

Анализ публикаций в области высокопроизводительных вычислений и моделирование факторов цитируемости

Акопян Микаэла Тиграновна

Научный руководитель: к.т.н. Григорьева Мария Александровна

Введение

Высокопроизводительные вычисления (HPC) — это область информатики, занимающаяся разработкой и применением суперкомпьютеров и параллельных алгоритмов для решения сложных задач. Быстрое развитие вычислительных технологий привело к резкому росту количества публикаций и исследовательских направлений в данной области.

В наши дни возникает необходимость в автоматизированном анализе публикационной активности, что позволяет систематизировать научные данные, выявлять тенденции развития и определять факторы, влияющие на цитируемость работ.

На основе таких данных возможно создание рекомендательных моделей, которые помогают исследователям повышать видимость и влияние своих публикаций.

Формальная постановка задачи работы

Входными данными являются метаданные публикаций в области НРС размещённых на платформе arXiv за определённый промежуток времени.

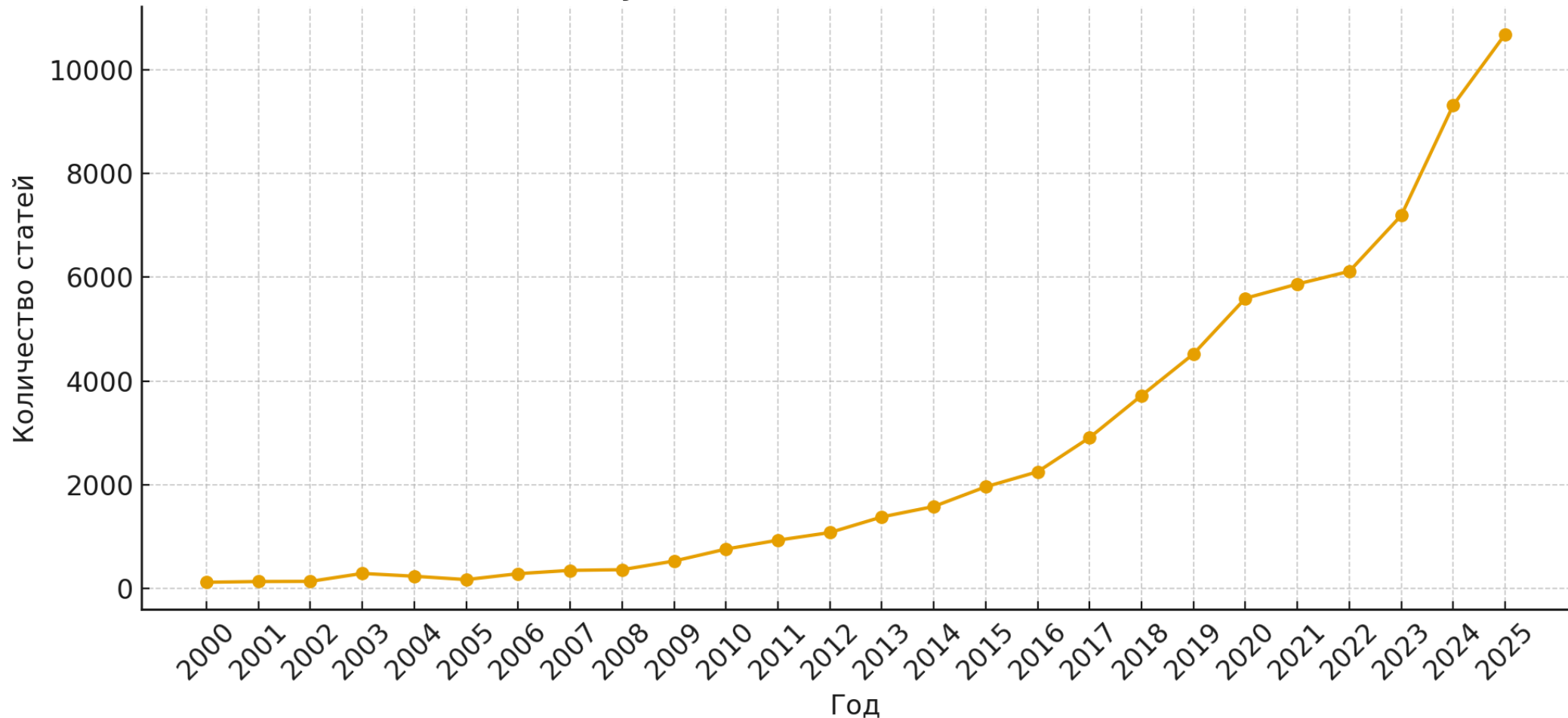
Для анализа цитируемости и факторов, влияющих на популярность, данные из arXiv дополняются информацией из открытых научных баз, таких как Semantic Scholar и OpenAlex, включающих количество цитирований, ссылки на другие публикации и аффилиации авторов.

Выходными данными являются статистические зависимости и параметры, которые будут использоваться в создании рекомендательной модели, помогающей исследователям выбирать наиболее эффективные стратегии для повышения цитируемости научных работ.

Результаты запроса на сайте arxiv.org

query = (
'all:("high performance computing" OR supercomputer OR "parallel computing") 'f"AND
submittedDate:[200001012359 TO 202512312359]"')

Публикации по годам (HPC)



С 2000 по 2025 год было опубликовано 73 997 статей с указанными ключевыми словами

Неудачный запрос ключа доступа Semantic Scholar

Dear Mikaela,

Thank you for your interest in a Semantic Scholar API key. After review, we are unable to approve your request at this time due to limited resources and our mission to maintain the Semantic Scholar API as an open-access research and educational tool in service of the public interest.

We appreciate your understanding and thank you for your interest in Semantic Scholar.

Thank you,
Semantic Scholar team

Следующие шаги

- Сбор данных о вышедших статьях из API Arxiv.org
- Интеграция с информацией о цитируемости и метаданных авторов с API OpenAlex
- Формирование признаков для анализа (тематика, количество соавторов, площадка, ключевые слова)
- Построение базовой модели прогноза цитируемости на основании модели из курсовой работы
- Разработка рекомендательной системы для исследователей в области НРС