

# Мониторинг качества

**Антон Репушко**

ML-инженер, независимый исследователь

Skillbox

# Антон Репушко

**8+ лет опыта**  
в машинном обучении

---

**Был участником**  
**CERN**

**Сделал свой**  
**ML-стартап**

---

**Победитель многих**  
**ML-соревнований**



# Цель модуля

Научиться отслеживать деградацию ML-моделей после их внедрения.

# Задачи модуля

- 1 Разобрать, что такое деградация ML-моделей
- 2 Определить её основные причины
- 3 Выяснить, какие методы существуют для её детектирования
- 4 Научиться пользоваться специальными инструментами для детектирования

# Почему ML-модели деградируют

Почему ML-модели деградируют

# Цель темы

Разобрать основные причины деградации ML-моделей.

# Два типа событий

Характерные для всех систем

Характерные для ML-систем

# Характерные для всех систем

- 1 Сбои в оборудовании
- 2 Проблемы с масштабированием
- 3 Проблемы с сетевым подключением
- 4 Уязвимости и проблемы безопасности
- 5 Ошибки в коде



Почему ML-модели деградируют

# Способы решения

- Мониторинг
- Отказоустойчивая архитектура
- Регулярные обновления

Почему ML-модели деградируют

# Характерные для ML-систем

**Data Drift**

Почему ML-модели деградируют

# Data Drift

Ситуация, когда распределение реальных данных начинает отличаться от распределения обучающих данных.

# Подробнее о Data Drift

- ❶ Изменение распределения входных признаков (Feature Drift)
- ❷ Изменение взаимосвязи между признаками и целевой переменной (Concept Drift)
- ❸ Появление новых категорий и значений, которых раньше не было в данных

# Методы для выявления Data Drift

- ❶ Тест Колмогорова — Смирнова
- ❷ Дивергенция Кульбака — Лейблера
- ❸ Метод скользящего окна

# Вывод темы

Разобрались в основных причинах деградации ML-моделей и в специальных терминах, описывающих этот процесс.