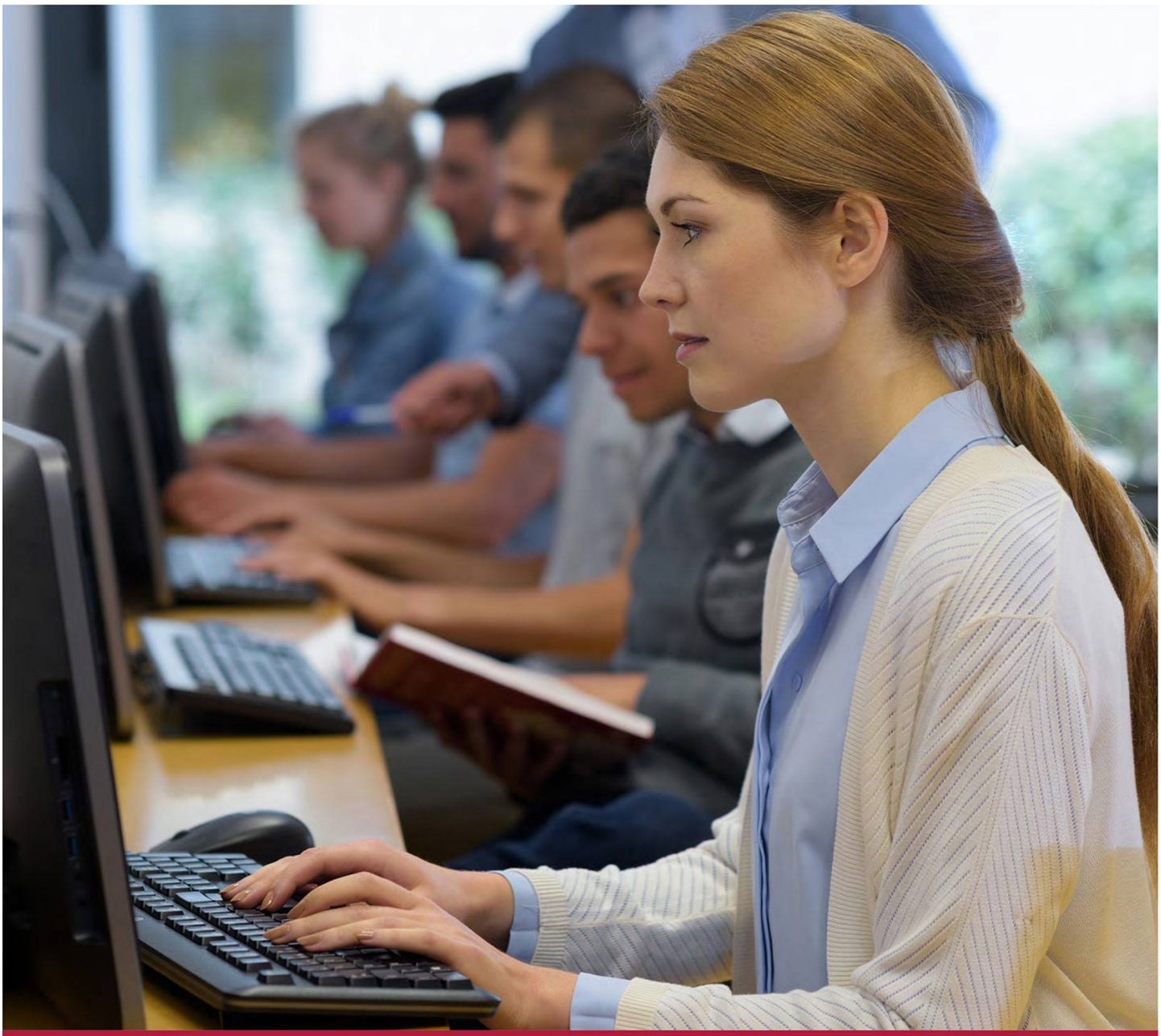
Академия АйТи
a Softline Company



MLSecOps. Обеспечение безопасности и отказоустойчивости систем машинного обучения

