

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Исаев Игорь Магомедович
Должность: Проректор по учебной работе
Дата подписания: 31.07.2023 14:01:17
Уникальный идентификатор документа:
d7a26b9e8ca85e98ec3de2eb454b4659d061f249

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение
высшего образования

«Национальный исследовательский технологический университет «МИСИС»

Аннотация рабочей программы учебной дисциплины

Программные платформы и технологии больших данных

Закреплена за подразделением Кафедра магистерская школа информационных бизнес систем

Направление подготовки 09.04.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Профиль Экосистема больших данных для цифровой трансформации

Квалификация **Магистр**

Форма обучения **очная**

Общая трудоемкость **4 ЗЕТ**

Часов по учебному плану 144

в том числе:

аудиторные занятия 52

самостоятельная работа 56

часов на контроль 36

Формы контроля в семестрах:
экзамен 2

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	2 (1.2)		Итого	
	18			
Неделя	18			
Вид занятий	уп	рп	уп	рп
Лекции	16	16	16	16
Практические	36	36	36	36
В том числе инт.	36		36	
Итого ауд.	52	52	52	52
Контактная работа	52	52	52	52
Сам. работа	56	56	56	56
Часы на контроль	36	36	36	36
Итого	144	144	144	144

1. ЦЕЛИ ОСВОЕНИЯ

1.1	Целями освоения дисциплины являются формирование у студентов исчерпывающего представления о подходах и методах хранения, организации и обработки больших объемов данных, а также получение практических навыков по выполнению проектов в области извлечения данных и анализа данных полученных данных в интересах организации.
1.2	Дисциплина основана на открытых технологиях и вендор-нейтральных стандартах, в качестве иллюстрации рассматриваемых концепций и продукты лидеров индустрии, таких как - Dell Technologies и д.р.
1.3	В результате изучения дисциплины обучающийся должен сформировать представление о понятии «Большие Данные» (Big Data) и роли специалиста по обработке данных (Data Scientist) в проектах, связанных с большими данными, о некоторых инструментах, входящих в экосистему Hadoop (Pig, Hive, HBase и т.п.), которые могут быть полезны в аналитических проектах; изучить способы хранения и организации данных в современной компании, структуру и основные этапы жизненного цикла проектов, связанных с обработкой больших данных, основные средства статистического анализа данных, которые могут применяться в современной компании, и основ работы в них (на примере пакета R), технологию MapReduce и экосистему Hadoop, применяемых для хранения и анализа данных большого объема, базовые алгоритмы, связанные с анализом данных, такие как кластеризация (k-mean, иерархическая), извлечение правил ассоциаций, регрессионного анализа (линейная и логистическая регрессии), алгоритмы основные на вероятностных графовых моделях (в том числе Наивный Байес), деревья принятия решений, анализ временных рядов (ARMA, ARIMA и т.п.), алгоритмы и инструменты анализа текстовой информации (регулярные выражения, метрики релевантности и т.п.), а также овладеть соответствующим инструментарием из пакета R и библиотеки MADlib.

2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Блок ОП:		Б1.В
2.1	Требования к предварительной подготовке обучающегося:	
2.1.1	Архитектура бизнеса	
2.1.2	Информационные технологии в офисной деятельности	
2.1.3	Новые направления и технологии современных СУБД	
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:	
2.2.1	Интеллектуальный анализ данных	
2.2.2	Научно-исследовательская работа. Исследовательский проект	
2.2.3	Практика машинного обучения. Дизайн новых материалов	
2.2.4	Практика машинного обучения. Материаловедение	
2.2.5	Прикладные области анализа больших данных. Дизайн новых материалов	
2.2.6	Прикладные области анализа больших данных. Материаловедение	
2.2.7	СУБД. Продвинутый уровень SQL	
2.2.8	Научно-исследовательская работа	
2.2.9	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.10	Производственная проектная практика	

3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ

ПК-4: Способен выполнять работы по сбору, обработке и анализу больших данных в междисциплинарных областях
Знать:
ПК-4-32 Типы больших данных: метаданные, полуструктурированные, структурированные, неструктурированные данные
ПК-4-31 Теоретические и прикладные основы анализа больших данных
ПК-4-34 Содержание и последовательность выполнения этапов аналитического проекта
ПК-4-33 Современные методы и инструментальные средства анализа больших данных
УК-1: Способен осуществлять критический анализ новых и сложных инженерных объектов, процессов и систем в междисциплинарном контексте, проблемных ситуаций на основе системного подхода, выбрать и применить наиболее подходящие и актуальные методы из существующих аналитических, вычислительных и экспериментальных методов или новых и инновационных методов, выработать стратегию действий
Знать:
УК-1-31 Решаемые в процессе анализа данных задачи

ПК-4: Способен выполнять работы по сбору, обработке и анализу больших данных в междисциплинарных областях
Уметь:
ПК-4-У1 Оформлять результаты аналитического исследования для представления заказчику
УК-1: Способен осуществлять критический анализ новых и сложных инженерных объектов, процессов и систем в междисциплинарном контексте, проблемных ситуаций на основе системного подхода, выбрать и применить наиболее подходящие и актуальные методы из существующих аналитических, вычислительных и экспериментальных методов или новых и инновационных методов, вырабатывать стратегию действий
Уметь:
УК-1-У1 Планировать работы по выполнению проектов, связанных с анализом данных
Владеть:
УК-1-В2 Навыками использования инструментария для анализа данных (статистические пакеты и т.п.), в том числе в рамках современных парадигм обработки данных больших объемов (map-reduce и т.п.)
УК-1-В1 Навыками использования инструментария для извлечения данных из различных источников (БД, публичные web-сервисы и т.п.)