Бауманский учебный центр «Специалист»

Программа курса Оркестрация ETL - процессов в Apache Airflow для инженеров данных

Код курса: **ЭИРФЛО**

Тема	Ак. часов
Модуль 1. Знакомство с Airflow • Что такое ETL. • Оркестрация вручную - Cron и скрипты. • Введение в AirFlow. Предназначение. • Архитектура Airflow (Scheduler, Executor, Worker). • Локальная установка AirFlow. • Установка в Docker. • Настройка подключений к внешним сервисам. • Создание пользователей. • Ключевые параметры конфигурации. • **Практика 1.** Установка и получение доступа к AirFlow.	**2**
Модуль 2. MLOps: оркестрация для машинного обучения • Оркестрация для машинного обучения. Роль Airflow в MLOps пайплайнах. • Архитектура ML пайплайна в Airflow: компоненты и зависимости. • **Практика 2.** Начало работы над сквозным проектом: настройка подключения к сервису ML модели.	**2**
Модуль 3. Создание первого задания • Что такое DAG, из чего он состоит? • Обзор веб-интерфейса. • Backfill & catchup. • Основные сущности (Таски, Хуки, операторы). • Процесс исполнения дага. • Планирование, ручной запуск и мониторинг выполнения заданий. • Отслеживание статуса выполнения задач. • **Практика 3.** Продолжение работы над сквозным проектом: создание DAG для подключения к API.	**4**
Модуль 4. Виды операторов • Виды операторов. • PythonOperator. • BashOperator. • HttpOperator. • SQLOperator. • Сенсоры. • Хуки. • **Практика 4.** Продолжение работы над сквозным проектом: описание простых заданий.	**4**
Модуль 5. Запуск задач в контейнерах • Знакомство с DockerOperator. • **Практика 5.** Продолжение работы над сквозным проектом: создание задачи и запуск в Docker-контейнере.	**4**

Модуль 6. Зависимости между задачами • Взаимодействие заданий (XCom). • Последовательное выполнение. • Параллельное выполнение. • Проверка условий (BranchOperator). • Передача данных между операторами: переменные. • Шаблоны и макросы. • **Практика 6.** Продолжение работы над сквозным проектом: создание ETL-процесса с передачей данных через XCom.	**4**
Модуль 7. Написание новых операторов, хуков, сенсоров • Написание оператора. • Написание хука. • Написание сенсора. • **Практика 7.** Продолжение работы над сквозным проектом: написание собственных сущностей.	**4**
Модуль 8. Тестирование и оптимизация заданий • Оптимизация заданий. • Динамические задания. • Управление параллелизмом и пулингом. • Тестирование заданий. • Плагины. • **Практика 8.** Продолжение работы над сквозным проектом: тестирование и оптимизация заданий.	**4**
Модуль 9. Мониторинг и алерты • Мониторинг. • Логирование. • Уведомления. • **Практика 9.** Продолжение работы над сквозным проектом: сбор метрик, отправка логов и создание алерта.	**4**
Аудиторная нагрузка в классе с преподавателем	**32 ак.ч. + 32 ак.ч. самостоятельно**

По окончании обучения на курсе проводится итоговая аттестация. Аттестация проводится в виде теста на последнем занятии или на основании оценок практических работ, выполняемых во время обучения на курсе.

Контакты центра

Для частных лиц
+7 (495) 232-32-16

Для организаций
+7 (495) 780-48-44

E-mail
info@specialist.ru