якщо система виявляє високий коефіцієнт подібності, формули заносяться до списку потенційного плагіату, а якщо подібність недостатня, формули класифікуються як унікальні. Усі результати згодом фіксуються у фінальному звіті, який містить висновки щодо рівня подібності між формулами. Завершення процесу настає, коли всі формули перевірено, результати опрацьовано, а формули класифіковано згідно зі ступенем їхньої ідентичності.

Виклад результатів дослідження

Програма аналізу впливу математичних операцій на схожість формул

Напишемо програму, яка допоможе детальніше дослідити виявлення математичного плагіату. Для тестування було підготовано наступний датасет набору формул:

Таблиця 1

Датасет формул, що передаються для аналізу

Назва змінної в коді	Значення змінної
F1	$2x + y^2$
F2	$2 \cdot x + y^{**2}$
F3	$x^2 + 2x + 1$
F4	$(a+b)^2$
F5	$a^2 + 2ab + b^2$
F6	$x^3 - 3x^2 + 3x - 1$
F7	$2x + y + \sqrt{z}$
F8	$\sum_{i=1}^n i^2$
F9	$\prod_{k=1}^n k$
F10	$\int_{(0 \text{ to } 1)} x^2 dx$
F11	$a/b + cd$
F12	$(x-1)^2 + (y+1)^2 = 1$
F13	$1/(1+x^2)$
F14	$\tan(x) + \cos(x)$
F15	$\sin(x)^2 + \cos(x)^2$

У результаті виконання програми було побудовано графік у якому було наведено інформацію кількості операцій у кожній формулі що були розпізнані (Рис. 2). Також була виведена кругова діаграма для демонстрації частоти появи різних операцій у наведених формулах, що є дуже інформативним саме для дослідницької мети даної роботи (Рис. 3). Найважливішим аспектом в плані результативності є теплова мапа (Рис. 4), адже нас саме цікавить який відсоток схожості є у попарному порівнянні всіх формул і стовпчикова діаграма, яка демонструє розподіл коефіцієнтів схожості для заданого датасету (Рис. 5).



