

Введение в MLOps

Ирина Степановна Трубчик

<https://t.me/+PsC-JDrwrvsxNmVi>

Лекция 1

О чем эта лекция

1

Что такое MLOps и почему это актуально

2

Эволюция: от DevOps к MLOps

3

Ключевые проблемы production-систем с ML

4

Жизненный цикл ML-проекта

5

Уровни зрелости MLOps

6

Инструменты и компоненты MLOps-экосистемы



MLOps (Machine Learning Operations) — это совокупность практик и инструментов для автоматизации, масштабирования и управления жизненным циклом машинного обучения в производственной среде.

Это культура взаимодействия между Data Scientists, ML-инженерами и DevOps-специалистами.

Аналогия



DevOps объединяет разработку и эксплуатацию ПО



MLOps добавляет в эту цепочку данные и модели, которые становятся first-class citizens

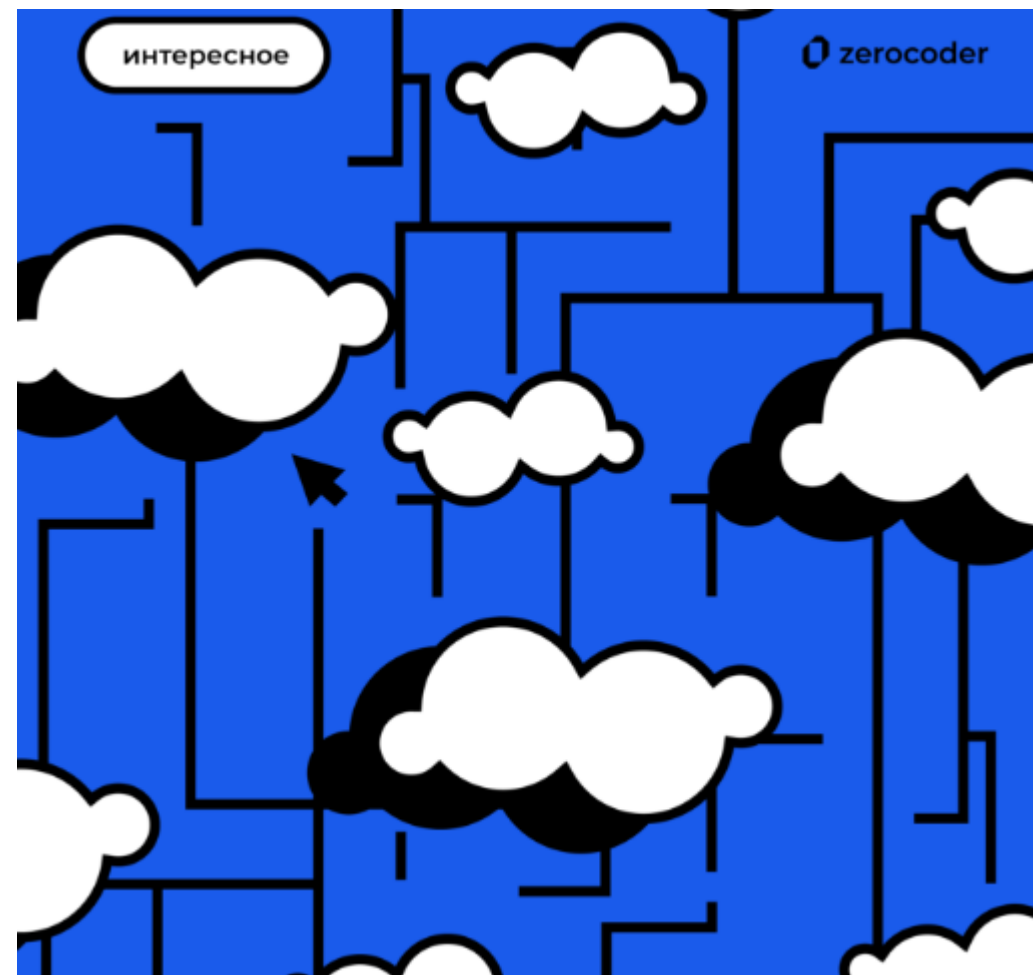


Почему MLOps — это не просто DevOps для ML

- *ML-системы зависят от данных, а не только от кода.*
- *Модели деградируют со временем (дрейф данных).*
- *Воспроизводимость экспериментов критически важна.*

Пример провала без MLOps:

Модель, работающая локально, не может быть развернута из-за проблем с окружением или данными



Жизненный цикл ML-проекта



Проблемы при развертывании

- Отсутствие стандартизированных процессов.
- Сложность настройки окружения
- Необходимость ручной настройки параметров.
- Отсутствие мониторинга и управления.
- Нехватка ресурсов для поддержки.



Последствия

- Снижение точности модели
- Увеличение времени на внедрение
- Рост затрат
- Снижение доверия к системе



Эволюция: от DevOps к MLOps

- ✓ DevOps: Автоматизация развертывания и мониторинга ПО
- ✓ DataOps: Автоматизация ETL-процессов и управления данными
- ✓ MLOps: Интеграция ML-моделей в production, включая данные, обучение и мониторинг

Ключевое отличие: В ML код — это лишь часть системы. Данные и модели требуют отдельного управления версиями и мониторинга.

Проблемы Production ML

1. Дрейф данных (Data Drift)
2. Концептуальный дрейф (Concept Drift)
3. Проблема воспроизводимости (Reproducibility)
4. Масштабируемость



Жизненный цикл ML-проекта



Уровни зрелости MLOps



Архитектурная схема

- ✓ Версионирование: Данные (DVC), код (Git), модели (MLflow).
- ✓ Оркестрация: Пайплайны (Airflow, Prefect).
- ✓ Сервинг: Развертывание (FastAPI, Docker, Kubernetes).
- ✓ Мониторинг: Метрики (Prometheus), дрейф (Evidently).
- ✓ Автоматизация: CI/CD (GitHub Actions).



Итоги



MLOps решает проблемы эксплуатации ML-моделей.



- Основные вызовы: дрейф, воспроизводимость, масштабируемость



Жизненный цикл ML — это непрерывный процесс



Контакты



Телеграм <https://t.me/+PsC-JDrwrvsxNmVi>



СКИФ

