УДК 004.8

МЕТОД ИЗВЛЕЧЕНИЯ ГЛАВНЫХ РЕЗУЛЬТАТОВ И ФОРМУЛ ИЗ ТЕКСТОВ НАУЧНЫХ СТАТЕЙ НА ОСНОВЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

Э. В. Саитов[1] [0009-0009-1980-8740], О. А. Невзорова[2] [0000-0001-8116-9446]

[1, 2]*Казанский (Приволжский) федеральный университет, г. Казань, Россия*

[1]saitoverik@mail.ru, [2]onevzoro@gmail.com

Аннотация

Представлен метод извлечения главных результатов и формул из текстов математических научных статей на основе большой языковой модели DeepSeek. Разработанный метод состоит из следующих шагов: формирование базы аннотаций статей, выделение из них формулировок главных результатов, поиск этих результатов в основном тексте, извлечение в найденных фрагментах ключевых формул с переводом в LaTeX.

Получены оценки построенного решения по метрикам точности, полноты и F-меры. Средние значения F-меры по тестовой коллекции математических статей для основных результатов и связанных с ними формул равны соответственно 0.93459 и 0.9294.

В качестве способа хранения полученных главных результатов и формул предложена система управления базами данных MySQL, приведено описание соответствующей структуры базы данных.

Ключевые слова*: большая языковая модель, математическая статья, основная формула, главная формула, формат LaTeX, извлечение формулы, семантический поиск, аннотирование формул, обработка PDF-документов, информационный поиск, искусственный интеллект, наукометрический анализ.*

ВВЕДЕНИЕ

Автоматизированное извлечение и семантический анализ ключевых результатов из научных публикаций представляют собой серьезную исследовательскую проблему в современную эпоху. Основные результаты математических работ, как правило, представляются в виде формул. Прямое извлечение

этих формул из формата PDF, стандартного для научной коммуникации, и последующее определение роли и значимости формул в контексте всей работы сопряжены с рядом технических и семантических сложностей.

Развитие больших языковых моделей (Large Language Model, LLM), обладающих продвинутыми способностями к пониманию контекста и генерации структурированных ответов, создает принципиально новые возможности для решения указанной проблемы.

В настоящей работе предложен и оценен подход к извлечению из математических статей главных результатов и связанных с ними формул с использованием LLM DeepSeek[1]. Разработанный метод использует возможности LLM как семантического поисковика, что обеспечивает поиск как внутри аннотации статьи, так и в ее содержимом. Результатами поиска внутри аннотации являются описания основных результатов статьи. Далее выполняется поиск выделенных из аннотации результатов внутри тела статьи, локализация соответствующих фрагментов и извлечение формул, связанных с основными результатами.

Структура статьи такова: в разделе 1 дан обзор основных подходов и методов, применяемых для решения задачи анализа математических документов; в разделе 2 описан разработанный метод извлечения главных результатов и формул из текстов математических научных статей; в разделе 3 даны оценки разработанного метода, в разделе 4 описана концепция формирования базы данных хранения полученных результатов; в заключении представлены основные выводы и направления будущих разработок.

1. БЛИЗКИЕ ПО ТЕМАТИКЕ РАБОТЫ

Исследования, связанные с построением методов обработки математических текстов, достаточно широко представлены в литературе. Представим наиболее значимые результаты.

Цикл работ по применению математических онтологий в задачах обработки математических текстов включает [1–4].

В [1] рассмотрены проблемы семантического аннотирования формул в формате PDF. Авторы предложили метод, который комбинирует синтаксический анализ внутренней структуры формулы (построение дерева операторов)

[1] https://chat.deepseek.com

с лингвистическим анализом окружающего ее текста, и сопоставление элементов формулы с концептами математической онтологии OntoMath^{PRO}. В [2] описаны семантические сервисы цифровой экосистемы OntoMath, применяемые в задачах математического образования. Функционал этих сервисов основан на использовании математических онтологий OntoMath^{Edu} и OntoMath^{PRO}. В частности, для задач математического образования разработаны сервисы семантического поиска по математическим формулам и семантического аннотирования учебных материалов, что позволило представить последние как интерактивные элементы, связанные с пояснениями и определениями. В работе [3] методы, ранее разработанные в [1], применены для аннотирования материалов научной конференции, что подтверждает масштабируемость и универсальность подхода к семантической разметке для различных видов математических PDF-документов. В [4] предложен сервис семантического поиска по коллекции математических статей в формате PDF. Поиск формул выполняется посредством поиска понятий, входящих в математическую формулу. В качестве источника понятий используется онтология профессиональной математики OntoMath^{PRO}. Результаты поиска не зависят от авторских обозначений переменных формулы и содержат все формулы, в которых присутствует понятие из поискового запроса.

В [5] рассмотрена задача поиска математических формул, имеющих древовидную структуру. Предложены три типа мер, представляющих отличительные особенности семантического сходства математических формул, разработаны эффективные алгоритмы на основе хеширования для приближенный вычислений. Выполнены эксперименты с использованием набора данных NTCIR-11 Math-2 Task – крупномасштабной тестовой коллекции для поиска математической информации, содержащей около 60 млн формул. Показано, что предложенный метод повышает точность поиска, сохраняя при этом высокую масштабируемость и эффективность по времени выполнения.

В обзорной работе [6] даны классификация и анализ ранних математических поисковых систем. Авторы систематизировали системы по архитектуре и типу поддерживаемых запросов. Главным результатом стало выявление имеющихся фундаментальных проблем: неэффективность при обработке концеп-

туальных (неточных) запросов и семантический разрыв между текстом и формулой.

Работа [7] относится к области семантических веб-технологий и связана с успешной интеграцией более 100 тыс. математических объектов в глобальную базу знаний Wikidata.

В настоящее время активно разрабатываются нейросетевые методы анализа математических документов.

В статье [8] рассмотрен метод извлечения математических концептов из текста с оценкой F-меры, равной 0.87, комбинирующий синтаксический парсинг зависимостей (для анализа структуры предложения) с контекстными языковыми моделями BERT. В [9] описан метод векторизации математической лексики с применением векторной модели word2vec, обученной на крупном корпусе математических текстов из ресурса arXiv. Конкретным результатом стало создание “math-aware” эмбеддингов, использование которых повысило точность классификации математических документов на 5–10%.

Обзорная статья [10] систематизирует более 150 современных методов извлечения информации, основанных на глубоком обучении. Авторы классифицируют методы по типам нейросетевых архитектур (например, RNN, CNN, Transformers) и решаемым задачам (NER, relation extraction). Этот обзор задает общий контекст развития решаемых задач в области машинного обучения.

В статьях [11, 12] представлена реализация гибридной поисковой системы MathIRs, объединяющей различные методы. Одним из интересных результатов является использование векторных представлений формул (Formula Embedding). Модуль встраивания формул предлагаемой системы использует таблицу битовой информации для преобразования математических формул, содержащихся в научных документах, в бинарные векторы формул. Каждый установленный бит вектора формулы обозначает наличие определенной математической сущности. Математический запрос пользователя преобразуется в вектор запроса аналогичным образом и извлекаются соответствующие релевантные документы. Релевантность результата поиска характеризуется степенью сходства между индексированным вектором формулы и вектором запроса. Этот подход продемонстрировал эффективность в поиске точных формул запроса, родительских формул, подформул и похожих формул. Модификация

критерия меры сходства и включение поддержки индексирования и поиска текстовых терминов являются важными будущими модификациями, которые могут еще больше повысить эффективность поиска.

