



Академия АйТи
a Softline Company



Специалист Data Science

Код курса: pp_ds

Специалист Data Science

Код курса: pp_ds

Длительность	256 ак. часов
Формат	
Разработчик курса	Академия АйТи
Тип	Учебный курс
Способ обучения	Под руководством тренера

О курсе

Хотите стать экспертом в области искусственного интеллекта и анализа данных? Представляем программу профессиональной переподготовки "Data Science специалист" (256 часов) – ваш путь к освоению одной из самых востребованных профессий современности. Программа разработана для тех, кто стремится освоить полный спектр компетенций в области Data Science: от фундаментальных основ до продвинутых технологий искусственного интеллекта. Независимо от того, являетесь ли вы программистом, аналитиком или специалистом из смежной области – наш курс поможет вам приобрести необходимые навыки для успешной карьеры в сфере данных и машинного обучения. В ходе обучения вы освоите современный стек технологий, включая Python и его ключевые библиотеки для анализа данных, алгоритмы машинного обучения, нейронные сети и методы глубокого обучения. Научитесь работать с большими данными, создавать системы компьютерного зрения и обработки естественного языка, а также познакомитесь с последними достижениями в области искусственного интеллекта, включая работу с большими языковыми моделями. Особое внимание в программе уделяется практическому применению полученных знаний. 70% времени обучения посвящено работе над реальными проектами, что позволит вам сформировать внушительное профессиональное портфолио. Вы научитесь проводить полный цикл разработки ML-проектов: от сбора и анализа данных до развертывания моделей в промышленную эксплуатацию. Программа включает изучение современных инструментов и методологий MLOps, что позволит вам эффективно внедрять ML-решения в производственную среду. Вы освоите автоматизацию ML-пайплайнов, научитесь работать с системами версионирования данных и моделей, познакомитесь с лучшими практиками развертывания AI-систем. По окончании обучения вы будете готовы к работе в роли Data Scientist, ML Engineer или AI Developer в ведущих компаниях. Наши выпускники успешно решают задачи прогнозирования, оптимизации бизнес-процессов, создания рекомендательных систем и автоматизации принятия решений. Обучение проводится в удобном онлайн-формате с возможностью совмещения с работой. Программа включает индивидуальные консультации с экспертами, работу в команде над реальными проектами и подготовку к собеседованиям в ведущие компании. Инвестируйте в свое будущее – станьте специалистом в одной из самых перспективных областей современных технологий. Количество мест ограничено. Запишитесь на программу уже сегодня и сделайте первый шаг к успешной карьере в Data Science! Особенности курса: Углубленная фундаментальная подготовка: расширенный блок математики и программирования больше внимания базовым концепциям Практическая ориентированность: каждый модуль включает практические задания работа с реальными датасетами кейсстади из индустрии Современные технологии: актуальные инструменты и

фреймворки облачные технологии работа с большими данными Проектная работа: сквозной проект через весь курс портфолио проектов командная работа Дополнительные активности: консультации с экспертами кодревью хакатоны Итоговая аттестация: защита финального проекта теоретический экзамен практическое задание

Подробная информация

Профиль аудитории:

- программа предназначена для аналитиков, разработчиков, IT менеджеров и специалистов по бизнесу, стремящихся использовать искусственный интеллект для оптимизации процессов и повышения эффективности компании. Курс подойдет как для профессионалов, уже имеющих базовые знания Python, так и для тех, кто хочет освоить современные методы создания ИИ моделей для бизнес-приложений

Цель:

- целевая аудитория курса включает специалистов с базовыми знаниями Python, аналитиков данных, желающих перейти в Data Science, программистов, стремящихся освоить анализ данных, студентов технических специальностей и специалистов, желающих сменить профессию

Предварительные требования

- базовое знание Python
- основы математической статистики
- базовое понимание SQL
- английский язык на уровне чтения технической документации

После успешного завершения программы слушатель будет:

1. знать:

- фундаментальные концепции и методологии Data Science
- математические основы машинного обучения и статистики
- принципы работы алгоритмов машинного обучения и нейронных сетей
- методы обработки и анализа больших данных
- современные инструменты и технологии в области Data Science
- принципы MLOps и промышленной разработки ML систем
- этические аспекты работы с данными и AI

2. уметь:

работа с данными:

- проводить сбор, очистку и предобработку данных различных типов
- выполнять комплексный разведочный анализ данных (EDA)
- создавать эффективные ETL пайплайны
- работать с различными форматами данных и базами данных
- применять методы feature engineering

анализ и моделирование:

- выбирать оптимальные алгоритмы для решения конкретных задач
- обучать и оптимизировать модели машинного обучения
- проводить оценку качества моделей и их валидацию
- интерпретировать результаты моделей
- применять методы глубокого обучения для различных задач

разработка и внедрение:

