

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



УТВЕРЖДАЮ
Декан ЭФ

УТВЕРЖДАЮ /Н.М. Стрельникова/
(Ф.И.О. декана (директора института))

19.03.2026 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б.1.2.13 Обработка больших данных

(код и наименование дисциплины по учебному плану)

Направление подготовки
(специальность)

09.03.03 Прикладная информатика

Квалификация выпускника

Бакалавр

(бакалавр/магистр/специалист)

Направленность

Искусственный интеллект в бизнес-аналитике

Курс 3
Семестр 6

Распределение учебного времени

Трудоемкость по учебному плану	180 / 5	часов/зачетных единиц
Лекции	16	часов
Лабораторные работы	32	часов
Практические занятия	-	часов
Иная контактная работа	-	часов
Всего контактной работы (без учета экз.)	48	часов
Контактная работа по экзамену	6	часов
Курсовой проект (работа)	6	семестр
Самостоятельная работа обучающихся (без учета экз.)	96	часов
Самостоятельная работа по подготовке к экзамену	30	часов
Экзамен	6	семестр
Зачет	-	семестр
БРК, ДЗ	-	семестр

(год)

Программа составлена в соответствии с требованиями ФГОС ВО направления подготовки (специальности) 09.03.03 Прикладная информатика

Программу составили:

доцент	ИСЭ	СОГЛАСОВАНО	О.В. Порядина
(должность)	(кафедра)		(И.О. Фамилия)

РАССМОТРЕНА и ОДОБРЕНА на заседании кафедры, за которой закреплена дисциплина
Кафедра информационных систем в экономике

(наименование кафедры)		
17.02.2026	протокол №	10
(дата)		

Заведующий кафедрой	СОГЛАСОВАНО	Т.А. Уразаева
		(И.О. Фамилия)

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

СООТВЕТСТВУЕТ действующей ОП.

Заведующий кафедрой	СОГЛАСОВАНО	Т.А. Уразаева
		(И.О. Фамилия)

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

СОГЛАСОВАНО	О.Е. Иванов
	(И.О. Фамилия)

Эксперт(ы): Галочкин Александр Владимирович, генеральный директор ООО «Трэвэл Лайн»

Рабочая программа проверена и зарегистрирована в УМЦ 19.02.2026 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. BD-1 Способен осуществлять поиск, сбор, очистку и предварительный анализ данных	BD-1.2. Применяет методы анализа данных для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ	знания: Знает методы заполнения пропусков в данных и удаления выбросов в табличных данных. умения: Умеет применять методы заполнения пропусков в данных и удаления выбросов в табличных данных. навыки: Владеет методикой использования методов анализа данных для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ.
	BD-1.3. Применяет методы понижения размерности для первичной интерпретации и визуализации многомерных данных	знания: Знает библиотеки, реализующие методы понижения размерности, для первичной интерпретации и визуализации многомерных данных. умения: Умеет применять библиотеки, реализующие методы понижения размерности, для первичной интерпретации и визуализации многомерных данных. навыки: Владеет методикой использования методов понижения размерности для первичной интерпретации и визуализации многомерных данных.
2. BD-2 Способен определять требования к наборам данных для решения задач машинного обучения, проводить разметку и анализ наборов данных, оценивать качество данных, обеспечивать непрерывную интеграцию данных	BD-2.1. Определяет требования к наборам и качеству данных для решения задач машинного обучения	знания: Знает методы решения задач разметки и оценки качества разведочного анализа данных. умения: Умеет ставить задачу разметки и оценки качества данных для решения задач машинного обучения. навыки: Владеет методикой разработки требований к наборам и качеству данных для решения задач машинного обучения.
	BD-2.2. Работает с данными, в том числе собирает данные из разрозненных источников, проверяет данные на корректность	знания: Знает инструменты разметки данных под условия задачи, методы краудсорсинга разметки. умения: Умеет подбирать метод и инструментарий разметки данных под условия задачи и организовывать краудсорсинг разметки данных. навыки: Владеет методиками разработки требований для инструментария разметки и проведения оценки качества данных на корректность.

