

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



УТВЕРЖДАЮ
Декан ФИиВТ

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

25.08.2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б.1.2.2 Анализ данных

(код и наименование дисциплины по учебному плану)

Направление подготовки
(специальность)

09.03.04 Программная инженерия

Квалификация выпускника

Бакалавр

(бакалавр/магистр/специалист)

Направленность

Разработка программных систем и систем искусственного
интеллекта

Курс 2
Семестр 3

Распределение учебного времени

Трудоемкость по учебному плану	108 / 3	часов/зачетных единиц
Лекции	18	часов
Лабораторные работы	-	часов
Практические занятия	36	часов
Иная контактная работа	-	часов
Всего контактной работы (без учета экз.)	54	часов
Контактная работа по экзамену	-	часов
Курсовой проект (работа)	-	семестр
Самостоятельная работа обучающихся (без учета экз.)	54	часов
Самостоятельная работа по подготовке к экзамену	-	часов
Экзамен	-	семестр
Зачет	3	семестр
БРК, ДЗ	-	семестр

(год)

Программа составлена в соответствии с требованиями ФГОС ВО направления подготовки (специальности) 09.03.04 Программная инженерия

Программу составили:

заведующий кафедрой	ИиСП	СОГЛАСОВАНО	А.В. Бородин
(должность)	(кафедра)		(И.О. Фамилия)

РАССМОТРЕНА и ОДОБРЕНА на заседании кафедры, за которой закреплена дисциплина
Кафедра информатики и системного программирования

(наименование кафедры)		
22.08.2025	протокол №	1
(дата)		

Заведующий кафедрой	СОГЛАСОВАНО	А.В. Бородин
		(И.О. Фамилия)

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

СООТВЕТСТВУЕТ действующей ОП.

Заведующий кафедрой	СОГЛАСОВАНО	А.В. Бородин
		(И.О. Фамилия)

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

СОГЛАСОВАНО	А.А. Кречетов
	(И.О. Фамилия)

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

Рабочая программа проверена и зарегистрирована в УМЦ 25.08.2025 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. BD-1 Способен осуществлять поиск сбор очистку и предварительный анализ данных	BD-1.1 Обосновывает способы и варианты применения методов предварительного анализа данных в задачах ИИ, включая их математическое (алгоритмическое) преобразование и адаптацию к специфике задачи	знания: (Б) Знает основные статистические характеристики данных и основные методы визуализации данных. умения: (П) Умеет проводить одномерный и многомерный анализ признаков, в том числе с использованием средств визуализации. навыки: (П) Владеет инструментарием одномерного и многомерного анализа признаков, в том числе инструментами визуализации.
	BD-1.2 Применяет методы анализа данных для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ	знания: (Б) Знает методы заполнения пропусков в данных и удаления выбросов в табличных данных умения: (П) Умеет производить очистку зашумленных временных рядов и изображений. Умеет обнаруживать и устранять выбросы в данных временных рядов. Умеет заполнять пропуски в данных временных рядов и изображений. навыки: (П) Владеет инструментарием очистки зашумленных временных рядов и изображений, обнаружения и устранения выбросов в данных временных рядов. Владеет методами заполнения пропусков в данных временных рядов и изображений.
	BD-1.3 Применяет методы понижения размерности для первичной интерпретации и визуализации многомерных данных	знания: (Б) Знает методы понижения размерности и способы их применения для первичной интерпретации и визуализации многомерных данных умения: (П) Умеет применять основные методы понижения размерности и подбирать оптимальную размерность в зависимости от необходимой доли объяснённой дисперсии. навыки: (П) Владеет инструментарием, реализующим методы понижения размерности и позволяющим подбирать оптимальную размерность в зависимости от необходимой доли объяснённой дисперсии.