Экспериментальное исследование [13] доказало важность использования мультимодальности. Авторы показали, что комбинирование визуальных признаков формулы (ее расположения на странице – layout) с семантическими признаками дает значительный прирост точности поиска. Конкретным результатом стало повышение этой точности на 15–20% по сравнению с использованием только одного типа признаков.

В [14] представлен итоговый отчет по проекту DFG “MathIR” (2018–2022). В его рамках были исследованы новые подходы и технологии повышения доступности математического контента и его семантической информации для широкого спектра приложений информационного поиска. Для достижения этой цели решались три основные исследовательские задачи: (1) синтаксический анализ математических выражений, (2) семантическое обогащение математических выражений и (3) оценка с использованием метрик качества и демонстраторов. Проект MathIR внес значительный вклад в исследования различных задач и систем математического информационного поиска, включая системы обнаружения плагиата и рекомендации, поисковые системы, системы первой помощи для математических типов данных, системы ответов на математические вопросы и обучения, автоматическую проверку правдоподобности математических выражений в Википедии, автоматическую вычислимость математического контента с использованием систем компьютерной алгебры (CAS).

В работе [15] предложен комплексный фреймворк, интегрирующий компьютерное зрение (для детектирования и распознавания символов) и лингвистический анализ (для классификации). Конкретными результатами на публичном датасете стали точность детектирования формул в 96.7% и точность распознавания их структуры в 89.2%. Этот фреймворк решает задачу нижнего уровня – перевод формулы из изображения в структурированный формат, что является необходимым этапом обработки математических текстов.

В [16] выполнена масштабная систематизация всей области Mathematical Language Processing (MLP). Авторы проанализировали свыше 200 работ, пред-

ложив единую классификацию по задачам (распознавание, разбор, поиск, генерация), методам и наборам данных.

Исследование [17] посвящено созданию предобученной языковой модели MathBERT, специально адаптированной для математики. Для обучения представлениям формул выполнялись две задачи предварительного обучения: языковое моделирование маскированных токенов (masked language modeling, MLM) и предсказание соответствия контекста (context correspondence prediction, CCP). В качестве входных данных были выбраны деревья операторов (operator trees) – иерархические структуры, представляющие математической формулы. Для обучения был подготовлен большой набор данных, содержащий более 8.7 млн пар «формула – контекст», извлеченных из научных статей, размещенных на arXiv.org1. Модель была оценена по трем задачам, включая поиск математической информации, классификацию тем формул и генерацию заголовков формул. Экспериментальные результаты показали, что MathBERT значительно превосходит существующие методы по всем трем задачам.

В работе [18] решена критическая проблема неоднозначности. Авторы разработали метод семантификации идентификаторов, который анализирует контекст для определения смысла символа (например, что означает "f" в конкретной статье). Значимым результатом стало достижение точности в 88.5% на задаче разрешения неоднозначности.

В [19] предложена модель, которая строит совместные векторные представления (эмбеддинги) для формулы и ее текстового контекста. Получено улучшение точности поиска релевантных формул на 12%.

Обзор [20] сфокусирован на применении LLM в информационном поиске. Авторы проанализировали 85 работ, систематизировав паттерны использования LLM (например, переформулирование запросов, реранжирование) и оценили прирост эффективности. Сделан вывод о возможном увеличении метрики nDCG до 15% при грамотном использовании LLM.

Таким образом, можно заключить, что существующие работы детально решают отдельные подзадачи: высокоточное извлечение и аннотирование всех формул, создание эффективных алгоритмов поиска по формулам, разрешение неоднозначности или построение векторных моделей. Однако задача целенаправленного автоматического выделения именно главных формул статей, ос-

нованного на понимании ее основных результатов и их иерархической связи с текстом, по-прежнему остается нерешенной.

Предлагаемый нами подход, направлен на решение указанной задачи. Мы используем большую языковую модель DeepSeek не как узкоспециализированный инструмент для одной операции, а как универсальное ядро для сквозного семантического анализа, которое последовательно выстраивает связи от аннотации к конкретным формальным утверждениям. Экспериментальная оценка на коллекции из 50 статей с использованием метрик точности, полноты и F-меры подтвердила эффективность такого подхода для создания структурированных баз знаний, пригодных для интеграции в современные системы семантического поиска и анализа научной литературы.

2. ОПИСАНИЕ РАЗРАБОТАННОГО МЕТОДА

Целью проведенного исследования стала разработка метода автоматизированного извлечения основных математических результатов и формул из текстов научных публикаций с применением современных больших языковых моделей, в частности модели DeepSeek. Практическим результатом применения разработанного метода является подготовка базы данных, содержащей метаданные математических статей. Метаданные включают основные результаты, главные формулы, связанные с основными результатами, фрагменты текста с главными формулами с указанием логических единиц (леммы, теоремы и др.) и страниц документа, а также аннотацию статьи.

Рис. 1. Схема разработанного решения.

Решение построено на последовательном вызове трех разработанных скриптов для выполнения задач каждого из выделенных этапов – построения выборки статей, извлечения метаданных и выделения основных результатов и соответствующих формул из аннотаций статей. Схема разработанного решения,

включающего анализ математических статей и формирование базы данных, представлена на рис. 1.

На первом этапе формируется выборка научных статей на основе первого разработанного скрипта. Выборка содержит статьи в формате Portable Document Format (PDF). Скрипт переходит на портал Math-Net.ru и по заданным параметрам фильтра (рис. 2) осуществляет поиск и извлечение pdf-файлов научных статей. Извлеченные файлы записываются в папку «Научные статьи».

