Санкт-Петербургский государственный университет

Кафедра системного программирования

Обзор принципов, подходов и инструментов MLSecOps, используемых при реализации ML-пайплайнов

Хокимзода Муборакшои Иноятулло, группа 24.М41-мм

Научный руководитель: к.ф.-м.н. Д.В. Луцив, доцент кафедры системного программирования
Консультант: В.А. Андриенко, старший преподователь кафедры системного программирования

Санкт-Петербург
2024

Введение

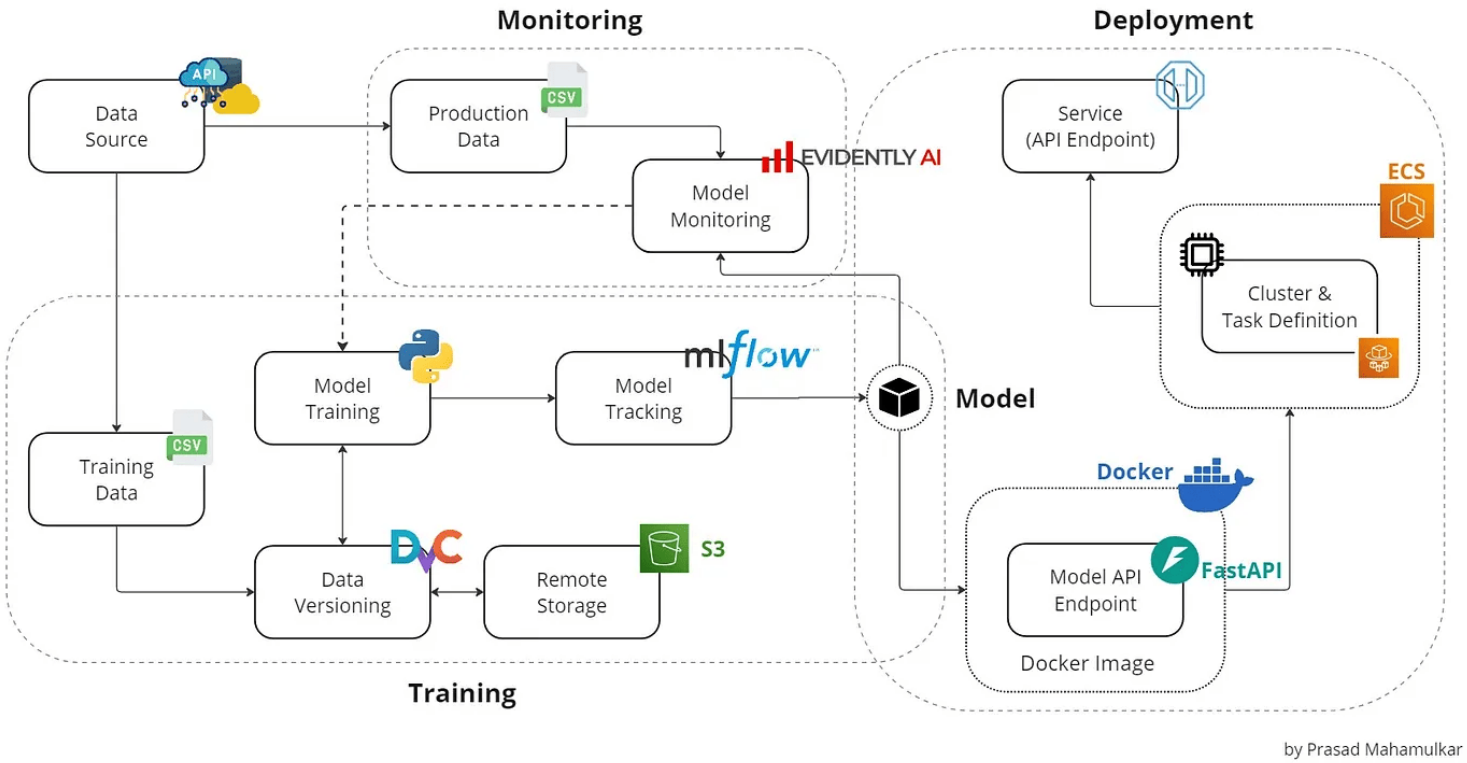
- Машинное обучение (ML) активно используется в различных отраслях, однако связано с новыми вызовами безопасности.
- **MLSecOps:** Комплексный подход для защиты данных, моделей и инфраструктуры.
- **Основная цель:** обеспечение безопасности и надёжности ML-моделей на всех этапах жизненного цикла.

Постановка задачи

- Исследование концепций и принципов MLSecOps
- Анализ существующих архитектурных решений и инструментов для обеспечения безопасности в ML
- Разработка комплексной архитектурной схемы MLSecOps включающей все этапы жизненного цикла моделей (от сбора данных до развертывания)
- Визуализация разработанной архитектуры для удобства понимания и внедрения
- Формирование подробных рекомендаций по практическому внедрению системы MLSecOps, включая методы управления рисками, мониторинг безопасности на всех этапах и реагирования на инциденты.