Код курса: MLSecOps

+7 (495) 150 96 00 | academy@academyit.ru | academyit.ru

MLSecOps. Обеспечение безопасности и отказоустойчивости систем машинного обучения

Код курса: MLSecOps

Длительность	48 ак. часов
Формат	
Разработчик курса	Академия АйТи
Тип	Учебный курс
Способ обучения	Под руководством тренера

О курсе

В сфере ИИ и больших данных укрепляется новое направление – MLSecOps. Оно находится на стыке машинного обучения и информационной безопасности, требует комбинированных навыков как в области информационной безопасности, так и управления данными. К таким навыками относятся: анализ угроз и уязвимостей в системах машинного обучения, оценка эффективности мер безопасности и контроль разработки и внедрения решений MLSecOps. Специалисты MLSecOps защищают как сами модели искусственного интеллекта, так и весь процесс их сборки и эксплуатации. Интерес к направлению стремительно растет, а спрос на специалистов ежегодно увеличивается более чем на 100%. Актуальность также основывается на развитии сферы искусственного интеллекта, а также аналитики и изучения данных (Data Analysis и Data Science) и активизации злоумышленников.

Подробная информация

Профиль аудитории:

- федеральные служащие, государственные гражданские и муниципальные служащие
- работники министерств, ведомств, департаментов, чья работа связана с ИИ
- руководители крупных предприятий и организаций, в которых используется или внедряются системы ИИ
- работники бюджетных организаций регионального и федерального уровня
- сотрудники предприятий крупного и среднего бизнеса, специалисты кадровых служб
- начинающие IT-специалисты и студенты технических специальностей
- сотрудники силовых структур и финансовых учреждений, внедряющие в работу нейронные сети и машинное обучение

Предварительные требования:

- навыки работы на компьютере на уровне опытного пользователя
- базовые навыки обеспечения информационной безопасности в качестве пользователя (сложные пароли, антивирусная защита, блокировка экрана)
- базовые навыки работы с нейронными сетями

По окончании курса слушатели будут уметь:

- классифицировать атаки MLSecOps и выбирать необходимые меры противодействия
- применять теоретические знания при решении прикладных задач обеспечения безопасности искусственного интеллекта
- применять знания об анализе данных, машинном обучении, информационной безопасности для выработки управленческих, исследовательских, хозяйственных и других решений в рамках своей профессиональной деятельности
- обосновывать профессиональное мнение о роли и потенциале использования технологий MLSecOps в государственном управлении
- подготавливать предложения в плановые, нормативные и другие документы по обеспечению безопасности искусственного интеллекта в государственном управлении

Программа курса

Модуль 1. Введение в MLSecOps и LLMSecOps

- Современные тренды IT-сферы. Появление сферы MLSecOps в IT
- Обзор концепций MLSecOps и LLMSecOps. Определение, цели, ключевые принципы
- Место MLSecOps в общей архитектуре кибербезопасности. Интеграция MLSecOps с существующими процессами безопасности
- Основные угрозы и уязвимости в ML и LLM, ущерб от ML-инцидентов. Обзор ресурсов для изучения MLSecOps

Модуль 2. Основы анализа больших данных в MLSecOps

- Источники и форматы данных для MLSecOps. Обзор типов данных, используемых в задачах безопасности ML
- Сбор, обработка и подготовка данных для ML. Техники очистки, преобразования и нормализации данных
- Визуализация данных для анализа аномалий. Использование графиков и диаграмм для выявления подозрительной активности
- Основы статистического анализа для MLSecOps. Применение статистических методов для анализа данных безопасности

Модуль 3. Основы машинного обучения в MLSecOps

- Основные понятия и типы машинного обучения. Обучение с учителем, без учителя, с подкреплением
- Алгоритмы машинного обучения и MLSecOps-риски. Классификация, кластеризация, обнаружение аномалий
- Безопасность моделей машинного обучения. Метрики точности, полноты, F1-меры, AUC-ROC
- Популярные библиотеки машинного обучения Scikit-learn, TensorFlow, PyTorch и другие

Модуль 4. Основы обеспечения информационной безопасности в MLSecOps

- Основные принципы информационной безопасности. Конфиденциальность, целостность, доступность.
- Управление доступом и аутентификация. Методы контроля доступа к данным и ресурсам

