

# ДМИТРИЙ БЕРЕСНЕВ

## ML-инженер и Data Scientist

@ d.beresnev.work@gmail.com

dsomni

Telegram

Иннополис, Россия

## ОБО МНЕ

Магистрант по направлению Computer Science, специализирующийся на AI и DS с глубоким исследовательским фокусом на оптимизации в машинном обучении, LLMs и Deep Learning. Обладаю экспертизой в разработке новых ML-моделей и функциональных пайплайнов с использованием PyTorch, практическим опытом работы с классическими DRL-алгоритмами (DQN, A2C, REINFORCE). Есть опыт руководства проектами от идеи до внедрения, включая разработку EdTech-платформы с AI возможностями, обслуживающей 200+ активных пользователей. В поиске амбициозных R&D позиций для внесения вклада в создание передовых решений в области машинного обучения и ИИ.

## ОБРАЗОВАНИЕ

### Магистр, Компьютерные науки

#### ИИ и наука о данных

2024 – 2026

Университет Иннополис

- **Дипломная работа:** Новые и эффективные методы идентификации на основе граней для факторизации матриц с симплексной структурой и неполным рангом [в процессе]
- **Руководители:** Valentin Leplat, Nicolas Gillis
- **Ключевые курсы:** Анализ многомерных данных, Линейная оптимизация, Продвинутое статистика, Продвинутое машинное обучение, Обучение с подкреплением

### Бакалавр, Компьютерные науки

#### ИИ и наука о данных

2020 – 2024

Университет Иннополис

- **Дипломная работа:** Обнаружение текстового плагиата в области больших языковых моделей с использованием обучения с подкреплением
- **Руководитель:** Армен Бекларян
- **Ключевые курсы:** Методы оптимизации в машинном обучении, Обучение с подкреплением, Обработка естественного языка, Практическое глубокое обучение

## ОПЫТ ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ

### Huawei: Беспроводная передача данных

#### Исследователь, ML-инженер

2024 – наст. время

ИСП РАН и Университет Иннополис

- Проектирование и симуляция глубоких ИИ моделей для распределения устройств беспроводной связи на базовые станции с учетом временных и ресурсных ограничений для Huawei
- Разработка и внедрение моделей на PyTorch, создание пайплайна обучения-тестирования и проведение экспериментов
- **Стек:** PyTorch, Numpy
- **Руководитель:** Александр Безносовых

## Diligent Learning: Перспективы и Применения

### Исследователь, ML-инженер

📅 2024 — наст. время

📍 ИИ Центр МГУ

- Реализация и тестирование Diligent Learning — нового подхода к файнтьюнингу LLM для решения задач логического мышления на основе статьи [From Reasoning to Super-Intelligence: A Search-Theoretic Perspective](#)
- Разработка пайплайна diligent learning и файнтьюнинг больших языковых моделей в новой парадигме
- **Стек:** PyTorch, TRL, Transformers
- **Руководитель:** [Петр Анохин](#)

## Applied AlphaEvolve: CAD Reconstruction

### Исследователь

📅 Лето 2025

📍 Skoltech Summer School of Machine Learning (SMILES-2025)

- Применение OpenEvolve (открытая реализация AlphaEvolve) для реконструкции CAD-моделей и задач комбинаторной геометрии с использованием эволюционного поиска на основе LLM
- Достижение оптимальных результатов разбиения шара, соответствующих теоретическим границам в размерностях 2–13
- Превосходство над zero-shot LLM базовыми моделями на множестве сложных 3D-форм
- Разработка комплексного benchmark-пайплайна с 7 метриками оценки
- **Стек:** PyTorch, OpenEvolve, CAD libraries
- **Руководитель:** [Петр Анохин](#)

## Обнаружение текстового плагиата с использованием DRL

### Исследователь

📅 2024

📍 Университет Иннополис

- Разработка нового подхода на основе DRL для обнаружения плагиата, достигнув наилучшего MSE 0.108 на синтетическом наборе данных
- Предложение трех архитектур на основе DQN, A2C и REINFORCE с лучшими результатами у модели REINFORCE
- **Стек:** PyTorch, Numpy, Pandas
- **Руководитель:** [Армен Бекларян](#)