Основы MLSecOps

- Основные направления:
  - **Безопасность данных:** шифрование, контроль доступа, анонимизация.
  - **Защита моделей:** от атак и утечек данных.
  - **Мониторинг:** управление конфигурациями, отслеживание аномалий.



Архитектуры внедрения ML-пайплайнов

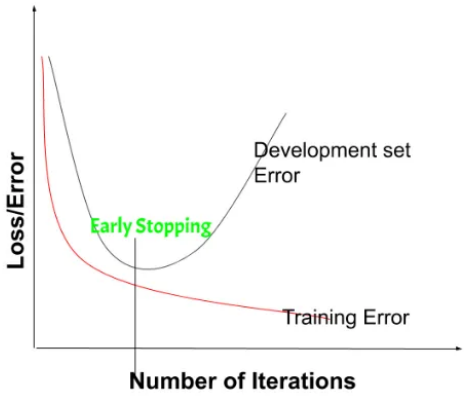
- **Batch Deployment:** предсказания выполняются с периодичностью.
- **Real-time Deployment:** мгновенные ответы через HTTP API.
- **Streaming Deployment:** использование брокеров сообщений (Kafka).
- **Edge Deployment:** выполнение моделей на клиентских устройствах.

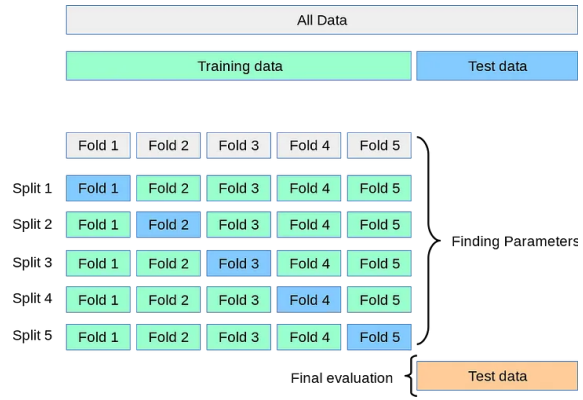
Инструменты безопасности

- **Проверка артефактов:**
  - Gitleak, SonarQube, Trivy, OWASP Dependency-Check.
  - Compliance-Checker: проверка датасетов.
- **Защита данных и моделей:**
  - Использование HashiCorp Vault для управления секретами.
  - OpenPubKey для подписания моделей.

Обучение и тестирование моделей

- **Обучение:**
  - Ранняя остановка (Early Stopping).
  - Кросс-валидация (KFold).