3. BD-4 Способен применять различные модели и (или) технологии хранения данных	BD-4.1. Осуществляет выбор технологий обработки больших данных, приемлемых для создания прикладной системы ИИ с заданными требованиями	знания: Знает принципы модели MapReduce для параллельной обработки больших данных. умения: Умеет организовывать распределенное хранилище и параллельную обработку на базе современных технологий (Hadoop, Spark) больших данных. навыки: Владеет методикой параллельной обработки больших данных с использованием модели MapReduce, а также способен осуществлять выбор и применять другие технологии обработки больших данных для решения поставленной задачи.
	BD-4.2. Разрабатывает и отлаживает прикладные решения с элементами ИИ с применением различных технологий хранения данных	знания: Знать: принципы разработки прикладных решений с элементами ИИ с применением различных технологий данных. умения: навыки:

Раздел 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к части, формируемой участниками образовательных отношений ОПОП.

Дисциплина является обязательной

Для продолжения формирования заявленных компетенций необходимы знания предшествующих дисциплин: Анализ данных (BD-1)

Изучаемая дисциплина является основой для продолжения формирования указанных компетенций в следующих практиках: Преддипломная практика (BD-1), Преддипломная практика (BD-2), Преддипломная практика (BD-4); государственной итоговой аттестации в форме: Выполнение и защита выпускной квалификационной работы (BD-1), Выполнение и защита выпускной квалификационной работы (BD-2), Выполнение и защита выпускной квалификационной работы (BD-4)

Раздел 3. ОПИСАНИЕ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

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

Основными стратегическими технологиями являются: имитационное моделирование, исследовательские, лекционные занятия, практические и лабораторные занятия, процедуры самообучения, лекция с обсуждением новых научных достижений в сфере ии по результатам конференции и публикаций, открытые образовательные ресурсы

На достижение конкретных целей обучения направлены применяемые тактические технологии: задания, информационные, классическая лекция, лекция с элементами мозгового штурма, мини-проекты, case-study, проблемная лекция, семинары с ит-практиками, конкурсы, хакатоны, кейс-чемпионаты, мастер-классы, командные проектные задания

Раздел 4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Виды и темы занятий	Количество часов	Формируемые компетенции
1. Введение в экосистему Big Data.	33	BD-1, BD-2, BD-4
Лекция. Лекция 1. Предмет курса. Понятие Big Data: 3V (Volume, Velocity, Variety) и расширенные V (Veracity, Value). Жизненный цикл данных (Data Lifecycle).	2	
Лекция. Лекция 2. Обзор типичных задач обработки больших данных: очистка данных, разведочный анализ, визуализация, машинное обучение. Знакомство с библиотеками Python для работы с большими данными: NumPy, Pandas, Matplotlib и надстройка Seaborn, Plotly.	2	
Лабораторная работа. Лабораторная работа 1. Изучение кейсов (примеров) применения технологий больших данных в организациях и органах государственной власти РФ и их применения в бизнес-аналитике.	2	
Лабораторная работа. Лабораторная работа 2. Изучение примеров визуализации больших данных для дальнейшего их применения в бизнес-аналитике.	2	
Лабораторная работа. Лабораторная работа 3. Применение библиотек NumPy, Pandas для обработки больших данных.	2	
Лабораторная работа. Лабораторная работа 4. Применение библиотек Plotly, Matplotlib и надстройки Seaborn для обработки больших данных.	4	
Лабораторная работа. Лабораторная работа 5. Методы и инструменты визуализации данных (Tableau, Power BI).	4	

