|  |  |  |
| --- | --- | --- |
| Университетская 1, Иннополис  Tatarstan, Russia, 420500  <https://github.com/kezouke> | **Смирнов Елисей** | +7 (917) 888 5135  eliwhatthe@gmail.com  [linkedin.com/in/elisei-smirnov-ds](https://www.linkedin.com/in/elisei-smirnov-ds) |
| **Кратко обо мне**  ML инженер в области NLP. Обладаю опытом построения интеллектуальных LLM-агентов, оптимизации инференса и автоматизации сложных пайплайнов. Отличаюсь системным мышлением, активным участием в open-source проектах и способностью превращать исследовательские идеи в масштабируемые решения.  **Образование** | | |
| **г. Иннополис, Россия** | **Университет Иннополис** | **Осень 2022 - Май 2026** |
| * Бакалавриат в области анализа данных и искусственного интеллекта. * Средний балл: 4,86. Текущий курс: третий курс. * Бакалаврские курсы: Операционные системы, Проектирование и анализ программных систем, Базы данных, Алгоритмы, Архитектура компьютера, Математический анализ, Дискретная математика, Дифференциальные уравнения, Теория вероятности и статистика, Оптимизация, Машинное обучение, Статистические методы для науки о данных, MLOps, Глубокое обучение. | | |
| **Опыт Работы** | | |
| **Инженер по машинному обучению (NLP)** | **Передовая Инженерная Школа** | **Апрель 2024 - Сейчас** |
|  | | |
| **BetaGuidance** (Ограниченная генерация LLM)   * Разработал методы ограниченной генерации текста, которые повысили точность вывода LLM на 90%. * Интеграция проекта с моделями Hugging Face и VLLM, повышение совместимости и эффективности асинхронных операций на 30%. * Внес [вклад](https://github.com/vllm-project/vllm/pull/4985) в библиотеку VLLM с открытым исходным кодом, позволяющую BetaGuidance использовать идентификаторы токенов в Logits Processors.     **Graphit** (Автоматизация создания образовательных курсов)   * Создал ETL-пайплайн для генерации курсов: парсинг PDF (Nougat) → дедупликация данных → обогащение контента с помощью LLM. * Настроил кластер из 6 серверов (4×A100 каждый, суммарно 24 GPU) с мониторингом через Prometheus/Grafana.   **Agnia** (Разработка многофункционального AI-агента)   * Спроектировал архитектуру LLM-агента, преобразующего NLP-запросы в последовательность API-вызовов для автоматизации задач. * Реализовал гибридный механизм извлечения аргументов API call-ов, интегрировав систему с Google Calendar, Todoist и Gmail. * Внес [вклад](https://github.com/vllm-project/vllm/pull/4985) в библиотеку VLLM с открытым исходным кодом, позволяющую BetaGuidance использовать идентификаторы токенов в Logits Processors. | | |
|  | | |
| **Личные проекты** | | |
| **Brain MRI Tumor Detection and Classification** - <https://github.com/kezouke/MRI-Diagnosis-API> | | |
| * Разработал систему классификации опухолей мозга по МРТ-изображениям. * Создал API CNN модели на основе FastAPI и веб-интерфейс с использованием Streamlit для взаимодействия с моделью. * Развернул систему с помощью Docker, обеспечивающих плавную масштабируемость и развертывание. | | |
| **Навыки** (Hard): Python, FastAPI, Streamlit, Docker, ООП, PyTorch, Hugging Face Transformers, VLLM, LLM Orchestration, Quantization, Speculative Decoding, CNN, NLP, PostgreSQL, Pandas, MLflow, Apache Airflow, Prometheus, Grafana, RabbitMQ, REST/JSON API, Git, Continuous Integration, Английский (C1)  **Навыки** (Soft): Работа в команде, Анализ и декомпозиция задач, Открыт к новым идеям и готов их предлагать. | | |