



ИНСТИТУТ
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
УНИВЕРСИТЕТА ИННОПОЛИС



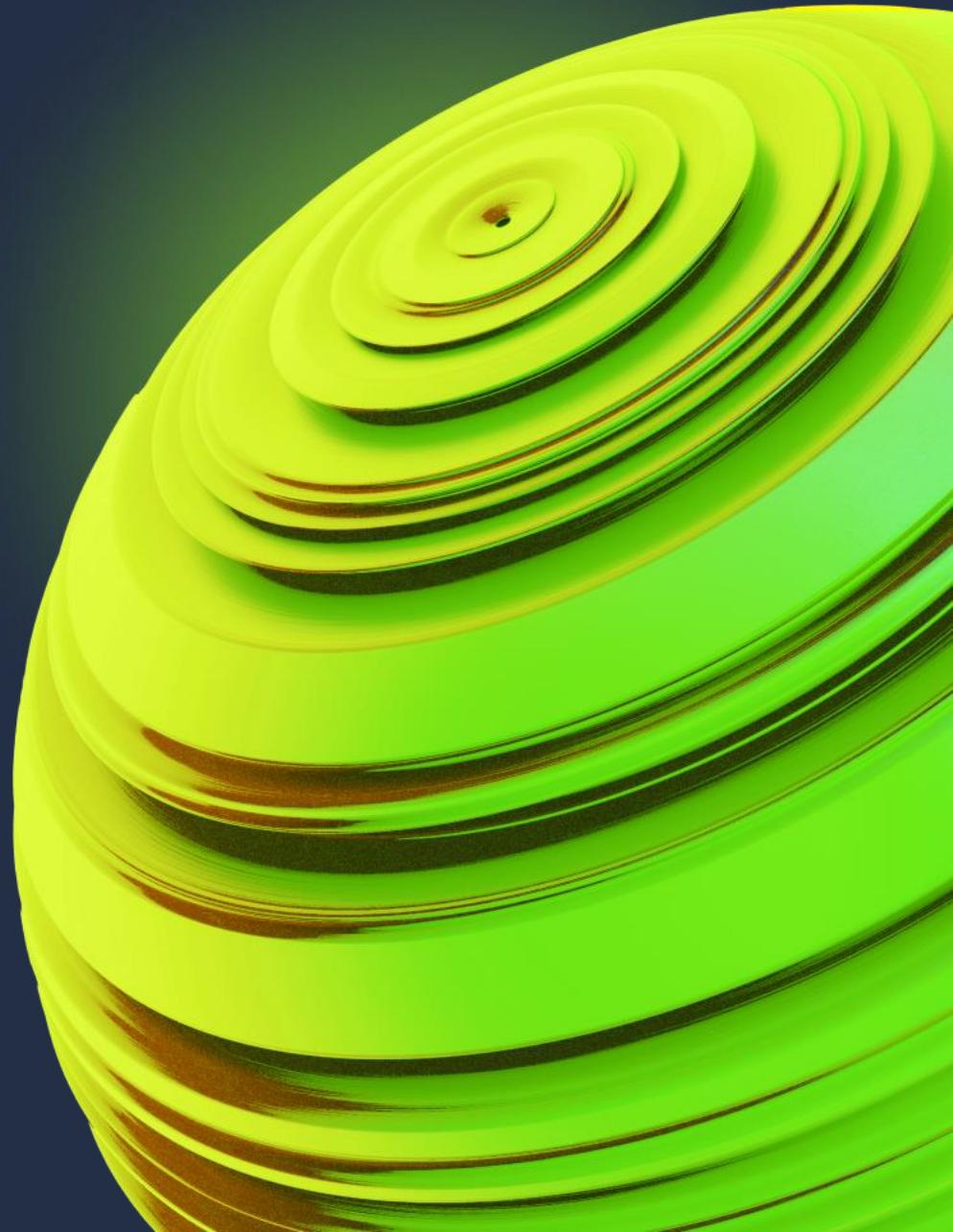
УНИВЕРСИТЕТ
ИННОПОЛИС

Презентация по итоговой аттестации курса «Архитектор в области искусственного интеллекта»

Название проекта: Система предиктивного обслуживания оборудования

👤 Макаров Алексей

📅 2025 год



Оглавление:

1. Проблематика и актуальность
2. Цели и задачи проекта
3. Архитектура системы
4. Анализ данных и EDA
5. Методология и план экспериментов
6. Результаты сравнения моделей
7. Финальная модель и её оценка
8. Достижения проекта
9. Ограничения и выявленные проблемы
10. Рекомендации для внедрения и развития
11. Итоги и выводы