Задания для самостоятельной работы, в том числе выполнение курсового проекта/работы Задания для самостоятельной работы, в том числе выполнение курсового проекта/работы ГОСТ Р 71476-2024 ИИ. Концепции и терминология искусственного интеллекта Учебное пособие по Python: https://docs.python.org/3/tutorial/index.html Разведочный анализ (EDA): https://habr.com/ru/companies/otus/articles/752434/ Юре, Л. Анализ больших наборов данных [Электронный ресурс] / Л. Юре, Р. Ананд, Д.У. Джеффри. — Электрон. дан. — Москва : ДМК Пресс, 2016. — 498 с. — Режим доступа: https://e.lanbook.com/book/93571 Что нам стоит диаграмму в Python построить: 5 вариантов привлекающей внимание визуализации данных и кое-что ещё https://habr.com/ru/companies/ru_mts/articles/738208/ EDA в Python для новичков: что почитать и где попрактиковаться https://habr.com/ru/companies/yandex_praktikum/articles/952084/ Изучение литературы по темам раздела. Выполнение домашних заданий. Подготовка к лабораторным работам. Выполнение курсовой работы по индивидуальному заданию. Изучение исследований и разработок по темам лекционных и лабораторных занятий, представленных в научных публикациях на конференциях уровня А*/А, статьях, опубликованных в журналах 1, 2, 3 уровня Белого списка, научных журналов Минобрнауки России. выполнение курсового проекта/работы	15 15	
2. Технологии анализа больших данных и их подготовки для задач машинного обучения.	38	BD-1, BD-2, BD-4
Лекция. Лекция 3. Структурированные и неструктурированные данные. Алгоритмы сбора и предварительного анализа данных: парсинг веб-страниц, работа с API, сбор логов.	2	
Лекция. Лекция 4. Алгоритмы предобработки больших данных: процессы очистки данных, работа с пропусками, дубликатами, выбросами. Разметка данных. Инструменты разметки данных.	2	
Лекция. Лекция 5. Разметка и анализ качества данных, обеспечение непрерывной интеграции данных.	4	
Лабораторная работа. Лабораторная работа 6. Изучение примеров структурированных и неструктурированных больших данных.	2	
Лабораторная работа. Лабораторная работа 7. Работа с инструментами разметки данных, изучение принципов работы: CVAT, LabelMe и др. Решение задач.	8	

Задания для самостоятельной работы, в том числе выполнение курсового проекта/работы Задания для самостоятельной работы, в том числе выполнение курсового проекта/работы Муромцев Д.И., Шилин И.А., Исаев И.В., Структурирование, разметка и обогащение данных– СПб: Университет ИТМО, 2024. – 72 с. https://books.ifmo.ru/file/pdf/3329.pdf 5 лучших инструментов для разметки данных в 2024 году. https://habr.com/ru/companies/data_light/articles/850770/ Первый этап внедрения ML: как провести разметку данных. https://embedika.ru/blog/razmetku-dannyh Изучение литературы по темам раздела. Выполнение домашних заданий. Подготовка к лабораторным работам. Выполнение курсовой работы по индивидуальному заданию. Изучение исследований и разработок по темам лекционных и лабораторных занятий, представленных в научных публикациях на конференциях уровня А*/А, статьях, опубликованных в журналах 1, 2, 3 уровня Белого списка, научных журналов Минобрнауки России. выполнение курсового проекта/работы	20 16	
3. Параллельные вычисления. Параллельная обработка данных.	32	BD-1, BD-2, BD-4
Лекция. Лекция 6. Параллельная обработка больших данных. Технология MapReduce для работы с BigData.	2	
Лекция. Лекция 7. Знакомство с методами: метод Монте-Карло, метод Монте-Карло на основе Марковской цепи, методы блужданий.	2	
Лабораторная работа. Лабораторная работа 8. Решение задач по технологии MapReduce.	4	
Лабораторная работа. Лабораторная работа 9. Решение задач на Python методом Монте-Карло.	2	
Лабораторная работа. Лабораторная работа 10. Знакомство с системами параллельной обработки больших данных Hadoop и Spark.	2	

