

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 15.08.2024 17:26:39

Уникальный программный ключ:

c098bc0c1041cb2a4cf926cf171d6715d99a

АННОТАЦИЯ рабочей программы дисциплины

Системы аналитики больших данных

1. Общая трудоёмкость

Трудоёмкость дисциплины составляет 5 зачётных единиц (180 часов), из них 18 часов лекционных занятий, 36 часов практических занятий.

2. Место дисциплины в структуре образовательной программы

Дисциплина относится к модулю обязательных профессиональных дисциплин обязательной части образовательной программы.

Для изучения данной дисциплины необходимы знания, умения и навыки, формируемые предшествующими элементами образовательной программы: Методология научной деятельности; Математические методы анализа больших данных; Методология проектирования и управления информационными системами.

Результаты обучения, формируемые данной дисциплиной, потребуются при освоении следующих элементов образовательной программы: производственная практика, проектно-технологическая практика; производственная практика, преддипломная практика.

3. Цель изучения дисциплины

формирование у обучающихся способности осуществлять руководство проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных с применением новых методов и алгоритмов машинного обучения.

4. Содержание дисциплины

Модуль 1. Основы построения и работы с системами аналитики больших данных

Архитектура систем аналитики больших данных.

Хранение больших данных в облаке.

Хранилища общего назначения. Форматы хранения данных. Облачное хранилище Microsoft Azure Storage. Облачное хранилище AWS. Реляционные базы данных. Azure SQL. AWS RDS. Нереляционные базы данных. Сервисы нереляционных баз данных от Azure и AWS. Корпоративные хранилища данных (DWH). Azure SQL DWH. AWS RedShift. Хранилища данных типа Data Lake. Azure Data Lake Store. AWS Data Lake Solutions.

Модуль 2. Разработка систем аналитики больших данных

Доставка больших данных в облако.

Прямая загрузка данных. Доставка данных в облачное хранилище общего назначения. Доставка данных в реляционные БД и хранилища. Доставка данных в нереляционные базы данных. Доставка данных в HDFS-совместимые хранилища. Прямая загрузка потоковых данных. Azure Event Hub. AWS Kinesis Data Streams. Облачные сервисы развертывания кластерных систем. Облачные сервисы копирования и трансформации данных. Azure Data Factory. AWS Data Pipeline. AWS Glue.

Аналитика больших данных в облаке.

Интерактивный анализ данных. Анализ реляционных данных. Azure Data Lake Analytics. AWS Athena. Apache Spark. Встроенные редакторы запросов сервиса CosmosDB. Поточковый анализ данных. Общие сведения. Azure Stream Analytics. Amazon Kinesis Analytics. Apache Storm. Пакетный анализ данных. Hadoop. Apache Pig. Apache Hive.

5. Дополнительная полезная информация

Дисциплина предназначена для формирования элементов следующих компетенций образовательной программы:

ПК-2. Способен руководить проектами по созданию систем искусственного интеллекта с применением новых методов и алгоритмов машинного обучения со стороны заказчика.

ПК-4. Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика.

Форма промежуточной аттестации: экзамен.

Наименование оценочного средства: собеседования по выполнению практических работ №1-3.