Рис. 2. Графік відношення кількості операцій до формул

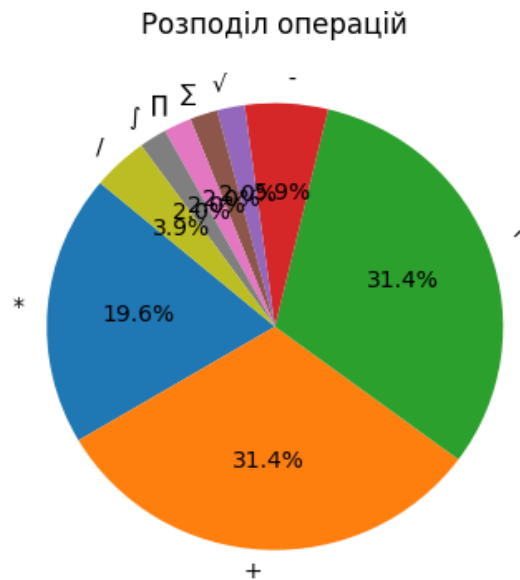


Рис. 3. Кругова діаграма відсотку складу операцій в датасеті

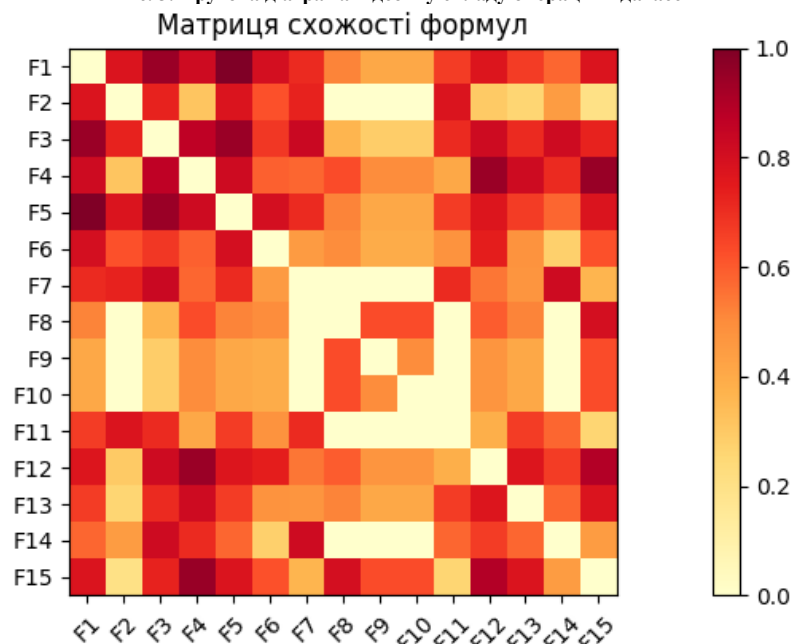


Рис. 4. Теплова мапа схожості формул

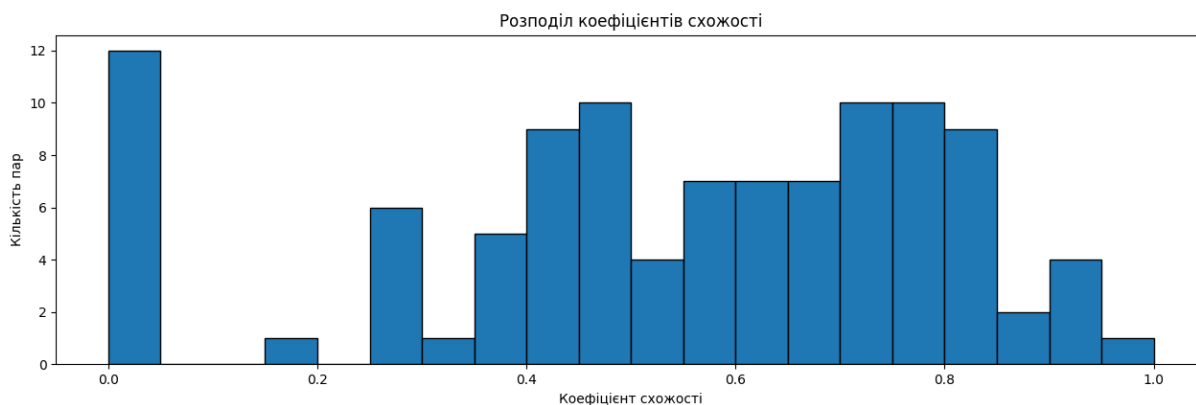


Рис. 5. Стовпчикова діаграма розподілу коефіцієнтів попарної схожості між формулами

Аналіз впливу математичних операцій

Дослідження показало, що різні типи математичних операцій неоднаково впливають на ефективність виявлення плагіату. Арифметичні операції характеризуються високою частотою використання та численними альтернативними позначеннями, що створює сприятливі умови для маскування плагіату шляхом простої заміни символів. Водночас, стандартизація нотації дозволяє досить ефективно виявляти такі трансформації, оскільки семантика операцій зберігається незалежно від їхнього синтаксичного представлення.

Степеневі операції демонструють середню частоту використання та значне різноманіття представлень, що ускладнює їх виявлення [5]. Крім того, степеневі операції мають специфічні властивості та правила перетворення, які можуть використовуватися для складних маскувальних трансформацій. Дослідження Шуботца та його колег показало, що виявлення плагіату у формулах зі степеневими операціями вимагає комбінації синтаксичного та семантичного аналізу для досягнення прийнятного рівня точності.

Функціональні операції (тригонометричні, логарифмічні) є одними з найскладніших для аналізу в контексті виявлення плагіату [4]. Ці операції характеризуються специфічними властивостями та численними тотожностями, які дозволяють створювати різноманітні еквівалентні представлення формул. Крім того, функціональні операції можуть мати різні позначення ($\log$, $\ln$, $\lg$) та форми запису параметрів, що додатково ускладнює їх аналіз.

Інтегрально-диференціальні операції представляють особливий виклик для систем виявлення плагіату через їхню високу семантичну складність та варіативність нотації [6]. Наприклад, похідна функції може бути записана різними способами (df/dx , f' , Df), а інтеграли можуть мати різні форми запису меж та диференціала. Крім того, застосування правил інтегрування та диференціювання дозволяє створювати математично еквівалентні, але візуально відмінні формули.

Кількісний аналіз впливу різних типів операцій на ефективність виявлення плагіату, проведений на основі експериментальних даних (Табл. 2).