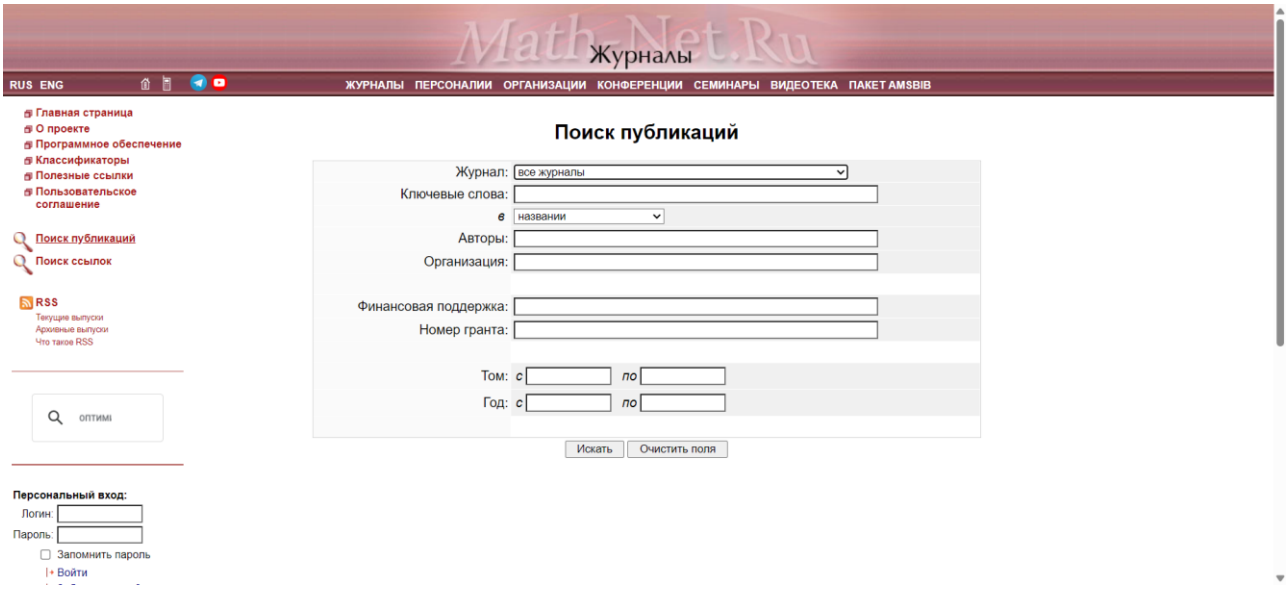

Рис. 2. Фильтры поиска статей на портале Math-Net.ru.

Вторым этапом является извлечение метаданных статей с помощью второго скрипта. Для каждого файла из папки «Научные статьи» скрипт формирует запрос к модели DeepSeek R1 по протоколу API (Application Programming Interface). Модель DeepSeek R1 обрабатывает запрос и передает ответ на вход скрипту по протоколу API. Скрипт выполняет парсинг ответа и записывает результаты в таблицу «Статья».

На третьем этапе для каждого файла из таблицы «Статья» третий скрипт извлекает содержимое поля «Аннотация» и передает это содержимое по протоколу API на вход модели DeepSeek R1. Модель DeepSeek R1 обрабатывает запрос, анализирует содержимое поля «Аннотация» и передает ответ на вход скрипту по протоколу API. Скрипт выполняет парсинг ответа и записывает его в таблицы «Результат» и «Формула».

Все этапы метода выполняются автоматически с предварительной ручной настройкой некоторых параметров, таких как указание фильтров отбора статей на первом этапе, настройка даты начала работы и периодичности запуска скрипта.

2.1. ПОДГОТОВКА ДАННЫХ ДЛЯ ИССЛЕДОВАНИЯ

В качестве данных для исследования были отобраны 50 научных статей с портала Math-Net.ru (https://www.mathnet.ru/) из различных разделов математики: это дифференциальные уравнения, математический анализ, теория вероятностей и математическая статистика

Для проводимого исследования была выбрана большая языковая модель DeepSeek из-за своей доступности на территории Российской Федерации, а также из-за высокой релевантности ответов, выдаваемых на запросы, вводимые пользователем, по сравнению с другими языковыми моделями.

2.2. ИЗВЛЕЧЕНИЕ ОСНОВНЫХ РЕЗУЛЬТАТОВ СТАТЬИ ИЗ АННОТАЦИИ

Источником основных результатов, представленных в статье, является ее аннотация. Она присоединяется в виде текста либо файла формата PDF к запросу большой языковой модели.

Далее во всех приводимых примерах взяты результаты из области обыкновенных дифференциальных уравнений, полученные в статье [21].

В рамках формирования запросов к LLM был применен подход к формулированию запроса, известный как zero-shot (запрос «с нулевыми примерами»). Суть этого подхода заключается в том, что большая языковая модель получает задачу, сформулированную на естественном языке, без какого-либо предварительного обучения на размеченных данных, специфичных для этой задачи, и без предоставления примеров ожидаемого выполнения непосредственно в теле запроса. Модель должна решить поставленную задачу, опираясь исключительно на обобщенные знания и языковые паттерны, приобретенные в процессе ее масштабного предобучения на разнородных текстовых корпусах.

Пример запроса к LLM DeepSeek на выделение главных результатов из аннотации, к указанной выше статьи, представлен в листинге 1.

Листинг 1. Промт 1. Выделение главных результатов из аннотации.

Дана следующая аннотация:

Получена формальная классификация типичных локальных перестроек неявного обыкновенного дифференциального уравнения в его особых точках при изменении одного внешнего параметра. Она состоит из четырех нормальных форм, каждая из которых содержит функциональный инвариант. Доказано, что любую деформацию квадратичного по производнойростка уравнения в классе контактной эквивалентности, при кото рой он остается квадратичным по производной, можно получить с помощью деформации независимой и зависимой переменных. На обобщении этого результата для семейств уравнений основана наша классификация. В качестве ее приложения получена формальная классификация типичных локальных перестроек на плоскости линейного уравнения смешанного типа с частными производными второго порядка в точках, где области эллиптичности и гиперболичности испытывают морсовские перестройки.

Выделить из данной аннотации основные результаты, полученные в ходе исследования.

Главные результаты, выделенные из аннотации статьи с помощью модели DeepSeek, приведены в листинге 2.

Модель DeepSeek верно выделила результаты из аннотации загруженной статьи, а также дала подробное описание каждого результата.

Листинг 2. Ответ модели на промпт 1.

Основные результаты исследования, выделенные из аннотации:

1. **Формальная классификация типичных локальных перестроек** неявного обыкновенного дифференциального уравнения (ОДУ) в его особых точках при изменении одного внешнего параметра. Классификация включает:
    - Четыре нормальные формы, каждая из которых содержит функциональный инвариант.
2. **Доказательство возможности сведения деформаций** квадратичного по производнойростка уравнения к деформациям независимой и зависимой переменных в классе контактной эквивалентности.
3. **Обобщение этого результата** на семейства уравнений, что легло в основу предложенной классификации.
4. **Применение классификации** к формальному описанию типичных локальных перестроек на плоскости для линейного уравнения смешанного типа (с частными производными второго порядка) в точках, где области эллиптичности и гиперболичности претерпевают морсовские перестройки.

Таким образом, ключевые достижения работы включают:

- Классификацию перестроек неявных ОДУ,

- Теоретическое обоснование сведения деформаций,
- Приложение результатов к уравнениям смешанного типа.

2.3. ЛОКАЛИЗАЦИЯ ФРАГМЕНТОВ ТЕКСТА С ОСНОВНЫМИ РЕЗУЛЬТАТАМИ

На следующем этапе решалась задача локализации фрагментов текста, содержащих основные результаты и связанные с ними формулы, посредством построения запросов к LLM.

Были сформулированы следующие требования к запросу к LLM. В качестве формата формул, выводимых в ответе, был выбран формат LaTeX. Требовалось также дать описание каждой переменной в выделенной формуле. При этом, если описание переменной не дано явно в тексте, LLM должна была найти описание в другом источнике с указанием этого источника. Кроме того, требовалось в ответе на запрос вывести ссылку в формате «страница статьи, на которой находится формула → тип логического элемента с номером (например, лемма, теорема и др.)». Текст запроса к LLM для выделения фрагментов текста и основных формул представлен в листинге 3.

