

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



УТВЕРЖДАЮ
Декан РТФ

УТВЕРЖДАЮ /А.Н. Дедов/
(Ф.И.О. декана (директора института))

10.03.2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

С.1.2.30 Обработка больших данных

(код и наименование дисциплины по учебному плану)

Направление подготовки
(специальность)

11.05.01 Радиоэлектронные системы и комплексы

Квалификация выпускника

Инженер

(бакалавр/магистр/специалист)

Специализация

Радиолокационные системы и комплексы

Курс 4
Семестр 7

Распределение учебного времени

Трудоемкость по учебному плану	108 / 3	часов/зачетных единиц
Лекции	16	часов
Лабораторные работы	-	часов
Практические занятия	32	часов
Иная контактная работа	-	часов
Всего контактной работы (без учета экз.)	48	часов
Контактная работа по экзамену	-	часов
Курсовой проект (работа)	-	семестр
Самостоятельная работа обучающихся (без учета экз.)	60	часов
Самостоятельная работа по подготовке к экзамену	-	часов
Экзамен	-	семестр
Зачет	7	семестр
БРК, ДЗ	-	семестр

(год)

Программа составлена в соответствии с требованиями ФГОС ВО направления подготовки (специальности) 11.05.01 Радиоэлектронные системы и комплексы

Программу составили:

доцент	РТиМБС	СОГЛАСОВАНО	К.О. Иванов
(должность)	(кафедра)		(И.О. Фамилия)

РАССМОТРЕНА и ОДОБРЕНА на заседании кафедры, за которой закреплена дисциплина
Кафедра прикладной математики и информационных технологий

(наименование кафедры)	
_____	_____
(дата)	протокол №

Заведующий кафедрой	И.В. Журавлева
_____	_____
(подпись)	(И.О. Фамилия)

Рабочая программа СОГЛАСОВАНА с факультетом (институтом), выпускающей(ими) кафедрой(ами).

СООТВЕТСТВУЕТ действующей ОП.

Заведующий кафедрой	СОГЛАСОВАНО	Р.Г. Хафизов
_____	_____	_____
		(И.О. Фамилия)

Председатель методической комиссии факультета (института), в который входит выпускающая кафедра

СОГЛАСОВАНО	А.Н. Дедов
_____	_____
	(И.О. Фамилия)

Эксперт(ы): Власов Никита Михайлович, заместитель главного конструктора АО
Марийский машиностроительный завод - заместитель начальника НТЦ «Коралл»

Рабочая программа проверена и зарегистрирована в УМЦ 12.03.2025 г.

Специалист учебно-методического центра СОГЛАСОВАНО /Т.А. Смирнова/

Раздел 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины является достижение планируемых результатов обучения, соответствующих установленным в ОПОП индикаторам достижения компетенций:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения
1. ПК-6 Способен решать задачи оптимизации существующих и новых технических решений в условиях априорной неопределенности с применением пакетов прикладных программ	ПК-6.1 Знать методы оптимизации существующих и новых технических решений в условиях априорной неопределенности	знания: Знание математических методов, статистики, теории вероятностей для анализа данных и разработки моделей в условиях неопределенности. умения: навыки:
	ПК-6.2 Уметь применять современный математический аппарат для решения задачи оптимизации	знания: умения: Умение применять численные методы для решения математических задач, таких как методы оптимизации, численное интегрирование и дифференцирование. навыки:
	ПК-6.3 Владеть методами оптимизации проектируемых радиоэлектронных систем и комплексо	знания: умения: навыки: Навыки работы с программами для моделирования и симуляции радиоэлектронных систем, такими как SPICE, MATLAB, Simulink и другими.

Раздел 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к элективным дисциплинам (модулям) ОПОП.

Дисциплина является элективной

Для продолжения формирования заявленных компетенций необходимы знания предшествующих дисциплин: Статистическая радиотехника (ПК-6)

Изучаемая дисциплина является основой для продолжения формирования указанных компетенций в следующих дисциплинах: Моделирование радиотехнических систем в LabView (ПК-6); практиках: Преддипломная практика (ПК-6); государственной итоговой аттестации в форме: Выполнение и защита выпускной квалификационной работы (ПК-6)

Раздел 3. ОПИСАНИЕ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

Для формирования заявленных компетенций используются методологические технологии, реализующие деятельностный, личностно-ориентированный, практико-ориентированный подходы.