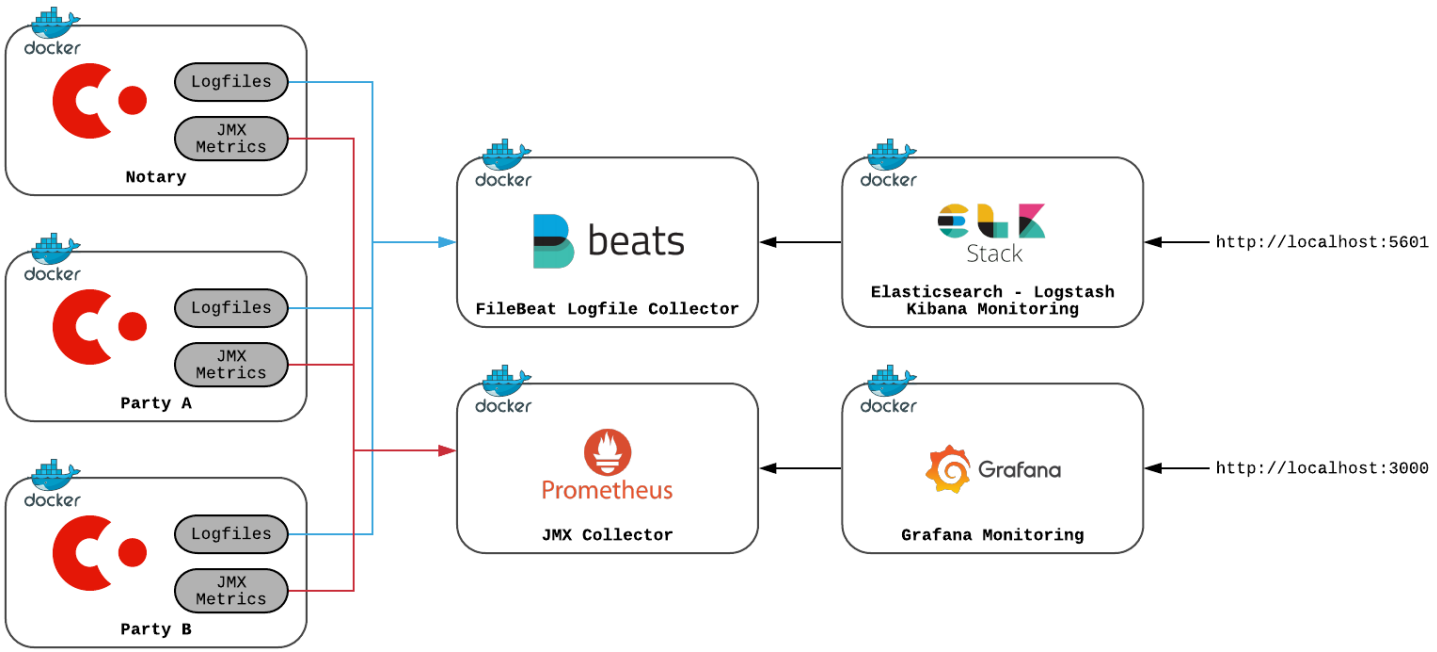




- **Тестирование:**
  - Shadow-режим: тестирование без влияния на продакшн.
  - Canary deployment: тестирование на части пользователей.
  - A/B тесты и многорукие бандиты.

Мониторинг и логирование

- **Инструменты:**
  - Prometheus, Grafana: визуализация и мониторинг.
  - ELK Stack: централизованное логирование.
  - Telegram/Slack: уведомления.

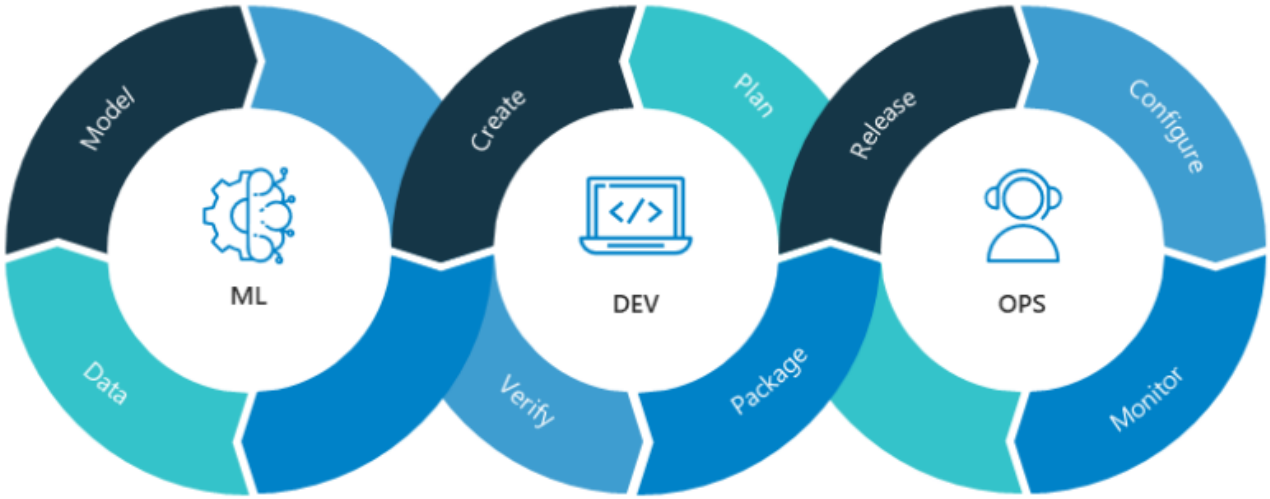


Заключительный этап: Деплоймент

- Подписание модели:
  - Использование OpenPubKey.
  - Хранение в AWS S3, Nexus, Google Cloud S3.
- Управление ключами:
  - HashiCorp Vault, AWS KMS.
- Контроль качества:
  - Полный аудит перед развертыванием.

Безопасность на всех этапах

- Интеграция безопасности на каждом этапе:
  - От сбора данных до мониторинга.
  - Устранение уязвимостей на ранних этапах.
- Результат:
  - Снижение рисков.
  - Повышение доверия к моделям.



Заключение

- Основной вывод работы заключается в том, что внедрение подхода MLSecOps на каждом этапе жизненного цикла модели позволяет минимизировать уязвимости, снизить риски утечек данных и атак, а также обеспечить стабильную и безопасную эксплуатацию моделей в реальных условиях.
- Описанные принципы, подходы, а также инструмен- тарий MLSecOps можно рассматривать как комплексное решение, способное повысить доверие к системам машинного обучения и обеспечить их надежность и отказоустойчивость в современных приложениях.

Источники

- Н. Гифт, К. Берман, А. Деза, Г. Георгу Python и DevOps: Ключ к автоматизации Linux. Перевел с английского И. Пальти. ООО "ПИТЕР". Россия, Санкт-Петербург 2022
- Github Repository: A curated list of awesome open-source tools, resources, and tutorials for MLSecOps. https://github.com/noobpk/MLSecOps-DevSecOps-Awesome
- Архитектура реальной системы машинного обучения. https://habr.com/ru/companies/vk/articles/673782/
- ML System Design: основные способы деплоя и тестирования моделей машинного обучения в продакшене. https://habr.com/ru/articles/739316/