Листинг 3. Пример промпта 2. Фрагмент текста и основные формулы

```
                                                                Статья 49.pdf

Выделить из текста статьи фрагменты текста, содержащие элементы выделенных из аннотации основных результатов. Фрагменты текста должны быть взяты из текста статьи. Необходимо выделить фрагменты текста по каждому полученному из аннотации результату. Выделить в найденных фрагментах математические формулы, связанные с основными результатами по списку результатов. В одном фрагменте может быть несколько основных формул. Формулы и переменные должны быть записаны в формате LaTeX, например, $z_k$. Дать описание каждой переменной, указанной в формуле, в формате LaTeX. Описание должно быть дано снизу формул и содержать слово "где". Если описание переменной формулы не содержится в тексте статьи, дать ссылку на внешний источник, откуда было взято описание. Ссылка на внешний источник должна располагаться после описания переменной и содержать раздел и номер страницы источника, откуда было взято описание. Также под каждой формулой указать ссылку в формате "Страница статьи, на которой находится формула --> Тип утверждения с номером (Например, лемма, теорема и т.п.)". Отчет страниц в статье начинать с первой страницы файла pdf. Вот пример, как должны быть представлены формулы в формате LaTeX:

\begin{equation}

-\Delta u(x) - 2(au_{x_1}(x) + bu_{x_1}(-x/p))_{x_2} = f(x)
```

```
\end{equation}
\begin{equation}
\xi_1^2 + \xi_2^2 + 2\xi_1\xi_2 \text{Re}(a \pm b\lambda) > 0
\end{equation}
```

Ответ модели DeepSeek на введенный запрос показан на листингах 4–7.

Листинг 4. Ответ модели на промпт 2 для основного результата 1 из аннотации (см. листинг 2): Фрагмент текста и основные формулы.

Основные результаты и соответствующие фрагменты текста статьи

1. Формальная классификация типичных локальных перестроек неявного ОДУ

Фрагмент текста:

"Получена формальная классификация типичных локальных перестроек неявного обыкновенного дифференциального уравнения в его особых точках при изменении одного внешнего параметра. Она состоит из четырех нормальных форм, каждая из которых содержит функциональный инвариант."

Математические формулы:

```
\begin{equation}
\pm(y')^2 \pm x^2 + z^2 = \varepsilon + c(\varepsilon)z^3, \quad \text{где } z
= y - \frac{1}{2}xy'
\end{equation}
```

где:

- y' – производная зависимой переменной y по независимой переменной x,
- x, y – независимая и зависимая переменные,
- ε – внешний параметр,
- $c(\varepsilon)$ – функциональный инвариант.

Ссылка:

Страница 8 --> Теорема 1

Листинг 5. Ответ модели на промпт 2 для основного результата 2 из аннотации (см. листинг 2): фрагмент текста и основные формулы.

2. Доказательство возможности сведения деформаций квадратичного ростка

Фрагмент текста:

"Доказано, что любую деформацию квадратичного по производной-ростка уравнения в классе контактной эквивалентности, при которой он остается квадратичным по производной, можно получить с помощью деформации независимой и зависимой переменных."

Математические формулы:

```
\begin{equation}
f(x, y, p) = 0, \quad \text{где } f(a_0) = \partial_p f(a_0) = 0 \text{ и }
\partial_p^2 f(a_0) \neq 0
\end{equation}
```

где:

- f – функция, задающая неявное ОДУ,
- a_0 – точка приложения ростка,
- p – производная dy/dx.

Ссылка:

Страница 9 --> Определение 3

Листинг 6. Ответ модели на промпт 2 для основного результата 3 из аннотации (см. листинг 2): Фрагмент текста и основные формулы.

3. Обобщение результата для семейств уравнений

Фрагмент текста:

"На обобщении этого результата для семейств уравнений основана наша классификация."

Математические формулы:

```
\begin{equation}
F(x, y, p, \varepsilon) = 0, \quad \text{где } F(A_0) = \partial_p F(A_0) = 0
\text{ и } \partial_p^2 F(A_0) \neq 0
\end{equation}
```

где:

- F – функция, задающая семейство неявных ОДУ,
- A_0 – точка приложения ростка,
- ε – параметр семейства.

Ссылка:

Страница 12 --> Теорема 4

Листинг 7. Ответ модели на промпт 2 для основного результата 4 из аннотации (см. листинг 2): Фрагмент текста и основные формулы.

4. Применение классификации к уравнениям смешанного типа

Фрагмент текста:

"В качестве ее приложения получена формальная классификация типичных локальных перестроек на плоскости линейного уравнения смешанного типа с частными производными второго порядка в точках, где области эллиптичности и гиперболичности испытывают морсовские перестройки."

Математические формулы:

```
\begin{equation}
\alpha(x, y, \epsilon)u_{xx} + 2\beta(x, y, \epsilon)u_{xy} + \gamma(x, y,
\epsilon)u_{yy} = \sigma(x, y, \epsilon, u, u_x, u_y)
```

```
\end{equation}
```
где:

- α, β, γ – коэффициенты уравнения,
- ϵ – параметр,
- u_{xx}, u_{xy}, u_{yy} – частные производные функции u.

Ссылка:
```
Страница 16 --> Теорема 5
```

Анализ ответов (листинги 4–7) показывает, что модель DeepSeek корректно выделяет фрагменты текста и основные формулы по каждому результату из аннотации. LLM также корректно представляет формулы в формате LaTeX и для каждой формулы указывает ссылку на страницу и логическую единицу, в которой дана формула.

3. ОЦЕНКА ЭФФЕКТИВНОСТИ ИЗВЛЕЧЕНИЯ ОСНОВНЫХ РЕЗУЛЬТАТОВ И ФОРМУЛ

На данных тестовой коллекции из 50 статей была проведена оценка эффективности извлечения основных результатов и формул с помощью модели DeepSeek. Сравнивалось количество извлеченных основных результатов и формул статей на основе экспертной оценки и средствами модели DeepSeek по метрикам точности, полноты и F-меры. Результаты приведены на рис. 3 и табл. 1.

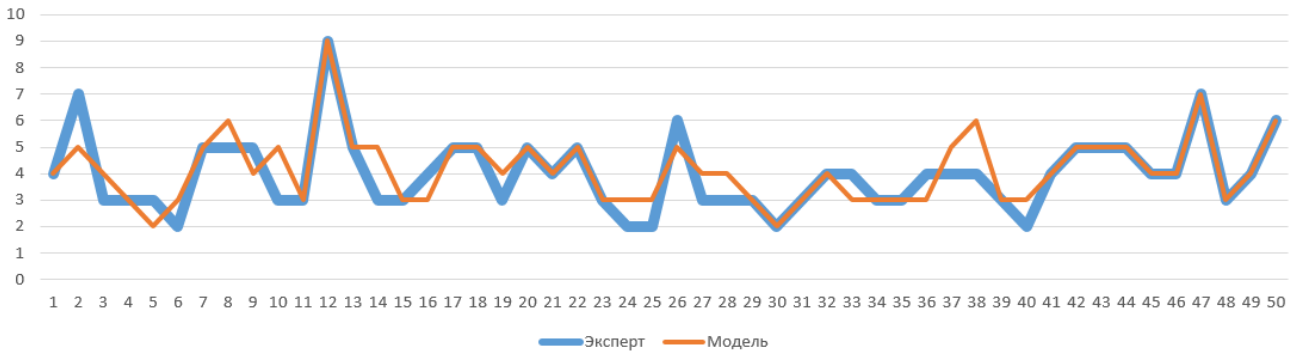

Рис. 3. Количество выделенных основных результатов.

