

Документ подписан простой электронной подписью
 Информация о владельце:
 ФИО: Косенок Сергей Михайлович
 Должность: ректор
 Дата подписания: 18.09.2026 08:36:05
 Уникальный программный ключ:
 e3a68f3eaa1e62674b54f4998099d3d6bfdcf836

Бюджетное учреждение высшего образования
 Ханты-Мансийского автономного округа-Югры
 "Сургутский государственный университет"

УТВЕРЖДАЮ
 Проректор по УМР

_____ Е.В. Коновалова

17 сентября 2026 г., протокол УМС №6

Большие данные

рабочая программа дисциплины (модуля)

Закреплена за кафедрой	Менеджмента и бизнеса
Учебный план	g380405-БизИнфор-27-1.plx 38.04.05 Бизнес-информатика Направленность (профиль): Управление технологическими изменениями
Квалификация	Магистр
Форма обучения	очная
Общая трудоемкость	7 ЗЕТ
Часов по учебному плану	252
в том числе:	
аудиторные занятия	32
самостоятельная работа	184
часов на контроль	36
Виды контроля в семестрах:	
контрольная работа 3	
экзамен 3	

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	3 (2.1)		Итого	
Неделя	17 1/6			
Вид занятий	уп	рп	уп	рп
Лекции	16	16	16	16
Лабораторные	16	16	16	16
Итого ауд.	32	32	32	32
Контактная работа	32	32	32	32
Сам. работа	184	184	184	184
Часы на контроль	36	36	36	36
Итого	252	252	252	252

Программу составил(и):

к. ф-м. н., доцент, Лысенкова Светлана Александровна

Рабочая программа дисциплины

Большие данные

разработана в соответствии с ФГОС:

Федеральный государственный образовательный стандарт высшего образования - магистратура по направлению подготовки 38.04.05 Бизнес-информатика (приказ Минобрнауки России от 12.08.2020 г. № 990)

составлена на основании учебного плана:

38.04.05 Бизнес-информатика

Направленность (профиль): Управление технологическими изменениями

утвержденного учебно-методическим советом вуза от 17 сентября 2026 г., протокол УМС №6.

Рабочая программа одобрена на заседании кафедры

Информатики и вычислительной техники

Зав. кафедрой к. ф-м. н., доцент, Лысенкова Светлана Александровна

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1	Сформировать у студентов системные знания, практические умения и навыки работы с большими данными, включая сбор, хранение, обработку и анализ.
-----	--

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП

Цикл (раздел) ООП:		Б1.В.ДВ.01
2.1	Требования к предварительной подготовке обучающегося:	
2.1.1	Основы научных исследований в области бизнес-информатики	
2.1.2	Аналитика данных и искусственный интеллект	
2.1.3	Управление ИТ-проектами, инновациями и технологическими изменениями	
2.2	Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:	
2.2.1	Машинное обучение	
2.2.2	Производственная практика, преддипломная практика	

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)**ПК-2.2: Интегрирует компоненты и сервисы в единую систему****ПК-2.3: Моделирует потоки данных и событий между информационно-аналитическими системам****В результате освоения дисциплины обучающийся должен**

3.1	Знать:
3.1.1	- основные понятия, характеристики и принципы работы с большими данными;
3.1.2	- признаки больших данных: объем, скорость поступления, разнообразие, достоверность и ценность данных;
3.1.3	- основные источники больших данных и форматы их представления;
3.1.4	- методы сбора, хранения, обработки и анализа больших данных;
3.1.5	- основы работы с реляционными и нереляционными базами данных;
3.1.6	- принципы построения распределенных систем хранения и обработки данных;
3.1.7	- назначение технологий и инструментов обработки больших данных, включая Hadoop, Spark, NoSQL-СУБД;
3.1.8	- основные требования к обеспечению безопасности данных при их хранении, обработке и передаче;
3.1.9	- принципы интеграции программных модулей и компонентов при разработке информационных систем, работающих с большими данными.
3.2	Уметь:
3.2.1	- определять источники данных для решения прикладных задач в области информационных систем;
3.2.2	- выполнять сбор, предварительную обработку и структурирование данных;
3.2.3	- выбирать подходящие методы и инструменты хранения и обработки больших данных;
3.2.4	- применять средства анализа данных для выявления закономерностей и получения полезной информации;
3.2.5	- использовать базы данных и программные средства для обработки больших массивов информации;
3.2.6	- проектировать простые схемы обработки данных в информационных системах;
3.2.7	- выполнять интеграцию программных компонентов, связанных с обработкой и хранением данных;
3.2.8	- оценивать качество данных и выявлять ошибки, дублирование и неполноту информации;