Задания для самостоятельной работы, в том числе выполнение курсового проекта/работы	
Задания для самостоятельной работы, в том числе выполнение курсового проекта/работы	
Метод Монте-Карло — один из самых полезных алгоритмов в ИТ. https://thecode.media/monte-carlo/	
Шаг за шагом: Исследование модели случайного блуждания https://www.techsyncer.com/ru/random-walk.html	
Основы работы со Spark DataFrame. https://habr.com/ru/articles/663690/	
https://spark.apache.org/	
Изучение литературы по темам раздела.	
Выполнение домашних заданий.	
Подготовка к лабораторным работам.	
Выполнение курсовой работы по индивидуальному заданию.	
Изучение исследований и разработок по темам лекционных и лабораторных занятий, представленных в научных публикациях на конференциях уровня А*/А, статьях, опубликованных в журналах 1, 2, 3 уровня Белого списка, научных журналов Минобрнауки России.	20
выполнение курсового проекта/работы	10
Иная контактная работа:	0
Подготовка к экзамену	30
Проведение экзамена	6

Раздел 5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Изучение дисциплины "Обработка больших данных" 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 "Обработка больших данных", 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. В рамках лекционных занятий запланировано ознакомление студентов и дальнейшее обсуждение новых научных достижений в сфере обработки больших данных, опубликованных в ведущих журналах по данной тематике.

Подготовка к занятиям **семинарского типа** включает ознакомление с планом лабораторного 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 "Обработка больших данных". Для изучения практического материала планируется приглашение ИТ-практиков и проведение мастер-классов, изучение практических кейсов обработки больших данных ведущими компаниями региона, демонстрация решения реальных задач обработки данных в бизнес-аналитике.

Содержание **самостоятельной работы** определяется рабочей программой дисциплины "Обработка больших данных", 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 "Обработка больших данных", к электронным библиотечным системам, профессиональным базам данных и информационным справочным системам.

В качестве дополнения к материалам студентам будет рекомендовано использование открытых образовательных ресурсов открытых образовательных платформ по тематике дисциплины.

В ходе обучения по дисциплине "Обработка больших данных" студенты под руководством ведущего преподавателя изучают исследования и разработки, представленные в научных публикациях на тематических конференциях уровня А*/А, статьях (например, конференции IEEE International Conference on Data Mining; 16. ACM Special Interest Group on Management of Data Conference; 17. International Conference on Data Engineering; International Conference on Very Large Databases; International Conference on Artificial Intelligence and Statistics; IEEE International Parallel and Distributed Processing Symposium (was IPPS and SPDP); SIAM International Conference on Data Mining), опубликованных в журналах 1, 2 и 3 уровней Белого списка научных журналов Минобрнауки России, в таких журналах как "Бизнес-информатика", "Информатика и автоматизация", "Информатика и ее применение", "Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления", "Известия Института математики и информатики Удмуртского государственного университета", "Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика". Изучение дисциплины "Обработка больших данных" включает выполнение лабораторных работ и домашних заданий. Периодичность проведения, формы текущего контроля успеваемости, система оценивания хода освоения дисциплин представлены в рабочей программе. В ходе обучения по дисциплине "Обработка больших данных" студенты выполняют курсовую работу по индивидуальному заданию, связанному с реальной практической задачей и применением современных методов и инструментов обработки больших данных.

Для формирования практических навыков студентам будет предложено участие в молодежных студенческих конференциях, конкурсах, хакатонах, кейс-чемпионатах.

Формой промежуточной аттестации по дисциплине "Обработка больших данных" является экзамен.

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.	Сузи, Р. А. Язык программирования Python	