## ОПЫТ РАБОТЫ

---

### ЦИПР Иннополис

#### ML-разработчик

📅 2025

- Проектирование и реализация RAG-пайплайна для проприетарных репозиторий фронтенда на Angular
- Создание индексеров: Inverse Index, BallTree с эмбедами, полученными от модели, и частично Faiss
- Подключение локальных генеративных моделей и проектирование полного пайплайна сбора данных, генерации эмбедингов, индексации и запроса
- Подтверждение качества на золотых запросах, предоставленных экспертами
- **Стек:** PyTorch, Docker, FastAPI, Faiss

### Газпром ЦПС

#### ML-инженер

📅 2024

- Разработка и обучение предсказательной ML-модели для выявления причин дефектов на строительных объектах, достигнув точности 80% на проприетарном наборе данных
- Ответственность за полный рабочий пайплайн: предварительная обработка данных, инжиниринг признаков, создание и валидация модели
- **Использованные концепции:** Модели на основе деревьев, MLP-модели, ансамблевые методы, Трансформеры
- **Стек:** PyTorch, Numpy, Scikit-learn, Pandas

## Передовая инженерная школа УИ

### ML-разработчик

📅 2023

- Разработка модели генерации кода с использованием трансформерной архитектуры
- Файнтьюнинг модели **Gorilla** на проприетарном наборе данных
- Значительный вклад в исследование
- **Использованные концепции:** LLM, Трансформеры, LoRA
- **Стек:** PyTorch, Numpy, Pandas

## ИЗБРАННЫЕ ПРОЕКТЫ

---

### Асепт School

#### Основатель, CEO

📅 2023 — наст. время

- Руководство полным циклом проектирования комплексной EdTech-платформы, объединяющей ML с современными веб-технологиями, в настоящее время используемой в образовательных организациях с примерно 200 активными пользователями
- Разработка системы обнаружения плагиата в коде на основе ML и внедрение генеративного ИИ для подсказок, генерации текста и изображений с использованием LLM с открытым исходным кодом
- Определение процессов разработки и эксплуатации, разработка бэкенда на FastAPI и MongoDB, создание фронтенда на Next.js
- **Стек:** PyTorch, FastAPI, Next.js, MongoDB, Docker, Apache Kafka

### DoWell

#### ML-разработчик, Технический руководитель

📅 2025

- Проектирование и реализация RAG-архитектуры для интеллектуальной диалоговой системы, имитирующей экспертные консультации в различных профессиональных областях
- Развертывание и подключение генеративных моделей, разработка бэкенда на FastAPI
- **Стек:** PyTorch, FastAPI, Docker

### Обнаружение ИИ-сгенерированного Python кода с помощью ML

#### Разработчик

📅 2025

- Достижение точности 95.9% с моделью CodeBERT на синтетическом наборе данных
- Разработка эффективного Random Forest на основе AST, достигающего точности 83.5% со временем вывода 2мс
- Создание пайплайна генерации набора данных с использованием 4 LLM со специализированными промптами
- Интеграция фреймворка интерпретируемости LIME для интерпретации модели
- **Стек:** PyTorch, Transformers, Tree-sitter, LIME

### PyFinder: быстрый поиск по документации Python

#### Разработчик

📅 2025

- Создание системы информационного поиска, объединяющей традиционное обратное индексирование с семантическим поиском на основе LLM
- Показатели: F1@1=0.53, nDCG@1=0.83 с LLM-эмбедами + индексатором Ball Tree
- Реализация RAG-пайплайна с prompt engineering и извлечением контекста
- **Стек:** PyTorch, sentence-transformers, FastAPI, Next.js

# ОПЫТ ПРЕПОДАВАНИЯ

Ассистент преподавателя

Университет Иннополис

📅 2024 – 2025

Ассистент преподавателя по курсу «Введение в методы оптимизации» для бакалавров второго курса. Проведение тестов и лабораторных работ.

Ассистент преподавателя

Летняя школа Яндекса по математике в ИИ

📅 Июль 2024

📍 Иннополис

Разработка и преподавание материалов для семинаров и домашних заданий по курсу «Методы оптимизации в машинном обучении» под руководством **Александр Безносикив**

# НАВЫКИ

ML & DS

PyTorch, TRL, Scikit-learn, PuLP, Numpy, Pandas, JAX

Языки программирования

Python, TypeScript, C/C++, Rust, Java, Haskell

Инструменты и платформы

Docker, Git, LaTeX, Postman, Faiss

Веб и базы данных

FastAPI, Next.js, Astro, PostgreSQL, MongoDB

# ЯЗЫКИ

Русский



Английский