Таблиця 2

Математичні операції у формульному плагіаті

Тип операції	Частота використання	Складність виявлення	Вплив на подібність
Додавання/віднімання	Висока	Низька	Помірний
Множення/ділення	Висока	Середня	Значний
Степеневі	Середня	Висока	Значний
Функціональні	Середня	Дуже висока	Критичний
Інтегрально-диф.	Низька	Дуже висока	Критичний

Ці результати підтверджують, що складність виявлення плагіату зростає зі зростанням семантичної складності операцій. Арифметичні операції, які мають просту семантику та обмежений набір властивостей, відносно легко піддаються аналізу, тоді як функціональні та інтегрально-диференціальні операції, що характеризуються складною семантикою та численними правилами перетворення, створюють значні труднощі для автоматизованих систем виявлення плагіату.

Висновки

Наукові роботи, насичені математичними формулами, особливо вразливі до плагіату. Використання базових властивостей математичних операцій та відомих тотожностей дозволяє створювати видимість оригінальності шляхом синтаксичних і структурних перетворень.

Розвиток систем для виявлення плагіату в математичних текстах має бути направлений на побудову алгоритмів, здатних поєднувати структурний, синтаксичний та семантичний аналіз. Особливої ефективності можна досягти завдяки алгоритмам нормалізації та технологіям машинного навчання для класифікації подібних виразів. У майбутньому дослідження можливості автоматичного виявлення регулярного повторення

певних шаблонів трансформацій матиме великий вплив на розвиток наукової етики. Також систему можна удосконалювати, використовуючи модулі, одним з яких, наприклад, може бути аналіз схожості на основі змінних у формулі [10].

Література

1. Altynbek A., Cemil T., Alfira M. (2025). Plagiarism types and detection methods: a systematic survey of algorithms in text analysis. <https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2025.1504725/full>
2. Ahmed M H., Abu Elnasr E S., (2023). Drivers and Consequences of ChatGPT Use in Higher Education: Key Stakeholder Perspectives. <https://www.mdpi.com/2254-9625/13/11/181>
3. Mohamed Ahmed E., Ramy G. M., Nawal A. E., Marwa A. S., (2023). An effective text plagiarism detection system based on feature selection and SVM techniques. <https://link.springer.com/article/10.1007/s11042-023-15703-4>
4. Ankit S., Andr'e G., Noah G., Isabel B., Moritz S., Olaf T., Akiko A., Bela G. (2024). Taxonomy of Mathematical Plagiarism. https://link.springer.com/chapter/10.1007/978-3-031-56066-8_2
5. Gunwoo L., Jindae K., Myung-seok C., Rae-Young J., Ryong L. (2023). Review of Code Similarity and Plagiarism Detection Research Studies. <https://www.mdpi.com/2076-3417/13/20/11358>
6. Anupama P., Parul G., Yogesh K. D. (2023). Plagiarism awareness efforts, students' ethical judgment and behaviors: a longitudinal experiment study on ethical nuances of plagiarism in higher education. https://www.researchgate.net/publication/373616091_Plagiarism_awareness_efforts_students'_ethical_judgment_and_behaviors_a_longitudinal_experiment_study_on_ethical_nuances_of_plagiarism_in_higher_education
7. Norman M., Moritz S., Felix H., Tomas S., Bela G. (2017). Analyzing Mathematical Content to Detect Academic Plagiarism. https://www.researchgate.net/publication/320882205_Analyzing_Mathematical_Content_to_Detect_Academic_Plagiarism
8. Mohamed A. E., Ramy G. M., Nawal A. E., Marwa A. S. (2024). An effective text plagiarism detection system based on feature selection and SVM techniques. https://www.researchgate.net/publication/370801210_An_effective_text_plagiarism_detection_system_based_on_feature_selection_and_SVM_techniques
9. Ahmed M. E., Khaled E., Saeed A. (2021). Some students plagiarism tricks, and tips for effective check. https://www.researchgate.net/publication/353483955_Some_students_plagiarism_tricks_and_tips_for_effective_check
10. Dyriv A., Lozynska O.(2024). analysis of the plagiarism issues in scientific publications and research of comparison method based on variable similarity. <https://heraldts.khmnu.edu.ua/index.php/heraldts/article/view/170>