На рис. 3 представлены данные о количестве выделенных основных результатов экспертом и моделью. На оси абсцисс отложены порядковые номера статей, отобранных для исследования, на оси ординат количество выделенных результатов. Синей линией на графике показано количество результатов, выде-

ленных экспертом, оранжевой линией – количество результатов, выделенных моделью. Из рисунка видно, что на некоторых интервалах количества результатов, выделенных моделью и экспертом, совпадают, например самый длинный интервал совпадения от 41 до 50 по оси абсцисс.

Точность *P* рассчитывается по следующей формуле:

$$P = \frac{N_{\text{с}}}{N_{\text{м}}},$$

где $N_{\text{с}}$ – количество результатов, выделенных моделью и совпадающих с результатами, выделенными экспертом; $N_{\text{м}}$ – количество результатов, выделенных моделью.

Полнота *R* рассчитана по формуле

$$R = \frac{N_{\text{с}}}{N_e},$$

где N_{e} – количество результатов, выделенных экспертом.

Для подсчета F-меры используется формула:

$$F = 2\frac{P * R}{P + R}.$$

Табл. 1. Оценка выделения основных результатов.

	Точность (сред)	Точность (max)	Точность (min)	Полнота (сред)	Полнота (max)	Полнота (min)	F-мера (сред)
Датасет из 50 статей	0.9233	1	0.6	0.9653	1	0.6667	0.9346

В табл. 1 представлена оценка выделения основных результатов. Наихудшее значение точности, равное 0.6, было получено для статьи с номером 10 [22]. Например, в этой статье эксперт выделил следующие результаты:

- построение дифференциальных уравнений равновесия в перемещениях;
- разработка биквадратичных физических зависимостей;
- анализ шести основных случаев физических зависимостей.

Модель смогла выделить лишь следующие результаты:

- построение дифференциальных уравнений равновесия в перемещениях;

- биквадратичная аппроксимация замыкающих уравнений;
- разработка биквадратичных физических зависимостей;
- предположение о независимости диаграмм объемного и сдвигового деформирования в общем случае;
- анализ шести основных случаев физических зависимостей.

Результаты «Биквадратичная аппроксимация замыкающих уравнений» и «Предположение о независимости диаграмм объемного и сдвигового деформирования в общем случае», дополнительно выделенные моделью, снижают значение точности на 0.4.

При выделении моделью дополнительных результатов требуется повторная экспертиза для принятия окончательного решения: относятся ли данные результаты к основным результатам научной статьи или нет.

Если модель выделяет меньше результатов, чем эксперт, то возможно, что такие результаты могут не относиться к основным. В этом случае точность понижается, но также требуется повторная экспертиза для принятия окончательного решения. Однако в настоящем исследовании такая ситуация не встретилась.

Наихудшее значение полноты, равное 0.67, было получено для статьи под номером 5 [23]. Например, в этой статье эксперт выделил следующие результаты:

- обзор недавних результатов в области нелинейных интегро-дифференциальных уравнений типа свертки;
- обзор недавних результатов в области сингулярных интегро-дифференциальных уравнений с ядрами Гильберта и Коши;
- для суммируемых решений произвольного знака применен метод максимальных монотонных операторов (по Браудеру – Минти).

Модель выделила только следующие результаты:

- обзор недавних результатов в области сингулярных интегро-дифференциальных уравнений с ядрами Гильберта и Коши;
- для суммируемых решений произвольного знака применен метод максимальных монотонных операторов (по Браудеру – Минти).

Модели не удалось выделить важный результат научной работы «Обзор недавних результатов в области нелинейных интегро-дифференциальных уравнений типа свертки», что понизило значение полноты на 0.33.

Основной результат, не выделенный моделью, также требует повторной экспертизы для принятия окончательного решения.

Наихудшее значение F-меры, равное 0.75, на графике было получено для статьи под номером 10.

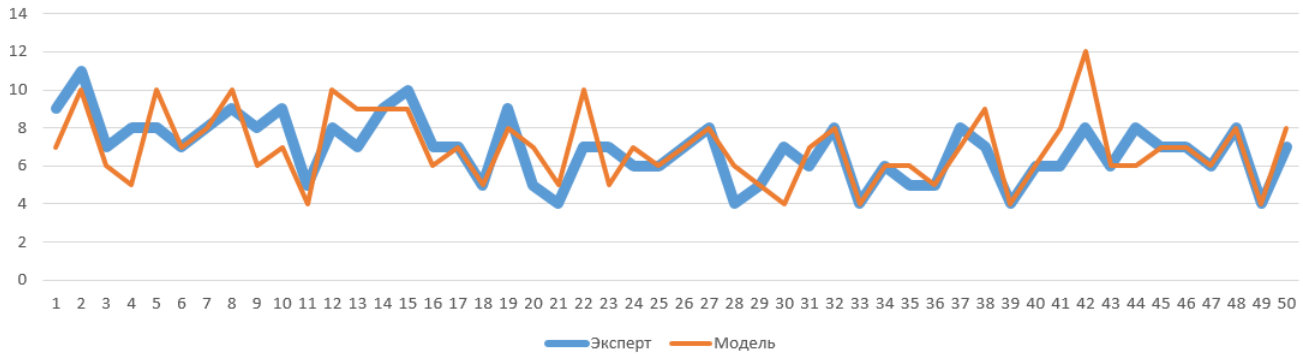

Рис. 4. Количество выделенных основных формул.

На рис. 4 представлены результаты по количеству основных формул, выделенных экспертом и моделью. На оси абсцисс отложены порядковые номера статей, отобранных для исследования, на оси ординат количество выделенных формул. Синей линией на графике обозначено количество формул, выделенных экспертом, оранжевой линией на графике количество формул, выделенных моделью. Из рисунка видно, что на некоторых интервалах количества формул, выделенных моделью и экспертом, совпадают, например самый длинный интервал совпадения от 45 до 49 по оси абсцисс.

Точность *P* определяется по формуле

$$P = \frac{N_c}{N_м},$$

где N_c – количество формул, выделенных моделью и совпадающих с формулами, выделенными экспертом; $N_м$ – количество формул, выделенных моделью.

Полнота *R* устанавливается по формуле

$$R = \frac{N_c}{N_e},$$

где N_e – количество формул, выделенных экспертом.

Для подсчета F-меры используется формула

$$F = 2\frac{P*R}{P+R}.$$

Табл. 2. Оценка выделения основных формул.