	BD-1.4 Отбирает признаки данных, значимые для исследования	знания: (П) Знает основы методов отбора признаков и выбора оптимального подмножества признаков. умения: (П) Умеет применять методы отбора признаков и выбора оптимального подмножества признаков. навыки: (П) Владеет инструментарием отбора признаков и выбора оптимального подмножества признаков.
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: (Б) Знает принципы проверки данных на корректность, принципы разметки данных, принципы формирования датасетов для обучения, валидации и тестирования. умения: (П) Умеет формулировать задачи разметки и оценивать качество работы разметчиков навыки: (П) Владеет инструментами разметки данных и средствами автоматизации оценки качества работы разметчиков
	BD-2.2 Работает с данными, в том числе собирает данные из разрозненных источников, проверяет данные на корректность	знания: (П) Знаком с вариантами инструментария разметки данных в разрезе условий задач. умения: (П) Умеет подбирать инструментарий разметки под условия задачи. Умеет организовать краудсорсинг разметки. навыки: (П) Владеет инструментариями разметки данных. Владеет навыками организации и управления краудсорсингом разметки.
	BD-2.3 Применяет инструменты и практики непрерывной интеграции данных (DataOps)	знания: (Б) Знаком с инструментами интеграции данных. умения: (П) Способен участвовать в процессе непрерывной интеграции данных (DataOps). навыки: (П) Владеет технологиями непрерывной интеграции данных (DataOps).
3. BD-4 Способен применять различные модели и (или) технологии обработки данных	BD-4.1 Осуществляет выбор технологий обработки больших данных, приемлемых для создания прикладной системы ИИ с заданными требованиями	знания: (П) Знаком с современными технологиями (Hadoop, Spark и т. п.) больших данных. умения: (П) Способен организовывать распределенное хранилище и параллельную обработку на базе современных технологий (Hadoop, Spark) больших данных . навыки: (Э) Владеет технологиями организации централизованных хранилищ данных (Data Lake), их распределенного хранения, параллельной обработки, а также обработки потоковых данных.

	BD-4.2 Разрабатывает и отлаживает прикладные решения с элементами ИИ с применением различных технологий обработки данных	знания: (П) Знаком с различными технологиями обработки данных, используемыми в сфере ИИ. умения: (П) Разрабатывает и отлаживает прикладные решения с элементами ИИ с применением различных технологий обработки данных. навыки: (П) Владеет инструментарием разработки и отладки прикладных решений с элементами ИИ.
	BD-4.3 Тестирует, испытывает и оценивает качество решений с элементами ИИ, реализованных с использованием технологий обработки данных	знания: (П) Знаком с концепциями качества программного обеспечения. Знает методы испытаний решений с элементами ИИ параллельной и потоковой обработки распределенных данных. умения: (П) Испытывает решения с элементами ИИ параллельной и потоковой обработки распределенных данных. навыки: (П) Владеет инструментарием испытаний решений с элементами ИИ параллельной и потоковой обработки распределенных данных.
	4. FC-5 Способен проводить фронтальные исследования в области безопасности, доверия и объяснимости	
	FC-5.1 Обеспечивает объяснения причин принятия тех или иных решений в результатах работы искусственного интеллекта	знания: (П) Знает методы post-factum объяснения (интерпретации) непрозрачных методов искусственного интеллекта на основе моделей машинного обучения. умения: (П) Создает методы post-factum объяснения (интерпретации) непрозрачных методов искусственного интеллекта на основе моделей машинного обучения, в т. ч. нейронных сетей. навыки: (П) Владеет инструментарием поддержки методов объяснения (интерпретации) непрозрачных методов искусственного интеллекта на основе моделей машинного
	FC-5.2 Обеспечивает отсутствие случайных или добавленных уязвимостей в системах искусственного интеллекта	знания: (П) Знает методы защиты от атак на датасеты (в т. ч. от отравления и закладок). умения: (П) Умеет применять методы и программные инструменты защиты от атак на датасеты (в т. ч. от отравления и закладок). навыки: (П) Владеет инструментарием разработки программных инструментов защиты от атак на датасеты (в т. ч. от отравления и