1. Проблематика и актуальность

Проблема: Незапланированные простоя → 1,35 млн руб./час убытков

Традиционные подходы неэффективны:

- По графику → избыточные затраты
- По факту отказа → катастрофические простоя

Решение: Предиктивная аналитика (Industry 4.0)

Ключевые драйверы:

- Снижение затрат на 30%
- Повышение ОЕЕ (Общей эффективности оборудования)
- Безопасность производства

2. Цели и задачи проекта

Технические цели:

- ML-модель: Recall ≥ 0.85 , FPR < 0.15
- Время инференса < 50 мс
- Интеграция с SCADA/MES

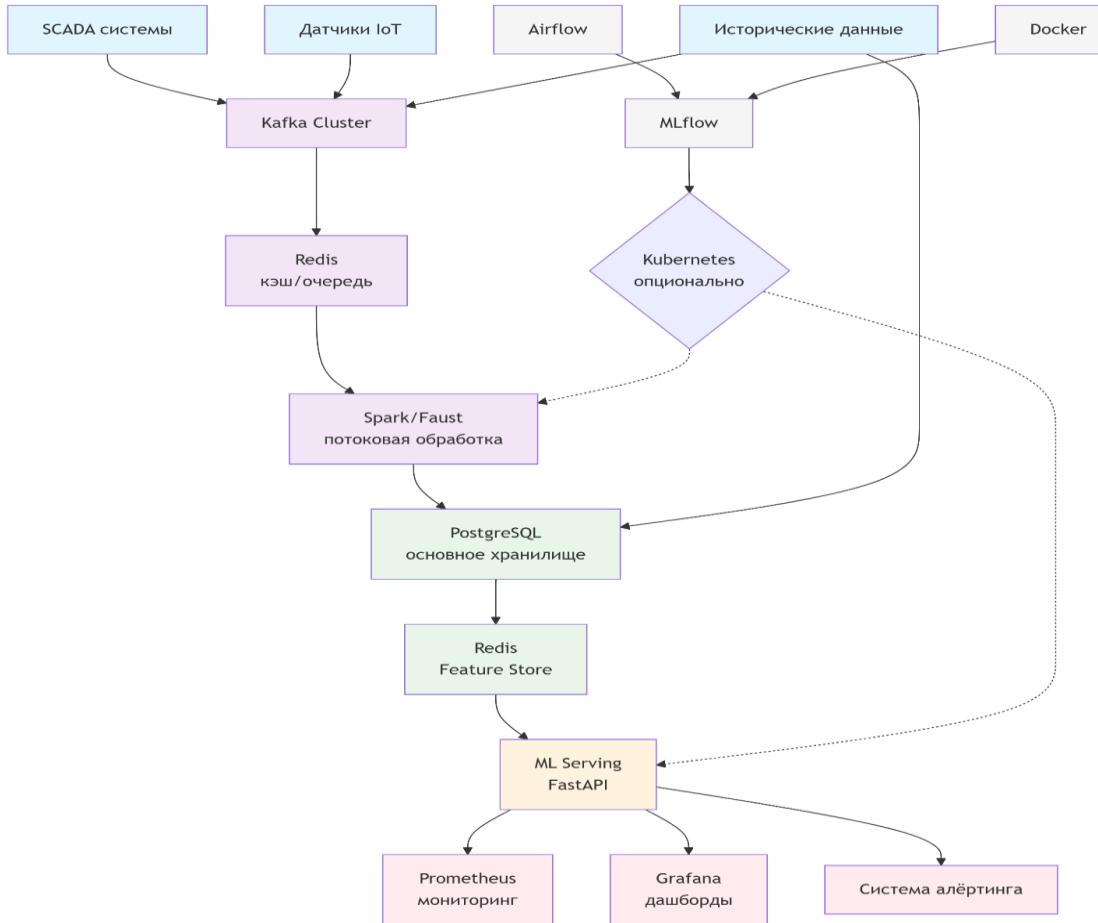
Бизнес-цели:

- Сокращение затрат на 30% в pilotной зоне
- Увеличение MTBF на 25%
- ROI $\geq 200\%$ в первый год

Архитектурные цели:

- Масштабируемая микросервисная платформа
- Доступность 99.95%

3. Архитектура системы



Компоненты:

- Сбор данных: Kafka, IoT-датчики, SCADA
- Обработка: Spark/Faust (реальное время)
- Хранение: PostgreSQL (основное), Redis (кэш/фици)
- ML-serving: FastAPI
- Мониторинг: Prometheus, Grafana
- Оркестрация: Airflow, Docker

Принципы:

- Микросервисы, event-driven design
- Многоуровневое хранение (hot/warm/cold)

4. Анализ данных и EDA

Датасет: Predictive Maintenance Dataset (10 000 записей, 10 признаков)

Ключевые проблемы:

- Дисбаланс классов: 3.4% отказов
- Высокая корреляция признаков (температуры, torque/speed)
- Отсутствие временных меток

Решение:

- SMOTE (создание синтетических примеров) + class weighting (взвешивание классов для обучения)
- Создание новых признаков:
 - temp_diff (разность температур) = Process temp – Air temp
 - $\text{torque_to_speed_ratio}$ = Torque / Rotational speed

5. Методология и план экспериментов

Этапы экспериментов:

1. Baseline: логистическая регрессия
2. Обработка дисбаланса (SMOTE, class weights)
3. Сравнение алгоритмов (Random Forest, XGBoost, LightGBM, SVM)
4. Feature engineering
5. Оптимизация порога классификации

Выбранные алгоритмы:

- LightGBM (основной)
- XGBoost, Random Forest (сравнение)

6. Результаты сравнения моделей

Модель	Recall	FPR	Precision	F1-Score	AUC-ROC
Random Forest	0.728571	0.001554	0.944444	0.822581	0.986247
XGBoost	0.785714	0.009845	0.743243	0.763889	0.989297
LightGBM Base	0.871429	0.005699	0.847222	0.859155	0.992339
LightGBM Opt	0.971429	0.063731	0.356021	0.521073	0.991203
SVM	0.942857	0.065285	0.34375	0.503817	0.97587
SMOTE+LGB	0.857143	0.011917	0.722892	0.784314	0.99205
Ensemble	0.828571	0.004145	0.878788	0.852941	0.991207

7. Финальная модель и её оценка

Конфигурация LightGBM:

```
n_estimators: 450  
num_leaves: 70  
learning_rate: 0.1  
class_weight: 'balanced'
```

Калибровка порога: Оптимальный порог = 0.18

Метрики на тестовой выборке:

- ✓ Recall = 0.88
- ✓ FPR = 0.008
- ✓ Precision = 0.80
- ✓ F1-Score = 0.84
- ✓ AUC-ROC = 0.98
- ✓ Время инференса = 3.46 мс

8. Достижения проекта

Бизнес-эффективность: Прогнозируемая годовая экономия 49.5 млн рублей при сроке окупаемости < 6 месяцев.

Технические показатели: Модель LightGBM с калиброванным порогом 0.18 обеспечивает оптимальный баланс между безопасностью (Recall) и экономической эффективностью.

Архитектурная готовность: Реализована распределенная система с поддержкой real-time обработки, мониторинга и автоматического переобучения.

9. Ограничения и выявленные проблемы

Качество данных: Зависимость от точности показаний датчиков требуется дополнительная система валидации входящих данных.

Дисбаланс классов: Несмотря на применение advanced техник, сохраняется сложность в детектировании редких типов отказов.

Оборудование Type H: Сниженная точность (Precision 0.58) для высококачественного оборудования требует отдельной доработки модели.

Временная зависимость: Отсутствие явных временных меток в датасете ограничивает анализ трендов и паттернов деградации.

10. Рекомендации для внедрения и развития

Поэтапное внедрение: Начать с пилотной зоны (10-20 единиц оборудования) для валидации в реальных условиях.

Мониторинг production: Реализовать систему детектирования дрейфа данных (Evidently AI) и автоматического переобучения

Улучшение модели:

- Разработка отдельных моделей для разных типов оборудования.
- Внедрение ансамблевых методов для редких типов отказов.
- Интеграция данных о предыдущих ремонтах и обслуживании.

Расширение функционала:

- Прогнозирование оставшегося срока службы (RUL) в дополнение к бинарной классификации.
- Рекомендательная система по оптимальному времени обслуживания.
- Интеграция с системами планирования производства (MES).

11. Итоги и выводы

Цели достигнуты:

- Модель превзошла целевые метрики (Recall 0.88, FPR 0.008)
- Архитектура готова к масштабированию
- Экономический эффект подтверждён

Готовность к внедрению:

- Все артефакты в репозитории
- Поддержка контейнеризации
- План внедрения и мониторинга

Перспективы:

- Расширение на другие типы оборудования
- Интеграция с AI-платформами предприятия



ИНСТИТУТ
ДОПОЛНИТЕЛЬНОГО
ОБРАЗОВАНИЯ
УНИВЕРСИТЕТА ИННОПОЛИС

Спасибо за внимание!

Контакты

📞 +7 916 3096883

✉ alex_makarov@rambler.ru



Сайт