	Точность (сред)	Точность (max)	Точность (min)	Полнота (сред)	Полнота (max)	Полнота (min)	F-мера (сред)
Датасет из 50 статей	0.9355	1	0.6667	0.9411	1	0.5714	0.9294

В табл. 2 представлена оценка выделения основных формул. Наихудшее значение точности, равное 0.67, было получено для статьи под номером 28 [24]. Например, в этой статье эксперт выделил следующие формулы:

- основное функционально-дифференциальное уравнение

$$\dot{x}\big(g(t)\big) = f\Big(t, x\big(h(t)\big)\Big),\ t \in [0,1];$$

- интегральное представление задачи Коши

$$y\big(g(t)\big) = f\Big(t, \alpha + \int_0^1 \chi_{[0,h(t)](s)y(s)ds}\Big),\ t \in [0,1];$$

- пример уравнений с оценкой решения (2 формулы)

$$\dot{x}(t^2) = \frac{x^2(t)}{t^2} + \frac{\lambda}{\sqrt{t}}, t \in [0,1];$$

$$\frac{k}{2t} \geq \frac{k^2}{t} + \frac{\lambda}{\sqrt{t}}.$$

Модель выделила следующие формулы:

- основное функционально-дифференциальное уравнение

$$\dot{x}\big(g(t)\big) = f\Big(t, x\big(h(t)\big)\Big),\ t \in [0,1],$$

- интегральное представление задачи Коши

$$y\big(g(t)\big) = f\left(t, \alpha + \int_0^1 \chi_{[0,h(t)](s)y(s)ds}\right), \qquad t \in [0,1];$$

- условие для функций-барьеров

$$\dot{V}\big(g(t)\big) \geq f\Big(t, v\big(h(t)\big)\Big), \qquad \dot{w}\big(g(t)\big) \leq f\Big(t, w\big(h(t)\big)\Big);$$

- пример уравнений с оценкой решения (2 формулы)

$$\dot{x}(t^2) = \frac{x^2(t)}{t^2} + \frac{\lambda}{\sqrt{t}}, t \in [0,1] ;$$

$$\frac{k}{2t} \geq \frac{k^2}{t} + \frac{\lambda}{\sqrt{t}};$$

Операторное уравнение

$$(\psi y)(t) := y\big(g(t)\big), \qquad (\varphi y)(t) := f\left(t, \alpha + \int_0^1 \chi_{[0,h(t)](s)} y(s) ds\right).$$

Дополнительно выделенные моделью формулы «Условие для функций-барьеров» и «Операторное уравнение» дают снижение значения точности на 0.33.

Дополнительно выделенные формулы требуют повторной экспертизы для принятия окончательного решения: относятся ли данные формулы к основным формулам научной статьи или нет.

Если модель выделила основных формул меньше, чем эксперт, то точность уменьшается, однако в настоящем исследовании модель выделяла основных формул больше, чем было определено экспертом.

Наихудшее значение полноты, равное 0.57, было получено для статьи под номером 4 [25]. Например, в этой статье эксперт выделил следующие основные формулы:

- интегро-дифференциальные уравнения Вольтера (2 формулы)

$$g(t, \phi) = \int_{-t}^{0} K(t,s) h\big((P_0\phi)(s)\big) ds , \qquad x'(t) = \int_0^t k(t,s) h\big(x(s)\big) ds ;$$

- уравнение дискретного запаздывания

$$x'(t) = ax(\lambda t) + bx(t) ;$$

- пространство Фреше непрерывных отображений

$$C = C\big((-\infty, 0], R^n\big) ;$$

- вариационное уравнение;

$$(D\,\Sigma_t(\varphi)\hat{\varphi})(s) = \hat{\varphi}(t+s), \qquad \text{если } t+s \leq 0 ,$$
$$(D\,\Sigma_t(\varphi)\hat{\varphi})(s) = (D\,\Sigma_{t+s}(\varphi)\hat{\varphi})(0), \qquad \text{если } 0 \leq t+s;$$

- полупото́к для автономных уравнений с запаздыванием

$$\Sigma(t, \phi) = x_t^{\phi};$$

- непрерывный процесс для неавтономных уравнений (2 формулы)

$$x'(t) = g(t, x_t), \qquad P(t, t_0, \phi) = p_n \Sigma_g(t - t_0, t_{0*}, \phi).$$

Модель смогла выделить следующие формулы:

- интегро-дифференциальные уравнения Вольтера (2 формулы)

$$g(t, \phi) = \int_{-t}^{0} K(t, s) h\big((P_0\phi)(s)\big) ds, \qquad x'(t) = \int_0^t k(t, s) h\big(x(s)\big) ds;$$

- полупото́к для автономных уравнений с запаздыванием

$$\Sigma(t, \phi) = x_t^{\phi};$$

- непрерывный процесс для неавтономных уравнений (2 формулы)

$$x'(t) = g(t, x_t), \qquad P(t, t_0, \phi) = p_n \Sigma_g(t - t_0, t_{0*}, \phi).$$

Модели не удалось выделить следующие важные формулы этой научной статьи: «Уравнение дискретного запаздывания», «пространство Фреше непрерывных отображений» и «Вариационное уравнение», что понизило значение полноты на 0.4.

Основные формулы, не выделенные моделью, требуют повторной экспертизы для принятия окончательного решения: относятся ли эти формулы к основным формулам научной статьи или нет.

Наихудшее значение F-меры, равное 0.73, было получено для статьи под номером 30.

Таким образом, согласно полученным результатам заключаем, что модель DeepSeek выдает релевантные ответы по главным результатам статей и связанным с ними формулам. Метрики точности, полноты и F-меры являются достаточно высокими, средние оценки F-меры по коллекции для основных результатов и связанных с ними формул равны 0.93459 и 0.9294 соответственно.

4. КОНЦЕПЦИЯ БАЗЫ ДАННЫХ

Основные результаты и формулы, выделенные из коллекции статей, сохраняются в базе данных. Предложена следующая структура базы данных, со-

держащая таблицы статей, основных результатов, фрагментов текста основных результатов, формул и переменных формул:

- статья (атрибуты: номер (int), название (varchar (500)) авторы (varchar (500)), год издания (year), наименование журнала (varchar (500)), УДК (varchar (50)), аннотация(text));
- результат (атрибуты: id (int), номер статьи (int), номер результата (int) наименование (varchar (500)), фрагмент текста (text), концепт (varchar(255)));
- формула (атрибуты: id (int), номер статьи (int), номер результата (int), номер формулы (int), наименование (varchar (500)), latex формула(text), описание переменных (text), концепт(varchar (255))).

Связи таблиц в базе данных задаются следующим образом:

- статья содержит один или множество результатов. Результат может быть отнесен только к одной статье;
- результат содержит одну или множество формул. Формула может быть отнесена только к одному результату.

ЗАКЛЮЧЕНИЕ

Предложен подход к выделению основных формул с применением современных больших языковых моделей. С помощью модели DeepSeek были проведены эксперименты по установлению основных результатов научных статей на основе аннотаций с последующей локализацией результатов в тексте статьи. В выделенных фрагментах из основного текста статей были извлечены и основные формулы, описывающие эти результаты.

Для оценки релевантности ответов LLM использованы метрики точности, полноты и F-меры. Результаты исследования показывают, что современные языковые модели, подобные DeepSeek, успешно справляются с поставленными задачами. Так, средние значения F-меры по тестовой коллекции равны соответственно 0.93459 для оценки эффективности извлечения основных результатов и 0.9294 для оценки эффективности извлечения основных формул.

Сформированная база данных аннотированных основных формул математических статей является новым математическим артефактом, который может быть встроен в систему математического поиска по коллекции научных ста-

тей. Предполагается, что данный поисковик будет составной частью исследовательской инфраструктуры в области математики, которая разрабатывается в проекте Казанского федерального университета.