5. ML-2 Способен применять фундаментальные принципы и методы машинного обучения включая подготовку данных оценку качества моделей и работу с признаками	ML-2.1 Различает основные типы задач машинного обучения и применяет на практике принципы их решения	знания: (Б) Знает основные типы задач машинного обучения (обучением с учителем, без учителя и с подкреплением). Применяет типовые подходы к решению базовых задач с использованием готовых инструментов и библиотек (ScikitLearn). умения: (П) Умеет выбирать и обосновывать методы решения задач машинного обучения с учётом характеристик данных и бизнес-контекста, настраивать базовые модели и проводить их оценку. навыки: (П) Владеет базовым инструментарием решения задач машинного обучения с учётом характеристик данных и бизнес-контекста.
	ML-2.2 Применяет методы предварительной обработки данных и работы с признаками	знания: (Б) Знаком с методами разведочного анализа данных, работы с пропущенными значениями и выбросами. умения: (П) Умеет проводить разведочный анализ данных, работать с пропущенными значениями и выбросами. навыки: (П) Владеет методами feature engineering: отбора создания и преобразования признаков.
	ML-2.3 Решает проблемы несбалансированных данных и оценивает качество моделей	знания: (Б) Знает основные метрики оценки качества для задач классификации и регрессии. умения: (П) Применяет различные типы кросс-валидации. Оценивает качество моделей с учетом bias-variance trade-off. навыки: (П) Владеет базовым инструментарием оценки качества для задач классификации и регрессии.
6. ПК-4 Владение навыками использования операционных систем, сетевых технологий, средств разработки программного интерфейса, применения языков и методов формальных	ПК-4.1 Знает принципы работы с основными операционными системами, базовые сетевые технологии, организацию средств разработки программного обеспечения	знания: Знает принципы работы с основными операционными системами, базовые сетевые технологии, организацию средств разработки программного обеспечения умения: навыки:
	ПК-4.2 Умеет применять современные средства и языки программирования	знания: умения: Умеет применять современные средства и языки программирования навыки:
	ПК-4.3 Имеет навыки использования операционных систем и современных сетевых технологий	знания: умения: навыки: Имеет навыки использования операционных систем, средств разработки программного обеспечения и современных сетевых технологий

Раздел 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к части, формируемой участниками образовательных отношений ОПОП.

Дисциплина является обязательной

Изучаемая дисциплина является основой для продолжения формирования указанных компетенций в следующих дисциплинах: Управление данными (BD-2), Управление данными (BD-4), Функциональное и логическое программирование (BD-4), Методы и средства защиты компьютерной информации (FC-5), Функциональное и логическое программирование (FC-5), Основы машинного обучения (FC-5), Перспективное программное обеспечение (FC-5), Основы технологий искусственного интеллекта (ML-2), Контроль качества (ML-2), Компьютерная вирусология (ML-2), Основы машинного обучения (ML-2), Основы генеративных систем искусственного интеллекта (ML-2), Большие языковые модели (ML-2); практиках: Производственная практика. Проектно-технологическая практика (BD-1), Производственная практика. Проектно-технологическая практика (BD-2), Производственная практика. Проектно-технологическая практика (BD-4), Производственная практика. Проектно-технологическая практика (FC-5), Производственная практика. Проектно-технологическая практика (ML-2); государственной итоговой аттестации в форме: Подготовка к процедуре защиты и защита выпускной квалификационной работы (BD-1), Подготовка к процедуре защиты и защита выпускной квалификационной работы (BD-2), Подготовка к процедуре защиты и защита выпускной квалификационной работы (BD-4), Подготовка к процедуре защиты и защита выпускной квалификационной работы (FC-5), Подготовка к процедуре защиты и защита выпускной квалификационной работы (ML-2)

Раздел 3. ОПИСАНИЕ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

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

Основными стратегическими технологиями являются: дискуссионные, игровые процедуры, исследовательские, лекционные занятия, практические и лабораторные занятия, процедуры самообучения, доступ обеспечивается с помощью электронного курса и сервисов образовательного портала пгту на платформе lms moodle (elearning.volgatech.net), интегрированного с федеральной платформой гис сдос online.edu.ru открытый образовательный портал (moored.net).

На достижение конкретных целей обучения направлены применяемые тактические технологии: задания, информационные, классическая лекция, лекция с элементами мозгового штурма, лекция-провокация, мини-проекты, проблемная лекция, лекция с обсуждением новых научных достижений в сфере ии по результатам конференции и публикаций и др.), семинары с ит-практиками, конкурсы, хакатоны, кейс-чемпионаты, мастер-классы, командные проектные задания и др.

