

МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ))



Рабочая программа дисциплины (модуля),
как компонент образовательной программы
базового высшего образования
по направлению подготовки
38.03.05 Бизнес-информатика,
утвержденной первым проректором РУТ (МИИТ)
Тимониным В.С.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Алгоритмы работы с большими данными

Направление подготовки: 38.03.05 Бизнес-информатика

Направленность (профиль): Цифровая экономика

Форма обучения: Очная

Рабочая программа дисциплины (модуля) в виде
электронного документа выгружена из единой
корпоративной информационной системы управления
университетом и соответствует оригиналу

Простая электронная подпись, выданная РУТ (МИИТ)
ID подписи: 564169
Подписал: заведующий кафедрой Каргина Лариса Андреевна
Дата: 11.06.2026

1. Общие сведения о дисциплине (модуле).

Целями освоения дисциплины являются:

- развитие общематематической культуры, логического и алгоритмического мышления студентов;
- выработка умения моделировать реальные финансово-экономические процессы;
- освоение приемов исследования и решения математически формализованных задач.

Задачами освоения дисциплины является формирование у обучающихся навыков:

- получение необходимого математического аппарата для изучения дисциплин естественнонаучного и профессионального циклов и применения этого аппарата в будущей профессиональной деятельности.

2. Планируемые результаты обучения по дисциплине (модулю).

Перечень формируемых результатов освоения образовательной программы (компетенций) в результате обучения по дисциплине (модулю):

ПК-3 - Способен проводить моделирование, анализ и совершенствование бизнес-процессов и информационно-технологической инфраструктуры предприятия в интересах достижения его стратегических целей с использованием современных методов и программного инструментария;

ПК-8 - Способен осуществлять контроль функционирования, анализ показателей результативности и эффективности функционирования информационной системы.

Обучение по дисциплине (модулю) предполагает, что по его результатам обучающийся будет:

Знать:

- современные методы и программный инструментарий для моделирования, анализа и совершенствования бизнес-процессов и информационно-технологической инфраструктуры при работе с большими данными;

- принципы контроля функционирования, показатели результативности и эффективности функционирования информационных систем, обрабатывающих большие объемы данных.

Уметь:

- проводить моделирование, анализ и совершенствование бизнес-процессов и информационно-технологической инфраструктуры предприятия с применением алгоритмов работы с большими данными;
- осуществлять контроль функционирования, анализировать показатели результативности и эффективности функционирования информационных систем в среде больших данных.

Владеть:

- навыками моделирования, анализа и совершенствования бизнес-процессов и информационно-технологической инфраструктуры с применением современных методов и алгоритмического инструментария для больших данных;
- навыками контроля функционирования, анализа показателей результативности и эффективности информационных систем при обработке больших объемов данных.

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 4 з.е. (144 академических часа(ов)).

3.2. Объем дисциплины (модуля) в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Тип учебных занятий	Количество часов	
	Всего	Семестр №7
Контактная работа при проведении учебных занятий (всего):	64	64
В том числе:		
Занятия лекционного типа	32	32
Занятия семинарского типа	32	32

3.3. Объем дисциплины (модуля) в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации составляет 80 академических часа (ов).

3.4. При обучении по индивидуальному учебному плану, в том числе при ускоренном обучении, объем дисциплины (модуля) может быть реализован

полностью в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации.

4. Содержание дисциплины (модуля).

4.1. Занятия лекционного типа.

№ п/п	Тематика лекционных занятий / краткое содержание
1	Введение в «большие данные» Рассматриваемые вопросы: - понятие «большие данные»: что это такое и с чем это связано; - характеристики больших данных: объем, разнообразие, скорость.
2	Введение в «большие данные» Рассматриваемые вопросы: - классические подходы; - подходы, ориентированные на данные.
3	Введение в «большие данные» Рассматриваемые вопросы: - классические подходы; - подходы, ориентированные на данные.
4	Введение в «большие данные» Рассматриваемые вопросы: - основные технологии работы с большими данными; - основные инструменты работы с большими данными.
5	Введение в «большие данные» Рассматриваемые вопросы: - принципы работы с большими данными; - сферы применения больших данных.
6	Технологии, связанные с большими данными Рассматриваемые вопросы: - уровни в системах обработки данных; - сбор данных; - преобразование данных.
7	Технологии, связанные с большими данными Рассматриваемые вопросы: - контроль данных; - анализ данных; - представление результатов.
8	Технологии, связанные с большими данными Рассматриваемые вопросы: - распределённые файловые системы: понятие и особенности; - виды распределённых файловых систем и их характеристики.
9	Технологии, связанные с большими данными Рассматриваемые вопросы:

№ п/п	Тематика лекционных занятий / краткое содержание
	- файловая система Hadoop; - принципы работы файловой система Hadoop.
10	Технологии, связанные с большими данными Рассматриваемые вопросы: - реляционные хранилища данных; - как реляционные хранилища данных используются в контексте больших данных.
11	Технологии, связанные с большими данными Рассматриваемые вопросы: - NoSQL- хранилища данных; - виды NoSQL; - HBase; - MongoDB.
12	Прикладные инструменты для работы с данными Рассматриваемые вопросы: - что такое лямбда-архитектура и какие преимущества она предоставляет; - как реализовать лямбда-архитектуру в системе обработки данных.
13	Прикладные инструменты для работы с данными Рассматриваемые вопросы: - что такое каппа-архитектура и как она отличается от лямбда-архитектуры; - как реализовать каппа-архитектуру в системе обработки данных.
14	Прикладные инструменты для работы с данными Рассматриваемые вопросы: - инструментарий Map/Reduce и что это такое; - как Map/Reduce используется в системах обработки данных.
15	Прикладные инструменты для работы с данными Рассматриваемые вопросы: - Apache Spark; - как Apache Spark используется в системах обработки данных.
16	Прикладные инструменты для работы с данными Рассматриваемые вопросы: - машинное обучение; - как машинное обучение используется в системах обработки данных.