Основными стратегическими технологиями являются: дискуссионные, лекционные занятия, практические занятия, процедуры самообучения

На достижение конкретных целей обучения направлены применяемые тактические технологии: задания, классическая лекция, лекция с элементами мозгового штурма

Раздел 4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

7 семестр

Виды и темы занятий	Количество часов	Формируемые компетенции
Технологии анализа данных	108	ПК-6
Лекция. Большие данные (Big Data): современные подходы к обработке и хранению. Проблема множественного сравнения данных.	2	
Лекция. Процесс анализа. Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представления наборов данных.	2	
Лекция. Технологии KDD и Data Mining. Подготовка данных к анализу. Методика извлечения знаний. Data Mining. Мультидисциплинарный характер Data Mining. Причины распространения KDD и Data Mining. Актуальность технологий Data Mining как средств обработки больших объемов информации.	2	
Лекция. Программное обеспечение в области анализа данных. Аналитические платформы: классификация и особенности применения. Языки визуального моделирования.	2	
Лекция. Ассоциативные правила. Аффинитивный анализ, предметный набор. Поддержка и достоверность ассоциативного правила. Значимость ассоциативных правил, лифт и левередж. Поиск ассоциативных правил. Частые предметные наборы и их обнаружение. Алгоритм генерации ассоциативных правил. Иерархические ассоциативные правила. Методы поиска иерархических ассоциативных правил.	4	
Лекция. Определение кластеризации. Постановка задачи кластеризации. Цели кластеризации в Data Mining. Примеры кластеризации в различных областях. Виды метрик. Шаги алгоритма. Меры расстояний. Пример работы алгоритма k-means. Проблемы алгоритмов кластеризации.	2	
Лекция. Применение классификации и регрессии. Обзор методов классификации и регрессии. Статистические методы. Методы, основанные на обучении, разнообразие подходов.	2	
Практическое занятие. Начало работы. Понятие сценария и узла обработки. Консолидация данных. Трансформация данных. Визуализация данных.	4	
Практическое занятие. Ассоциативные правила. Поиск ассоциативных правил.	4	
Практическое занятие. Кластеризация. Алгоритм кластеризации k-means.	6	
Практическое занятие. Прогнозирование с помощью линейной регрессии.	6	
Практическое занятие. Классификация с помощью деревьев решений.	6	
Практическое занятие. Классификация с помощью нейросети.	6	

Задания для самостоятельной работы, в том числе выполнение Проработка необходимых вопросов для подготовки к лекциям и практическим занятиям. Самостоятельное изучение таких разделов, как определение дерева решений; причины популярности и условия применимости; структура дерева решений; выбор атрибута разбиения в узле; алгоритм ID3; критерий выбора атрибута разбиения ID3.	60	
Иная контактная работа:	0	

Раздел 5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Изучение дисциплины рекомендуется начать с ознакомления с рабочей программой, ее структурой и содержанием разделов. Учебный материал структурирован, изучение дисциплины осуществляется в тематической последовательности. **Занятия лекционного типа** дают систематизированные знания по дисциплине, концентрируют внимание на наиболее сложных и важных вопросах. Во время лекционных занятий рекомендуется вести конспектирование учебного материала; обращать внимание на формулировки и категории, раскрывающие суть проблемы, явления или процесса; зафиксировать выводы и практические рекомендации. Подготовка к **занятиям семинарского типа** включает ознакомление с планом практического занятия; работу с конспектом лекций, выполнение домашнего задания, работу с учебной и учебно-методической литературой, научными изданиями и электронными образовательными ресурсами, рекомендованными рабочей программой дисциплины.

Содержание **самостоятельной работы** определяется рабочей программой дисциплины, оценочными и методическими материалами, заданиями и указаниями преподавателя. Самостоятельная работа может осуществляться в аудиторной и внеаудиторной формах. Эффективным средством осуществления самостоятельной работы является электронная информационно-образовательная среда университета, которая обеспечивает доступ к образовательной программе, рабочей программе дисциплины, к электронным библиотечным системам, профессиональным базам данных и информационным справочным системам.