- создавать масштабируемые ML-решения
- разрабатывать API для ML-моделей
- применять практики CI/CD в ML-проектах
- осуществлять мониторинг и поддержку ML-систем
- оптимизировать производительность ML-решений

3. владеть:

технический стек:

- Python и его основные библиотеки для DS (NumPy, Pandas, Scikit-learn)
- фреймворки глубокого обучения (PyTorch)
- инструменты визуализации (Matplotlib, Seaborn, Plotly)
- системы контроля версий (Git)
- инструменты MLOps (DVC, MLflow)
- SQL и NoSQL базы данных
- Docker и Kubernetes
- Cloud платформы для ML

Soft Skills:

- навыки решения комплексных бизнес-задач
- способность работать в команде
- навыки презентации результатов
- умение писать техническую документацию
- критическое мышление и аналитические способности

Практические навыки:

1. проектирование ML-систем:

- разработка архитектуры ML-решений
- проектирование пайплайнов обработки данных
- создание систем автоматического обучения моделей

2. работа с различными типами задач:

- классификация и регрессия
- кластеризация и снижение размерности
- обработка естественного языка
- компьютерное зрение
- анализ временных рядов

3. промышленная разработка:

- создание воспроизводимых исследований
- написание чистого и эффективного кода

- работа с системами контроля версий
- развертывание моделей в production

Программа курса

Модуль 1: Фундаментальные основы Data Science

Блок 1.1: Введение в Data Science и программирование

- Основы Data Science и его роль в современном мире
- Python для Data Science:
 - Основы синтаксиса и структуры данных
 - Функциональное программирование
 - ООП в Python
 - Git и командная работа
 - Работа с IDE и Jupyter notebooks

Блок 1.2: Математические основы

Линейная алгебра для DS:

- Векторы и матрицы
- Собственные значения и векторы
- Математический анализ:
 - Производные и градиенты
 - Оптимизация функций
 - Теория вероятностей

Блок 1.3: Инструменты работы с данными

NumPy углубленно:

Pandas продвинутый уровень:

- Сложные трансформации данных
- Оптимизация производительности
- SQL и базы данных:
 - Реляционные БД и сложные запросы
 - NoSQL решения

Модуль 2: Обработка и анализ данных

Блок 2.1: Продвинутая работа с данными

ETL процессы и pipeline:

- Архитектура ETL систем
- Apache Airflow
- Работа с большими данными:
 - PySpark основы

- Distributed Computing

Блок 2.2: Продвинутая визуализация

Статическая визуализация:

- Matplotlib углубленно
- Seaborn продвинутый уровень
- Интерактивная визуализация:
 - Plotly и Dash
 - Streamlit разработка

Блок 2.3: Статистический анализ

- Описательная статистика углубленно
- Статистические тесты и их применение
- Байесовская статистика
- Продвинутое A/B тестирование

Модуль 3: Классическое машинное обучение

Блок 3.1: Supervised Learning углубленно

- Линейные модели:
 - Продвинутая регрессия
 - Регуляризация и оптимизация
- Ансамблевые методы:
 - Random Forest углубленно
 - Gradient Boosting детально
 - SVM и kernel методы

Блок 3.2: Unsupervised Learning продвинутый

- Кластеризация:
 - Классические алгоритмы углубленно
 - Спектральная кластеризация
- Понижение размерности:
 - PCA и tSNE углубленно
 - UMAP и автоэнкодеры

Блок 3.3: Оценка и оптимизация моделей

- Продвинутые метрики
- Кроссвалидация и бутстрап
- Оптимизация гиперпараметров

Модуль 4: Deep Learning и нейронные сети

Блок 4.1: Основы Deep Learning

- Архитектура нейронных сетей

- Функции активации и оптимизаторы
- PyTorch углубленно

Блок 4.2: Компьютерное зрение

CNN архитектуры детально

- Object Detection
- Segmentation
- GANs основы

Блок 4.3: Обработка последовательностей

RNN и LSTM углубленно

- Attention механизмы
- Time Series прогнозирование

Модуль 5: Natural Language Processing

Блок 5.1: Классический NLP

- Предобработка текста углубленно
- Word Embeddings детально
- Topic Modeling

Блок 5.2: Современный NLP

Трансформеры архитектура

- BERT и его варианты
- LLM применение и файнтюнинг

Модуль 6: MLOps и производственное внедрение

Блок 6.1: MLOps практики

- ML Pipeline автоматизация
- Версионирование данных и моделей
- CI/CD для ML проектов
- Мониторинг и логирование

Блок 6.2: Развертывание моделей

REST API разработка (Flask/FastAPI)

- Контейнеризация с Docker
- Kubernetes для ML
- Cloud платформы для ML

[Посмотреть расписание курса и записаться на обучение](#)

Обращайтесь по любым вопросам

к менеджерам Академии АйТи

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