	ресурс] / Сузи Р. А. 2-е изд. Москва: ИНТУИТ, 2016. - 350 с. ISBN 5-9556-0058-2.	https://e.lanbook.com/book/100546
2.	Макшанов, А. В. Большие данные. Big Data [Электронный ресурс] : учебник для вузов / Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н.; Журавлев А. Е., Тындыкарь Л. Н. 3-е изд., стер. Санкт-Петербург: Лань, 2023. - 188 с. ISBN 978-5-507-46866-9.	https://e.lanbook.com/book/322664
3.	Парамонов, И. Ю. Методы и модели исследования сложных систем и обработки больших данных [Электронный ресурс] : монография / И. Ю. Парамонов, В. А. Смагин, Н. Е. Косых, А. Д. Хомоненко. Санкт-Петербург: Лань, 2020. - 236 с. ISBN 978-5-8114-4006-1.	https://e.lanbook.com/book/126938
4.	Кревецкий, Александр Владимирович. Основы технологий искусственного интеллекта [Электронный ресурс] : учебное пособие / А. В. Кревецкий, Ю. А. Ипатов, Н. И. Роженцова ; под общей редакцией А. В. Кревецкого; Министерство науки и высшего образования Российской Федерации, ФГБОУ ВО "Поволжский государственный технологический университет". Йошкар-Ола: ПГТУ, 2023. - 272 с. ISBN 978-5-8158-2358-	https://portal.volgatech.net/books/Krevetskiy_Osnovy_tekhnologii_iskusstvennogo_intellekta_2023.pdf
5.	Цехановский, В. В. Архитектор информационных систем: как проектировать большие системы [Электронный ресурс] / Цехановский В. В. Москва, Алматы: Ай Пи Ар Медиа, EDP Hub (Идипи Хаб), 2024. - 284 с. ISBN 978-5-4497-2315-4.	https://www.iprbookshop.ru/133454.html

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.	Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации (Минцифры)	https://digital.gov.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.	250а (III)	Комплект учебной мебели (1)	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, STATISTICA 6.1 for Windows Russian, 1С:Документооборот 8 КОРП, ERwin Process Modeler r7.3, 1С:Предприятие 8. Комплект для обучения., Программный комплекс "Компьютерная деловая игра "БИЗНЕС-КУРС: Максимум. версия 1", Project Expert 7.55 Tutorial, Arena, AnyLogic 7, Embarcadero RAD Studio XE2 Professional 10 Named Users ESD, 1С:Документооборот 8 КОРП, 1С:Предприятие 8. Комплект для обучения., Project Expert
2.	2506 (III)	ПК ICL RAY S902.1,клавиат.,мышь,патч корд 3м,монитор ViewSonic 21,5" VA2248-LED (16), Комплект учебной мебели (1)	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, STATISTICA 6.1 for Windows Russian, 1С:Документооборот 8 КОРП, ERwin Process Modeler r7.3, 1С:Предприятие 8. Комплект для обучения.,

			Программный комплекс "Компьютерная деловая игра "БИЗНЕС-КУРС: Максимум. версия 1", Project Expert 7.55 Tutorial, Arena, AnyLogic 7, Embarcadero RAD Studio XE2 Professional 10 Named Users ESD, 1С:Документооборот 8 КОРП, 1С:Предприятие 8. Комплект для обучения., Project Expert
3.	335 (III)	Доска маркерная 120x240 см (1), Персональный компьютер Power RaY P550 (16), Экран настенный рулонный 200x200 см (1), Комплект учебной мебели (1)	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, STATISTICA 6.1 for Windows Russian, 1С:Документооборот 8 КОРП, ERwin Process Modeler r7.3, 1С:Предприятие 8. Комплект для обучения., Программный комплекс "Компьютерная деловая игра "БИЗНЕС-КУРС: Максимум. версия 1", Project Expert 7.55 Tutorial, Arena, AnyLogic 7, Embarcadero RAD Studio XE2 Professional 10 Named Users ESD, 1С:Документооборот 8 КОРП, 1С:Предприятие 8. Комплект для обучения., Project Expert

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.	удовлетворительно
Продвинутый уровень	Обучающийся твердо знает программный материал, излагает его грамотно и по существу, не допускает существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения	хорошо
Высокий уровень	Обучающийся глубоко и прочно усвоил программный материал, грамотно и логически стройно его излагает, дает исчерпывающие ответы на поставленные вопросы. В ответе тесно увязывается теория с практикой, при этом обучающийся не затрудняется с ответом при видоизменении задания, свободно справляется с задачами, вопросами и другими видами применения знаний, показывает знакомство с монографической литературой, периодическими изданиями, правильно обосновывает принятые решения, свободно владеет разносторонними навыками, приемами выполнения практических работ	отлично

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 из перечисленного не относится к основным характеристикам больших данных (3V)?