Периодичность проведения, формы текущего контроля успеваемости, система оценивания

Раздел 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. Учебно-методическое обеспечение

№№ п/п	Список используемой литературы	Количество экземпляров печатных изданий, имеющих в библиотеке, или электронный адрес издания (ресурса) в сети Интернет
УЧЕБНЫЕ, УЧЕБНО-МЕТОДИЧЕСКИЕ И НАУЧНЫЕ ИЗДАНИЯ		
1.	Бояркина, Лариса Александровна. Проектирование баз данных [Текст] : лаб. практикум / Л. А. Бояркина. Йошкар-Ола: МарГТУ, 2005. - 63 с. Экземпляры: всего 150.	150
2.	Иванов, Олег Евгеньевич. Реляционные базы данных [Текст] : учеб. пособие / О. Е. Иванов, С. Э. Нам. Йошкар-Ола: МарГТУ, 2007. - 227 с. Экземпляры: всего 46.	46 / https://portal.volgatech.net/books/ivanov-nam-reljac-BD.pdf

3.	Галочкин, Владимир Иванович. Базы данных [Текст] : учеб. пособие / В. И. Галочкин. Йошкар-Ола: МарГТУ, 2009. - 199 с. ISBN 978-5-8158-0688-7. Экземпляры: всего 90.	90
4.	Кудинов, Ю. И. Практикум по основам современной информатики [Электронный ресурс] / Кудинов Ю. И., Пащенко Ф. Ф., Келина А. Ю. Санкт-Петербург: Лань, 2022. - 352 с. ISBN 978-5-8114-1152-8.	https://e.lanbook.com/book/210749
5.	Орлова, И. В. Информатика. Практические задания [Электронный ресурс] : учебное пособие / Орлова И. В. Санкт-Петербург: Лань, 2024. - 140 с. ISBN 978-5-8114-3608-8.	https://e.lanbook.com/book/358664
6.	Нестеров, Сергей Александрович. Базы данных [Текст : Электронный ресурс] : учебник и практикум для вузов / С. А. Нестеров. Москва: Юрайт, 2023. - 230 с ISBN 978-5-534-00874-6.	https://urait.ru/bcode/511650
ЭЛЕКТРОННЫЕ ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ		
1.	Научная электронная библиотека eLIBRARY.RU	http://elibrary.ru
2.	Научная электронная библиотека «Киберленинка»	http://cyberleninka.ru
ПРОФЕССИОНАЛЬНЫЕ БАЗЫ ДАННЫХ И ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ		
1.	Справочно-правовая система Консультант+	http://www.consultant.ru
2.	Информационно-правовой портал Гарант	http://www.garant.ru
3.	Профессиональные справочные системы Техэксперт	http://www.cntd.ru

6.2. Материально-техническая база и программное обеспечение

№№ п/п	Аудитории для проведения учебных занятий, самостоятельной работы и проведения государственной итоговой аттестации	Перечень основного оборудования	Программное обеспечение
1.	531 (III)	ПК B112,2 420W/Intel Celeron Dual-Core E3300/кл,мышь,фильт,мон. VA1931 (5)	Microsoft Windows Enterprise, Справочная правовая система "Консультант Плюс", Microsoft Office Standard, Агент Dr.Web, Комплект ГАРАНТ-Мастер, Microsoft Access, Microsoft Visio Professional, Microsoft Project Professional, Microsoft Visual Studio Enterprise, Комплект ПО для решения основных пользовательских задач

Раздел 7. ФОРМЫ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ/ ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Критерии оценивания индикаторов достижения компетенций направлены на:

- усвоение теоретического материала (объем знаний, глубина усвоения), предусмотренного рабочей программой;
 - умение излагать материал (четкость, грамотность изложения материала, точность и полнота воспроизведения учебного материала);
 - умение применять теоретические знания при решении практических заданий.
- Шкала оценивания представлена ниже.