Благодарности

Работа выполнена за счет предоставленного в 2026 году Фондом науки и технологий Республики Татарстан гранта на осуществление фундаментальных и поисковых исследований в научных и образовательных организациях, предприятиях и организациях реального сектора экономики Республики Татарстан (соглашение № 2 от 05.08.2026).

СПИСОК ЛИТЕРАТУРЫ

1. *Невзорова О.А., Николаев К.С.* Семантическое аннотирование математических формул в PDF-документах // Электронные библиотеки. 2022. Т. 25 (6). С. 616–639. https://doi.org/10.26907/1562-5419-2022-25-6-616-639

2. *Невзорова О.А., Липачев Е.К., Николаев К.С.* Семантические сервисы цифровой экосистемы OntoMath для математического образования // Электронные библиотеки. 2023. Т. 26, № 4. С. 538–564. https://doi.org/10.26907/1562-5419-2023-26-4-538-569

3. *Nikolaev K., Nevzorova O*. Annotation of mathematical formulas in pdf documents // Информационные технологии и нанотехнологии (ИТНТ–2023): сборник трудов по материалам IX Международной конференции и молодежной школы (г. Самара, 17–23 апреля 2023 г.): в 6 томах. – Том 5. Науки о данных. Самара: Издательство Самарского университета, 2023. С. 54–55.

4. *Николаев К.С.* Сервис семантического поиска формул по коллекции математических PDF-документов // Информационные технологии и вычислительные системы. 2025. № 3. С. 34–43.

5. *Ohashi S. et al.* Efficient algorithm for math formula semantic search // IEICE TRANSACTIONS on Information and Systems. 2016. Vol. 99, No. 4. P. 979–988. https://doi.org/10.1587/transinf.2015DAP0023

6. *Dhar S., Roy S., Das S.K.* A critical survey of mathematical search engines // International Conference on Computational Intelligence, Communications, and Business Analytics. Singapore: Springer Singapore, 2018. P. 193–207.

7. *Scharpf P. et al.* Mathematics in Wikidata // In: Proceedings of the 2nd Wikidata Workshop (Wikidata 2021) co-located with the 20th International Semantic Web Conference (ISWC 2021). CEUR Workshop Proceedings. 2021. https://doi.org/10.5281/zenodo.5589640

8. *Collard J. et al.* Extracting Mathematical Concepts from Text // arXiv. 2022. https://doi.org/10.48550/arXiv.2208.13830

9. *Greiner-Petter A., Youssef A., Ruas T. et al.* Math-word embedding in math search and semantic extraction // Scientometrics 125. 2020. P. 3017–3046. https://doi.org/10.1007/s11192-020-03502-9

10. *Yang M. et al.* A Survey of Information Extraction Based on Deep Learning // School of Management, China University of Mining and Technology (Beijing), Beijing 100083, China 2022. https://doi.org/10.3390/app12199691

11. *Pathak A., Pakray P. et al.* MathIRs: Retrieval System for Scientific Documents // Computación y Sistemas. 2017. Vol. 21, No. 2. P. 253–265. https://doi.org/10.13053/CyS-21-2-2743

12. *Pathak A. et al.* A Formula Embedding Approach to Math Information Retrieval // Computación y Sistemas. 2018. Vol. 22, No. 3. P. 819–833. https://doi.org/10.13053/CyS-22-3-3015

13. *Davila K., Zanibbi R.* Layout and semantics: Combining representations for mathematical formula search // Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2017. P. 1165–1168. https://doi.org/10.1145/3077136.3080748

14. *Gipp B. et al.* Methods and Tools to Advance the Retrieval of Mathematical Knowledge from Digital Libraries for Search-, Recommendation-, and Assistance-Systems // arXiv. 2023. https://doi.org/10.48550/arXiv.2305.07335

15. *Shah A.K., Dey A., Zanibbi R.* A Math Formula Extraction and Evaluation Framework for PDF Documents // Lecture Notes in Computer Science. Vol. 12822. Springer 2021. P. 19–34. https://doi.org/10.1007/978-3-030-86331-9_2

16. *Meadows J. et al.* A Survey in Mathematical Language Processing // arXiv. 2022. https://doi.org/10.48550/arXiv.2205.15231

17. *Gao L. et al.* Pre-trained models for mathematical formula understanding // Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2022. P. 3124–3132. https://doi.org/10.48550/arXiv.2105.00377

18. *Schubotz M. et al.* Semantification of Identifiers in Mathematics for Better Math Information Retrieval // SIGIR '16: Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. 2016. P. 135–144. https://doi.org/10.1145/2911451.2911503

19. *Pankaj D. et al.* Embedding and generalization of formula with context in the retrieval of mathematical information // Journal of King Saud University. 2022. P. 6624–6634. https://doi.org/10.1016/j.jksuci.2021.05.014

20. *Zhu Y. et al.* Large Language Models for Information Retrieval: A Survey // Arxiv.org. 2023. https://doi.org/10.48550/arXiv.2308.07107

21. *Богаевский И.А.* Неявные обыкновенные дифференциальные уравнения: перестройки и усиление эквивалентности // Известия РАН. Серия математическая. 2014. Т. 78 (6). С. 5–20.

22. *Бакушев С.В.* Дифференциальные уравнения равновесия сплошной среды для плоской деформации в декартовых координатах при биквадратичной аппроксимации замыкающих уравнений // Вестн. Томск. гос. ун-та. Матем. и мех. 2022. № 76. С. 70–86.

23. *Асхабов С.Н.* Нелинейные интегро-дифференциальные уравнения с разностными ядрами // Итоги науки и техн. Соврем. мат. и ее прил. Темат. обз. 2021. Т. 198. С. 22–32

24. *Жуковская Т.В., Филиппова О.В., Шиндяпин А.И.* О распространении теоремы Чаплыгина на дифференциальные уравнения нейтрального типа // Вестник российских университетов. Математика. 2019. Т. 24 (127). С. 272–280.

25. *Вальтер Х.-О*. Дифференциальные уравнения с запаздыванием с дифференцируемыми операторами решений на открытых областях в C((−∞,0],Rn) и процессы для интегро-дифференциальных уравнений Вольтера // СМФН. 2021. Т. 67 (3). С. 483–506.

AN APPROACH TO EXTRACTING BASIC FORMULAS FROM SCIENTIFIC TEXTS BASED ON LARGE LANGUAGE MODELS

E. V. Saitov[1] [0009-0009-1980-8740], O. A. Nevzorova[2] [0000-0001-8116-9446]

[1, 2]*Kazan (Volga region) Federal University, Kazan, Russia*

[1]saitoverik@mail.ru, [2]onevzoro@gmail.com

Abstract

This article studies the problem of extracting formulas of scientific articles presented in pdf format, as well as converting the selected formulas from pdf to LaTeX format. The extraction of the main formulas of a mathematical article and their subsequent presentation in LaTeX format using the large language model DeepSeek is considered. Extraction of the main results of the papers bases on the paper annotations. Then, using a large language model we search fragments of papers describing the main results and highlighting the main formulas within the found fragments. The solution was evaluated using the metrics of precision, recall, and F-measure; the average values of the F-measure for the test collection was 0.93459 and 0.9294. The MySQL database management system is given as a method for storing the obtained data and the description of the database structure is given in text form: database tables, their attributes, as well as types of stored data