- Шифрование и хеширование данных. Применение криптографических методов для защиты конфиденциальности
- Основы сетевой безопасности. Защита сетевой инфраструктуры и каналов передачи данных

Модуль 5. Оценка рисков и разработка стратегии защиты в MLSecOps

- Основные риски в MLSecOps и базовые методы защиты
- Специфические риски MLSecOps, характерные для разных отраслей
- Человеческий фактор как один из главных рисков MLSecOps. Социальная инженерия
- Методы оценки рисков и разработка стратегии защиты в MLSecOps

Модуль 6. OWASP TOP-10 угроз в MLSecOps и LLMSecOps и рекомендуемые методы защиты

- Атаки типа Prompt Injection: манипуляция запросами к LLM. Атаки типа JailBreaks. Компрометация LLM
- Атаки при обучении моделей (Data Poisoning). Примеры и методы защиты
- Атаки на конфиденциальность моделей ML. Извлечение конфиденциальных данных из моделей. Атаки типа Model Extraction: кража моделей LLM
- Уязвимости в интеграции LLM с внешними системами. Возможные риски и методы защиты

Модуль 7. Безопасный жизненный цикл MLSecOps

- Безопасная разработка моделей ML (Secure by Design). Встраивание безопасности на этапе проектирования
- Безопасное обучение моделей ML. Защита от отравления данных (Data Poisoning), атак на обучение
- Безопасное развертывание моделей ML. Обеспечение безопасности инфраструктуры
- Безопасное обслуживание и обновление моделей ML. Мониторинг и обновление моделей для защиты от новых угроз
- Облачные платформы как средство обеспечения надежной инфраструктуры MLSecOps

Модуль 8. Реагирование на инциденты безопасности MLSecOps

- Планы реагирования на инциденты. Процедуры и алгоритмы для быстрого реагирования на атаки
- Изоляция и анализ инцидентов. Методы анализа и расследования инцидентов безопасности
- Восстановление после инцидентов. Стратегии восстановления работоспособности систем
- Уроки, извлеченные из инцидентов. Postmortems. Анализ инцидентов для улучшения будущей защиты

Модуль 9. Обеспечение надежности и отказоустойчивости в MLSecOps

- Обеспечение высокой доступности ML-систем. Методы предотвращения простоев в работе. Техники SRE
- Резервирование и резервное копирование данных. Защита от потери данных и сбоев системы
- Мониторинг инфраструктуры ML. Отслеживание состояния серверов и другого оборудования
- Планирование аварийного восстановления. Процедуры для быстрого восстановления работоспособности

Модуль 10. Обеспечение безопасной работы пользователей с нейронными сетями

- Проверка и валидация входных данных. Правила безопасных запросов к нейросетям
- Ограничение доступа пользователей к чувствительной информации. Разделение прав доступа и управление полномочиями
- Предотвращение утечки данных. Защита от несанкционированного копирования и распространения данных
- Обучение пользователей правилам безопасной работы с нейронными сетями

Модуль 11. Нормативно-правовое регулирование MLSecOps. Этичное применение ИИ на предприятиях

- Международные стандарты регулирования ИИ
- Российские стандарты регулирования ИИ
- Этичное и ответственное использование ИИ в различных отраслях. Соблюдение комплаенс-политик и процедур. Предотвращение предвзятости и дискриминации
- Формирование культуры MLSecOps в организации. Корпоративная политика по безопасности ИИ-систем

Модуль 12. Обзор лучших практик и перспектив MLSecOps

- Обзор современных инструментов и технологий MLSecOps. Анализ передовых разработок в области безопасности
- Автоматизация задач MLSecOps. Применение инструментов для автоматизации рутинных операций
- Перспективы развития MLSecOps и LLMSecOps. Обзор будущих тенденций и направлений развития
- Практические кейсы и примеры внедрения MLSecOps. Анализ успешных проектов и лучших практик

Посмотреть расписание курса и записаться на обучение

Обращайтесь по любым вопросам
к менеджерам Академии АйТи

+7 (495) 150 96 00 | academy@academyit.ru