Уровень сформированности элементов компетенции	Критерии оценивания	Шкала оценивания
Пороговый уровень	Обучающийся имеет знания основного материала, проявляет умение логично его излагать, но может допускать неточности в изложении материала, недостаточно правильные формулировки, испытывает затруднения в выполнении практических заданий	Зачтено

7.1. Текущий контроль успеваемости

Текущий контроль успеваемости обеспечивает оценивание хода освоения дисциплины (модуля) и производится с применением технологии рейтингового контроля в соответствии с технологической картой дисциплины. Порядок составления технологической карты и алгоритм проведения процедуры оценивания видов деятельности обучающихся, направленных на освоение знаний, умений, навыков и/или опыта деятельности, по накопительной системе в баллах устанавливается положением о системе РИТМ в ФГБОУ ВО «ПГТУ»

7.2. Промежуточная аттестация обучающихся

Промежуточная аттестация обучающихся направлена на оценивание результатов обучения по дисциплине (модулю) и проводится с использованием фондов оценочных средств.

Примеры типовых контрольных заданий из базы фонда оценочных средств по образовательной программе.

1. Что такое SQL?

- Реляционная база данных
- Язык неструктурированных запросов
- Язык структурированных запросов
- Средство для создания пайплайнов

2. Одна из главных целей Big Data – это...

- привлечение все больше пользователей
- рост числа обработок данных
- снижение издержек операций
- таргетирование пользователей

3. Закончите следующее предложение: С точки зрения машины, информация становится структурированной, если...

- Машина проинструктирована, каким образом её обрабатывать
- Информация разделена на части и озаглавлена
- Информация имеет логическую взаимосвязь внутри себя
- Машина знает из каких частей состоит информация

4. Отметьте причину создания NoSQL баз данных...

- Высокая стоимость горизонтальной масштабируемости RDBMS при сохранении требования b)

высокой доступности

- с) недостаточная гибкость языка запросов SQL
- d) невозможность хранить большие объёмы
- e) данных дороговизна лицензий RDBMS

5. Искусственные нейронные сети (ИНС) — модели машинного обучения,использующие комбинации распределенных простых операций, зависящих отобучаемых параметров, для обработки входных данных. Какого вида ИНС не существует?

- a) Наивные
- b) Рекуррентные
- с) Импульсные
- d) Противоборствующие

6. Что не является целью процесса Business Intelligence?

- a) Интерпретация большого количества данных
- b) Моделирование исходов различных вариантов действий
- с) Модификация существующего программного обеспечения
- d) Отслеживание результатов решений

7. Что из перечисленного помогает следить за эволюцией документа, надсозданием которого работает одновременно большое количество авторов?

- a) Пространственный поток
- b) Исторический поток
- с) Визуальный поток
- d) Интерактивный поток

8. Выберите неверное высказывание...

- a) большие данные – это данные объёма свыше 1 Тб
- b) проблема больших данных – это проблема, когда при существующихтехнологиях хранения и обработки сущностная обработка данных затрудненаили невозможна
- с) большие данные – это тренд в области ИТ, подогреваемый маркетинговыми кампаниями крупных вендоров
- d) большие данные как правило не структурированы

Перечень вопросов для проведения промежуточной аттестации

1. Какова цель обработки больших данных в вашем проекте или организации?
2. Какие источники данных используются для сбора больших объемов информации?
3. Как вы определяете критерии отбора данных для анализа?
4. Какие методы сбора и хранения больших данных вы используете?
5. Какие технологии обработки данных (например, Apache Hadoop, Apache Spark) используются в вашем проекте?
6. Как происходит очистка и предварительная обработка данных перед анализом?
7. Каковы методы обеспечения безопасности данных в процессе их обработки?
8. Каким образом обеспечивается масштабируемость вашей системы обработки данных?
9. Какие алгоритмы и модели машинного обучения применяются при обработке больших данных?
10. Как вы решаете проблемы с производительностью при обработке больших объемов данных?
11. Как организован процесс мониторинга и управления качеством данных?
12. Как осуществляется визуализация и интерпретация результатов обработки больших данных?
13. Каковы методы резервного копирования и восстановления данных?
14. Как вы управляете и храните метаданные в процессе обработки данных?
15. Какие вызовы и проблемы с конфиденциальностью данных вы решаете в процессе обработки больших данных?
16. Как происходит управление версиями данных в вашей системе?

17. Какова ваша стратегия в отношении обучения моделей на обновленных данных?
18. Какие инструменты используются для управления рабочим процессом обработки больших данных?
19. Как вы определяете и измеряете эффективность процесса обработки данных?
20. Какова роль и влияние больших данных на стратегическое принятие решений в компании?