Keywords*: large language model, mathematical article, main formula, main formula, LaTeX format, extraction formula, semantic search, annotation formula, PDF document processing, information retrieval, artificial intelligence, scientometric analysis.*

REFERENCES

1. *Nevzorova O.A., Nikolaev K.S.* Cemanticheskoe annotirovanie matematicheskih formul v PDF-dokumentah // Russian Digital Library. 2022. Vol. 25 (6). P. 616–639. https://doi.org/10.26907/1562-5419-2022-25-6-616-639

2. *Nevzorova O.A., Lipachev E.K., Nikolaev K.S.* Cemanticheskie servisy cifrovoj ekosistemy ontomath dlya matematicheskogo obrazovaniya // Russian Digital Library. 2023. Vol. 26, No. 4. P. 538–564.
https://doi.org/10.26907/1562-5419-2023-26-4-538-569

3. *Nikolaev K., Nevzorova O.* Annotation of mathematical formulas in pdf documents // Informacionnye tekhnologii i nanotekhnologii (ITNT-2023): sbornik trudov po materialam IX Mezhdunarodnoj konferencii i molodezhnoj shkoly (g. Samara, 17–23 aprelya 2023 g.): v 6 tomah. – Tom 5. Nauki o dannyh. Samara: Izdatel'stvo Samarskogo universiteta, 2023. S. 54–55.

4. *Nikolaev K.S.* Servis semanticheskogo poiska formul po kollekcii matematicheskih PDF-dokumentov // Informacionnye tekhnologii i vychislitel'nye sistemy. 2025. No. 3. S. 34–43.

5. *Ohashi S. et al.* Efficient algorithm for math formula semantic search // IEICE TRANSACTIONS on Information and Systems. 2016. Vol. 99, No. 4. P. 979–988. https://doi.org/10.1587/transinf.2015DAP0023.

6. *Dhar S., Roy S., Das S.K.* A critical survey of mathematical search engines // International Conference on Computational Intelligence, Communications, and Business Analytics. Singapore: Springer Singapore, 2018. P. 193–207.

7. *Scharpf P. et al.* Mathematics in Wikidata // In: Proceedings of the 2nd Wikidata Workshop (Wikidata 2021) co-located with the 20th International Semantic Web Conference (ISWC 2021). CEUR Workshop Proceedings. 2021. https://doi.org/10.5281/zenodo.5589640

8. *Collard J. et al.* Extracting Mathematical Concepts from Text // arXiv. 2022. https://doi.org/10.48550/arXiv.2208.13830

9. *Greiner-Petter A., Youssef A., Ruas T. et al.* Math-word embedding in math search and semantic extraction/ Scientometrics. 2020. Vol. 125. P. 3017–3046. https://doi.org/10.1007/s11192-020-03502-9

10. *Yang M. et al.* A Survey of Information Extraction Based on Deep Learning // School of Management, China University of Mining and Technology (Beijing), Beijing 100083, China 2022. https://doi.org/10.3390/app12199691

11. *Pathak A., Pakray P. et al.* MathIRs: Retrieval System for Scientific Documents // Computación y Sistemas. 2017. Vol. 21, No. 2. P. 253–265. https://doi.org/10.13053/CyS-21-2-2743

12. *Pathak A. et al.* A Formula Embedding Approach to Math Information Retrieval // Computación y Sistemas. 2018. Vol. 22, No. 3. P. 819–833. https://doi.org/10.13053/CyS-22-3-3015

13. *Davila K., Zanibbi R.* Layout and semantics: Combining representations for mathematical formula search // Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2017. P. 1165–1168. https://doi.org/10.1145/3077136.3080748

14. *Gipp B. et al.* Methods and Tools to Advance the Retrieval of Mathematical Knowledge from Digital Libraries for Search-, Recommendation-, and Assistance-Systems // arXiv:2305.07335. https://doi.org/10.48550/arXiv.2305.07335

15. *Shah A.K., Dey A., Zanibbi R.* A Math Formula Extraction and Evaluation Framework for PDF Documents // Lecture Notes in Computer Science. Vol. 12822. Springer 2021. P. 19–34. https://doi.org/10.1007/978-3-030-86331-9_2

16. *Meadows J. et al.* A Survey in Mathematical Language Processing // arXiv. 2022. https://doi.org/10.48550/arXiv.2205.15231

17. *Gao L. et al.* Pre-trained models for mathematical formula understanding // Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2022. P. 3124-3132. https://doi.org/10.48550/arXiv.2105.00377

18. *Schubotz M. et al.* Semantification of Identifiers in Mathematics for Better Math Information Retrieva // SIGIR'16: Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. 2016. P. 135–144. https://doi.org/10.1145/2911451.2911503

19. *Pankaj D. et al.* Embedding and generalization of formula with context in the retrieval of mathematical information // Journal of King Saud University. 2022. P. 6624–6634. https://doi.org/10.1016/j.jksuci.2021.05.014

20. *Zhu Y. et al.* Large Language Models for Information Retrieval: A Survey // Arxiv.org. 2023. https://doi.org/10.48550/arXiv.2308.07107

21. *Bogaevskij I.A.* Neyavnye obyknovennye differencial'nye uravneniya: perestrojki i usilenie ekvivalentnosti // Izvestiya RAN. Seriya matematicheskaya. 2014. T. 78 (6). S. 5–20.

22. *Bakushev S.V.* Differencial'nye uravneniya ravnovesiya sploshnoj sredy dlya ploskoj deformacii v dekartovyh koordinatah pri bikvadratichnoj approksimacii zamykayushchih uravnenij // Vestn. Tomsk. gos. un-ta. Matem. i mekh. 2022. No. 76. S. 70–86.

23. *Askhabov S.N.* Nelinejnye integro-differencial'nye uravneniya s raznostnymi yadrami // Itogi nauki i tekhn. Sovrem. mat. i ee pril. Temat. obz. 2021. T. 198. S. 22–32

24. *Zhukovskaya T.V., Filippova O.V., Shindyapin A.I.* O rasprostranenii teoremy CHaplygina na differencial'nye uravneniya nejtral'nogo tipa // Vestnik rosijskih universitetov. Matematika. 2019. T. 24 (127). S. 272–280.

25. *Val'ter H.-O.* Differencial'nye uravneniya s zapazdyvaniem s differen ciruemymi operatorami reshenij na otkrytyh oblastyah v C((−∞,0],Rn) i processy dlya integrodifferencial'nyh uravnenij Vol'terra // SMFN. 2021. T. 67 (3). S. 483–506.

СВЕДЕНИЯ ОБ АВТОРАХ

САИТОВ Эрик Вадимович – магистрант Института вычислительной математики и информационных технологий Казанского федерального университета.

Erik Vadimovich SAITOV – Master's student at the Institute of Computational Mathematics and Information Technologies, Kazan Federal University.

email: saitoverik@mail.ru

ORCID: 0009-0009-1980-8740

НЕВЗОРОВА Ольга Авенировна — кандидат технических наук, доцент Казанского (Приволжского) федерального университета.

Olga Avenirovna NEVZOROVA — Candidate of Technical Sciences (PhD equivalent), Associate Professor, Kazan (Volga Region) Federal University.

ORCID: 0000-0001-8116-9446.

email: onevzoro@gmail.com

Материал поступил в редакцию 10 апреля 2026 года