

|   |   |  |
|---|---|--|
| <br>ШКОЛА БОЛЬШИХ ДАННЫХ | <b>ООО «Учебный центр «Коммерсант» ©</b><br><b>«Школа Больших Данных»</b><br><a href="http://www.bigdataschool.ru">www.bigdataschool.ru</a><br>2026 |  |
|---|---|--|

## Программа курса

### «AGENT: AI-агенты для оптимизации бизнес-процессов»

**О курсе:** 6-дневный интенсивный курс AI агенты для компаний по проектированию, разработке и внедрению AI-агентов на основе LLM (Llama), многоагентных фреймворков (LangGraph, CrewAI, AutoGen) и интеграции с корпоративными системами. Живые занятия с практикой, сопровождение домашних заданий, итоговый проект под ваш кейс индивидуально или в составе группы + консультации по проекту.

**Аудитория:**

- Разработчики и data инженеры
- ML/AI-специалистам
- Архитекторам/тимлидам
- ML инженерам

**Уровень подготовки:**

- Опыт программирования на любом языке
- Базовые знания System Design (не обязательный навык)

**Продолжительность курса:** 24 академических часа, 6 дней по 4 часа дистанционно

### Содержание программы

#### 1. Использование LLM

- Архитектура трансформеров. от BERT до Llama 3
- Autoencoders и принципы работы LLM
- Открытый vs проприетарный стек моделей
- Критерии выбора. размер модели, аппаратные требования, лицензирование
- Бенчмаркинг на целевых задачах
- Обзор инструментов Ollama, LM Studio, OpenWebUI

**Практическая часть:**

- Развёртывание моделей через Ollama на облачной инфраструктуре
- Сравнение производительности Llama 3 vs Mistral 7B
- Prompt инжиниринг на открытых моделях под задачи суммаризации встреч, анализа кода, генерации скелета презентации

#### 2. RAG для обогащения моделей внешними источниками знаний

- Использование RAG (Retrieval Augmented Generation) для обогащения моделей
- Архитектура RAG системы с opensource компонентами
- Выбор моделей эмбеддингов под различные задачи и языки
- Векторные базы данных: In-memory решения (FAISS, Annoy)
- Специализированные векторные БД (Chroma, Qdrant, Weaviate, Milvus/Zilliz)
- Расширения векторных БД PgVector для PostgreSQL
- Критерии выбора. производительность, масштабируемость

**Практическая часть:**

- Построение RAG с использованием Sentence Transformers
- Загрузка данных в Quadrant
- Настройка метрик расстояния и фильтрации
- Оценка качества релевантности

#### 3. Разработка ИИ агента на своей LLM модели

- Архитектура ИИ агента на основе LLM с контролем безопасности
- Техники снижения галлюцинаций



ШКОЛА БОЛЬШИХ ДАННЫХ

ООО «Учебный центр «Коммерсант» ©

## «Школа Больших Данных»

[www.bigdataschool.ru](http://www.bigdataschool.ru)

2026



- Мультимодальные возможности (Llava, Bakllava)
- СЦМ создания ИИ-агента на кастомных данных

### Практическая часть:

- Разработка специализированного ИИ агента на Llama 3
- Настройка system prompt для предметной области
- Интеграция с внешними API
- Тестирование и валидация ответов

## 4. Разработка кастомного ИИ агента

- Архитектура кастомных ИИ агентов
- Обзор opensource решений для разработки ИИ агентов
- Инструменты LangGraph, LangChain для создания агентов
- Графы состояний и многоагентные системы

### Практическая часть:

- Реализация ИИ агента на основе LangGraph
- Интеграция с открытыми моделями reasoning (Qwen2.5)
- Создание multistep workflow
- Встраивание векторной БД как памяти агента

## 5. Nocode платформы для быстрого прототипирования

- использование No-Code/Low-code платформ для ИИ разработки (n8n, Langflow, LangGraph)
- Кейсы использования, когда выбирать No-Code vs код
- Интеграция с внешними API и базами данных

### Практическая часть:

- Создание ИИ-агента в Langflow без программирования
- Интеграция с векторными базами данных через UI
- Развёртывание прототипов в облако

## 6. Развёртывание и эксплуатация ИИ агентов в проде

- Эксплуатация LLM в проде. фреймворки sglang/ollama
- MLOps для открытых LLM
- Развёртывание, мониторинг, версионирование
- Мониторинг дрифта и качества моделей

### Практическая часть:

- Развёртывание в продакшен с vLLM
- Настройка мониторинга с Prometheus/Grafana
- A/B тестирование моделей
- Оптимизация затрат на инференс