- a) Объем
- b) Скорость
- c) Вариативность
- d) Вероятность

Пояснение: Три основных характеристики — объем (Volume), скорость (Velocity) и разнообразие (Variety), а не "вариативность" или "вероятность".

2. Какой тип данных характеризуется отсутствием четкой структуры и формата?

- a) Структурированные
- b) Полуструктурированные
- c) Неструктурированные
- d) Реляционные

Пояснение: Неструктурированные данные не имеют фиксированного формата (например, видео, текст).

3. Какая из перечисленных систем является распределённым файловым хранилищем, часто используемым для больших данных?

- a) MySQL
- b) HDFS
- c) Oracle DB
- d) SQLite

Пояснение: HDFS — Hadoop Distributed File System, распределённое хранилище данных.

4. Что такое MapReduce?

- a) Модель потоковой обработки данных
- b) Алгоритм машинного обучения
- c) Парадигма пакетной обработки больших данных
- d) База данных

Пояснение: MapReduce — модель пакетной обработки больших данных.

5. Какая платформа предназначена для быстрой распределённой обработки данных в памяти?

- a) Apache Hadoop
- b) Apache Spark
- c) Apache Kafka
- d) MongoDB

Пояснение: Apache Spark работает с данными в памяти, обеспечивая высокую скорость обработки.

6. Какой из перечисленных инструментов используется для потоковой передачи сообщений?

- a) Apache Hive
- b) Apache Kafka
- c) Apache Pig
- d) Apache Flink

Пояснение: Kafka — распределённая система обмена сообщениями.

7. Какой тип базы данных наиболее подходит для хранения неструктурированных данных?

- a) Реляционная
- b) NoSQL
- c) Графовая

d) Табличная

Пояснение: NoSQL базы данных хорошо справляются с неструктурированными и полуструктурированными данными.

8. Какой из перечисленных вызовов не относится к проблемам больших данных?

- a) Масштабируемость
- b) Конфиденциальность
- c) Скорость печати документов
- d) Качество данных

Пояснение: Скорость печати документов не связана с обработкой больших данных.

9. Что из перечисленного является примером источника больших данных?

- a) Книги в библиотеке
- b) Логи серверов
- c) Печатные газеты
- d) Ручные записи

Пояснение: Логи серверов — типичный источник больших данных.

10. Какой из подходов лучше всего подходит для анализа исторических данных?

- a) Поточковая обработка
- b) Пакетная обработка
- c) Машинное обучение
- d) Облачные вычисления

Пояснение: Пакетная обработка эффективна для анализа больших объемов исторических данных.

11. Какой из следующих трендов связан с развитием технологий больших данных?

- a) Edge computing
- b) Текстовый редактор
- c) Классическая реляционная модель
- d) Локальное хранение данных

Пояснение: Edge computing — вычисления на периферии сети, важный тренд для больших данных.

Перечень вопросов для проведения промежуточной аттестации

Перечень вопросов для проведения промежуточной аттестации.

1. Понятие Big Data. Основные характеристики Big Data.

2. Жизненный цикл данных.

3. Примеры кейсов больших данных для решения задач бизнес-аналитики.

4. Обзор типичных задач обработки больших данных: очистка данных, разведочный анализ, визуализация, машинное обучение.

5. Библиотека Python для работы с большими данными: NumPy.

6. Библиотека Python для работы с большими данными: Pandas.

7. Библиотека Python для работы с большими данными: Matplotlib и надстройка Seaborn.

8. Библиотека Python для работы с большими данными: Plotly.

9. Структурированные и неструктурированные данные.

10. Алгоритмы сбора и предварительного анализа данных.
11. Алгоритмы предобработки больших данных: процессы очистки данных, работа с пропусками, дубликатами, выбросами.
12. Алгоритмы предобработки больших данных: Разметка данных.
13. Инструменты разметки данных.
14. Обеспечение непрерывной интеграции данных.
15. Параллельная обработка больших данных.
16. Технология MapReduce для работы с BigData.
17. Использование метода Монте-Карло для обработки больших данных.
18. Использование методов блужданий для обработки больших данных.