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература	Примечание
-------------	---	----------------	-------	-------------	------------	------------

	Раздел 1. Раздел 1. Введение в Большие данные. Анализ Больших данных					
1.1	Введение в большие данные. Модель 5V/7V /Лек/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
1.2	Знакомство с инструментами. Установка и настройка окружения (Anaconda + Jupyter + PySpark). Работа с первыми датасетами /Лаб/	3	4	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
1.3	Экосистема Hadoop. Архитектура HDFS, принципы распределённого хранения, отказоустойчивость. /Лек/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
1.4	Работа с HDFS: загрузка, копирование, удаление данных. /Лаб/	3	4	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
1.5	Модель программирования MapReduce. Примеры задач, ограничения и причины появления Spark. /Лек/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
1.6	Реализация алгоритма WordCount с помощью MapReduce /Лаб/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
1.7	Работа с RDD: создание, трансформации, действия. Кэширование. /Лаб/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
1.8	Apache Spark как универсальный движок. RDD API, DAG, планировщик задач /Лек/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
	Раздел 2. Раздел 2. Хранение больших данных и распределённая обработка					
2.1	Структурированная обработка в Spark. DataFrame, SQL, оптимизатор Catalyst, Tungsten /Лек/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
2.2	Работа с DataFrame: загрузка данных (CSV, Parquet, JSON), очистка, фильтрация, агрегация /Лаб/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
2.3	Потоковая обработка Spark. Интеграция с Kafka /Лек/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
2.4	Настройка Kafka + Spark Streaming. Обработка потока логов в реальном времени. /Лаб/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
	Раздел 3. Раздел3. Машинное обучение на больших данных					
3.1	Масштабируемое машинное обучение. MLlib, алгоритмы, распределённое обучение, оценка моделей /Лек/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
3.2	Построение и оценка моделей MLlib на большом датасете. Pipeline API /Лаб/	3	0	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
3.3	Современный стек: Lakehouse-архитектура, Delta Lake, Iceberg. Безопасность, мониторинг, российское законодательство /Лек/	3	2	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
3.4	Итоговый мини-проект: разработка полного пайплайна обработки больших данных. Защита проекта /Ср/	3	184	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	
	Раздел 4.					

4.1	Контрольная работа /Контр.раб./	3	0	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	Задание на контрольную работу
4.2	/Экзамен/	3	36	ПК-2.2 ПК-2.3	Л1.1Л2.1Л3.1 Э1 Э2 Э3	Задание на экзамен

5. ОЦЕНОЧНЫЕ СРЕДСТВА

5.1. Оценочные материалы для текущего контроля и промежуточной аттестации

Представлены отдельным документом

5.2. Оценочные материалы для диагностического тестирования

Представлены отдельным документом

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л1.1	Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н.	Большие данные. Big Data: учебник для вузов	Санкт-Петербург: Лань, 2023, электронный ресурс	1

6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л2.1	Ланских Ю. В., Ланских В. Г., Родионов К. В.	Введение в большие данные: учеб. пособие	Киров: ВятГУ, 2023, электронный ресурс	1

6.1.3. Методические разработки

	Авторы, составители	Заглавие	Издательство, год	Колич-во
Л3.1	Кучин И. А.	Технологии обработки и анализа больших данных: учебно-методическое пособие	Сургут: Издательский центр СурГУ, 2025, электронный ресурс	1

6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Документация Apache hadoop hadoop.apache.org
Э2	Документация Apache spark spark.apache.org
Э3	Документация Jupyter docs.jupyter.org

6.3.1 Перечень программного обеспечения

6.3.1.1	Операционная система РЕД ОС №03872000223220002740001 от «25» ноября 2022 г., бессрочно. Программное обеспечение SCAD Office S64 договор 2014.302750 от 20.10.2014 г., неисключительные права (лицензия) на неограниченный период. Программное обеспечение ГИС MapInfoProfessional 01872000017210014450001 от 01.10.2021 г., бессрочно. Графические пакеты CS5 AdobeDesignPremium 5, CorelDRAWGraphicsSuiteX6, бессрочно. Среда разработки EmbarcaderoDelphi, EmbarcaderoC++Builder 2010, договор 123/11-ГК от 12.12.2011 г., бессрочно. Программное обеспечение ABBYY FineReader 14 Профессиональная, неисключительные права (лицензия) на неограниченный период.
6.3.1.4	Свободно распространяемое ПО: LibreOffice; Яндекс браузер; Графические редакторы - Gimp, Inscapе; Математическое ПО - Freemat, Maxima; CAD системы - NanoCAD2, KiCAD, Scilab; Средства разработки - Python, Lazarus, Java, Notepad++, free pascal.