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Большие данные На лабораторном занятии студент изучит: - понятие «большие данные»; - методы формирования «больших данных».
2	Большие данные На лабораторном занятии студент изучит - примеры «больших данных»; - возможности получения «больших данных» на примерах.
3	Большие данные На лабораторном занятии студент

№ п/п	Тематика практических занятий/краткое содержание
	- изучит принципы работы с данными; - источники получения данных.
4	Большие данные. На лабораторном занятии студент изучит - получение первичных статистик набора данных; - методы получения первичных статистик данных.
5	Большие данные На лабораторном занятии студент изучит - возможности MS Excel по работе с данными; - работу с закладкой Данные в MS Excel.
6	Анализ В результате лабораторного занятия осваиваются - элементы кластерного анализа; - основные методы кластеризации.
7	Анализ В результате лабораторного занятия осваиваются - SQL запросы; - работа с SQL запросы.
8	Проведение анализа В результате лабораторного занятия осваиваются - NoSQL запросы; - работа с NoSQL запросы.
9	Проведение анализа В результате лабораторного занятия осваиваются - методы визуализации; - современные инструменты визуализации данных.
10	Проведение анализа данных В результате лабораторного занятия выполняются примеры и инструменты визуализации с применением: - MS Excel; - MS Bi; - иных инструментов по выбору обучающегося.
11	Инструменты для работы с данными В результате лабораторного занятия студент изучает: - инструментарий HDMR; - функции HDMR.
12	Инструменты для работы с данными В результате лабораторного занятия изучаются примеры применения операторов Map на: - языке Python; - языке R.
13	Инструменты для работы с данными В результате лабораторного занятия изучаются примеры применения операторов Reduce на: - языке Python; - языке R.
14	Инструменты для работы с данными В результате лабораторного занятия изучается инструментарий HDMR. Примеры применения операторов Map и Reduce на языках: - Python; - R.

№ п/п	Тематика практических занятий/краткое содержание
15	Прикладные инструменты для работы с данными В результате лабораторного занятия студент изучает: - техники, применяемые при работе с большими данными; - инструмент для работы с большими данными.
16	Машинное обучение и «большие данные» В результате лабораторного занятия студент изучает: - существующие алгоритмы машинного обучения; - связь машинного обучения и «больших данных».

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Подготовка к практическим занятиям
2	Самостоятельное изучение темы «База данных MongoDB» [1, с.49-95]
3	Подготовка к промежуточной аттестации.
4	Подготовка к текущему контролю.

5. Перечень изданий, которые рекомендуется использовать при освоении дисциплины (модуля).

№ п/п	Библиографическое описание	Место доступа
1	Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2022. — 121 с. — ISBN 978-5-534-09837-2.	— Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/492609 (дата обращения: 21.04.2025).
2	Линейное программирование : учебное пособие для среднего профессионального образования / И. А. Палий. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2024. — 175 с. — (Профессиональное образование). — ISBN 978-5-534-18270-5.	— Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/534655 (дата обращения: 21.04.2025).

6. Перечень современных профессиональных баз данных и информационных справочных систем, которые могут использоваться при освоении дисциплины (модуля).

Официальный сайт РУТ (МИИТ) (<https://www.miit.ru/>).

Научно-техническая библиотека РУТ (МИИТ) (<http://library.miit.ru>).

Образовательная платформа «Юрайт» (<https://urait.ru/>).
Электронно-библиотечная система издательства «Лань» (<http://e.lanbook.com/>).
Федеральная служба государственной статистики: <https://rosstat.gov.ru/>
КонсультантПлюс: <http://www.consultant.ru/>
Гарант: <http://www.garant.ru/>

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, необходимого для освоения дисциплины (модуля).

1. Офисный пакет приложений Microsoft Office;
2. PyCharm
3. PostgreSQL

8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю).

Для проведения лекционных занятий необходима аудитория с мультимедиа аппаратурой. Для проведения практических занятий требуется аудитория, оснащенная мультимедиа аппаратурой и ПК с необходимым программным обеспечением и подключением к сети интернет.

9. Форма промежуточной аттестации:

Экзамен в 7 семестре.

10. Оценочные материалы.

Оценочные материалы, применяемые при проведении промежуточной аттестации, разрабатываются в соответствии с локальным нормативным актом РУТ (МИИТ).

Авторы:

доцент, доцент, к.н. кафедры
«Математическое моделирование
сложных систем» Института
железнодорожного транспорта

А.С. Милевский

Согласовано:

Заведующий кафедрой ИСЦЭ
Председатель учебно-методической
комиссии

Л.А. Каргина

М.В. Ишханян