

Университет ИТМО, факультет программной инженерии и компьютерной техники
Двухнедельная отчётная работа по «Информатике»: аннотация к статье

Дата прошедшей лекции	Номер прошедшей лекции	Название статьи/главы книги/видеолекции	Дата публикации (не старше 2022 года)	Размер статьи (от 400 слов)	Дата сдачи
10.09.2025	1	Non-Standard Binary Representations and the Stern Sequence	14.08.2023	~3,800	24.09.2025
24.09.2025	2	LLM-Guided Search for Deletion-Correcting Codes	01.04.2025	~10000	08.10.2025
08.10.2025	3	Intelligent Interface: Enhancing Lecture Engagement with Didactic Activity Summaries	20.06.2024	~7500	22.10.2025
22.10.2025	4	Advancing vision-language models in front-end development via data synthesis	03.03.2025	~12000	05.11.2025
05.11.2025	5	Spreadsheet Quality Assurance: A Literature Review	22.01.2024	~7,500	19.11.2025
19.11.2025	6	NeuRaLaTeX: A machine learning library written in pure LaTeX	31.05.2025	~8500	03.12.2025
	7				

Выполнил(а) Эллити Мохамед Эмад Ахмед Авад, № группы P3131, оценка
Фамилия И.О. студента не заполнять

Прямая полная ссылка на источник или сокращённая ссылка (bit.ly, tr.im и т.п.)

Прямая ссылка на источник: <https://arxiv.org/abs/2503.24187>

Теги, ключевые слова или словосочетания (минимум три слова):

TeX programming, LaTeX automation, Reproducible research, Self-contained papers, Neural networks in LaTeX

Перечень фактов, упомянутых в статье (минимум четыре пункта)

- Авторы представляют первую нейросетевую библиотеку, полностью реализованную на LaTeX, включая вычисления, обучение и генерацию графиков.
- LaTeX рассматривается как Тьюринг-полный язык, способный выполнять циклы, вычисления и даже обратное распространение ошибки.
- При компиляции документа LaTeX генерирует данные, обучает MLP и создаёт таблицы и графики результатов.
- Концепция “self-contained paper” делает исходник статьи полностью воспроизводимым: весь код, данные и эксперименты находятся внутри .tex файлов.

Позитивные следствия и/или достоинства описанной в статье технологии (минимум три пункта)

- Обеспечивается **полная воспроизводимость**, так как весь ML-пайплайн встроен в исходный LaTeX документ.
- Метод уменьшает необходимость в стороннем коде, объединяя написание статьи и реализацию экспериментов в одной среде.
- Использование LaTeX как вычислительной платформы демонстрирует **границы автоматизации** и новые способы работы с научными публикациями.
- Подход подчёркивает важность открытого исходного кода и доступа к данным через arXiv.

Негативные следствия и/или недостатки описанной в статье технологии (минимум три пункта)

- Компиляция LaTeX-документа, выполняющего обучение нейросети, занимает чрезвычайно большое время, делая подход практически непригодным для реальных ML-задач.
- Отсутствие стандартных программных конструкций LaTeX значительно усложняет разработку и сопровождение библиотек.
- Ограничение загрузки данных ($\text{arXiv} \leq 50 \text{ MB}$) делает невозможным использование реальных датасетов и крупных моделей.
- Высокий порог вхождения: почти никто не владеет LaTeX как языком программирования, и отладка становится трудоёмкой.

Ваши замечания, пожелания преподавателю или анекдот о программистах

Same sheet, different compile.

(Особенно если LaTeX тренирует нейросеть 48 часов 😊)