6.3.2 Перечень информационных справочных систем

6.3.2.1	СПС «Гарант». – режим доступа: http://www.garant.ru/ СПС «КонсультантПлюс». – режим доступа: http://www.consultant.ru/ АИС "Студент СурГУ" - режим доступа: https://student.surgu.ru/
---------	---

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)	
7.1	Учебная аудитория – компьютерный класс для проведения лекционных, практических (лабораторных) занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. Номер помещения в соответствии с документами бюро технической инвентаризации: № 503. Перечень основного оборудования: комплект специализированной учебной мебели, меловая доска. Количество посадочных мест – 12. Технические средства обучения для представления учебной информации: комплект мультимедийного оборудования, компьютеры (13 шт.), проектор, проекционный экран. Используемое программное обеспечение: Операционная система РЕД ОС №03872000223220002740001 от «25» ноября 2022 г., бессрочно. Программное обеспечение SCAD Office S64 договор 2014.302750 от 20.10.2014 г., неисключительные права (лицензия) на неограниченный период. Программное обеспечение ГИС MapInfoProfessional 01872000017210014450001 от 01.10.2021 г., бессрочно. Графические пакеты CS5 AdobeDesignPremium 5, CorelDRAWGraphicsSuiteX6, бессрочно. Среда разработки EmbarcaderoDelphi, EmbarcaderoC++Builder 2010, договор 123/11-ГК от 12.12.2011 г., бессрочно. Программное обеспечение ABBYY FineReader 14 Профессиональная, неисключительные права (лицензия) на неограниченный период. Свободно распространяемое ПО: LibreOffice; Яндекс браузер; Графические редакторы - Gimp, Inscapе; Математическое ПО - Freemat, Maxima; CAD системы - NanoCAD2, KiCAD, Scilab; Средства разработки - Python, Lazarus, Java, Notepad++, free pascal. Обеспечен доступ к сети Интернет и в электронную информационную среду организации.
7.2	Учебная аудитория - компьютерный класс для проведения лекционных, практических (лабораторных) занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. Номер помещения в соответствии с документами бюро технической инвентаризации: № 605. Перечень основного оборудования: комплект специализированной учебной мебели, меловая доска. Количество посадочных мест – 12. Технические средства обучения для представления учебной информации: комплект (стационарный/переносной) мультимедийного оборудования, компьютеры (13 шт.), проектор, проекционный экран. Используемое программное обеспечение: Операционная система РЕД ОС №03872000223220002740001 от «25» ноября 2022 г., бессрочно. Программное обеспечение SCAD Office S64 договор 2014.302750 от 20.10.2014 г., неисключительные права (лицензия) на неограниченный период. Программное обеспечение ГИС MapInfoProfessional 01872000017210014450001 от 01.10.2021 г., бессрочно. Графические пакеты CS5 AdobeDesignPremium 5, CorelDRAWGraphicsSuiteX6, бессрочно. Среда разработки EmbarcaderoDelphi, EmbarcaderoC++Builder 2010, договор 123/11-ГК от 12.12.2011 г., бессрочно. Программное обеспечение ABBYY FineReader 14 Профессиональная, неисключительные права (лицензия) на неограниченный период. Свободно распространяемое ПО: LibreOffice; Яндекс браузер; Графические редакторы - Gimp, Inscapе; Математическое ПО - Freemat, Maxima; CAD системы - NanoCAD2, KiCAD, Scilab; Средства разработки - Python, Lazarus, Java, Notepad++, free pascal. Обеспечен доступ к сети Интернет и в электронную информационную среду организации.
7.3	Помещение для самостоятельной работы обучающихся – Зал экономической и юридической литературы. Номер помещения в соответствии с документами бюро технической инвентаризации: № 47. Перечень основного оборудования: комплект специализированной учебной мебели. Количество посадочных мест – 30. Технические средства обучения для представления учебной информации: компьютер – 5 шт. Используемое программное обеспечение: Операционная система РЕД ОС №03872000223220002740001 от «25» ноября 2022 г., бессрочно. Свободно распространяемое ПО: LibreOffice; Яндекс браузер. Обеспечен доступ к сети Интернет и в электронную информационную среду организации.