Раздел 4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3 семестр

Виды и темы занятий	Количество часов	Формируемые компетенции
Основные методы статистического анализа данных. Анализ одномерных данных	41	BD-1, BD-2, BD-4, FC-5, ML-2, ПК-4
Лекция. Лекция №1. Введение в анализ данных.	1	

1. Статистические данные и их роль в ИИ. 2. Классификация и структура наборов данных для ИИ-моделей. 3. Типы данных и переменных в контексте машинного обучения.	
Практическое занятие. Практическое занятие №1. Источники данных и предварительная автоматизированная обработка с применением ИИ. Виды представления данных для ИИ. Поиск, сбор и автоматизация предварительной обработки данных.	2
Лекция. Лекция №2. Разведочный анализ данных. 1. Анализ структуры и характеристики данных. 2. Выявление аномалий и выбросов в данных. 3. Идентификация связи и корреляции между переменными. 4. Подготовка данных к обучению моделей машинного обучения.	1
Практическое занятие. Практическое занятие №2. Автоматизированный сбор, очистка и разведочный анализ данных с AI-алгоритмами.	2
Лекция. Лекция №3. Статистические характеристики рядов распределения с ИИ-анализом. 1. Расчет и визуализация статистик средствами ИИ. 2. Использование AI-алгоритмов для создания гистограмм, кумулянт и полигонов.	1
Практическое занятие. Практическое занятие №3. Расчет статистических характеристик с применением ИИ-алгоритмов и визуализации данных.	2
Лекция. Лекция №4. Анализ статистических характеристик рядов распределения с ИИ-алгоритмами. Типические значения. 1. Структурные средние. 2. Расчет моды и медианы для дискретного ряда. 3. Расчет моды и медианы для интервального ряда. 4. Среднеквадратические отклонения. 5. Дисперсия. 6. Коэффициент вариации.	1
Практическое занятие. Практическое занятие №4. Глубокий статистический анализ с AI-элементами и расчет характеристик изменчивости для вариационного ряда.	2
Лекция. Лекция №5. Анализ эмпирических и теоретических моментов для повышения точности, обработки выбросов и нормализации данных. 1. Понятие моментов. 2. Условные эмпирические моменты. 3. Эмпирические и теоретические частоты.	1
Практическое занятие. Практическое занятие № 5. Применение ИИ для расчета типических значений, анализа дисперсии, проверки гипотез и построения вариационных рядов, проверка выборки на нормальность распределения.	2
Лекция. Лекция №6. Методы дисперсионного анализа с поддержкой ИИ. 1. Сущность дисперсионного анализа. 2. Математическая постановка. 3. Проверка гипотезы о нормальном распределении.	1

Практическое занятие. Практическое занятие № 6. Использование инструментов ИИ и библиотек машинного обучения для проверки гипотезы о равенстве групповых средних методом дисперсионного анализа.	2	
Лекция. Лекция №7. Анализ статистических рядов распределения и их видов на основе ИИ-алгоритмов. 1. Оценка и моделирование атрибутивных и вариационных рядов распределения. 2. Оценка и моделирование дискретных и интервальных рядов распределения. 3. Построение интервального вариационного ряда.	1	
Практическое занятие. Практическое занятие №7. Построение интервального вариационного ряда на основе алгоритмов машинного обучения, разведочного анализа.	2	
Задания для самостоятельной работы, в том числе выполнение Изучение лекционного материала и подготовка к текущему контролю. Подготовка к практике. Изучение исследований и разработок, представленных в научных публикациях на конференциях уровня А* (International Conference on Machine Learning, International Conference on Data Engineering, IEEE International Conference on Data Mining), статьях, опубликованных в журналах 1 уровня Белого списка научных журналов Минобрнауки России (МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПРИНЯТИЕ РЕШЕНИЙ)(УММ на основе публикаций в научных журналах 1 или 2 уровня Белого списка научных журналов Минобрнауки России, публикаций на конференциях уровня А*/А). Изучение дополнительного материала на онлайн-курсах Предварительная обработка данных в машинном обучении: инструкция, инструменты, полезные ресурсы для начинающих: https://habr.com/ru/companies/skillfactory/articles/848858/ , Разведочный анализ (EDA): https://habr.com/ru/companies/otus/articles/752434/	20	
Анализ двумерных данных. Анализ временных рядов	32	BD-1, BD-2, BD-4, FC-5, ML-2, ПК-4
Лекция. Лекция №8. Анализ временных данных и динамических рядов с применением методов ИИ. 1. Основные элементы временного ряда. 2. Основные аналитические характеристики динамических рядов: Абсолютный прирост, темп роста, темп прироста. 3. Использование ИИ для анализа трендов и сезонных компонент.	1	
Практическое занятие. Практическое занятие №8. Автоматизированное определение аналитических характеристик динамических рядов.	2	
Лекция. Лекция №9. Определение и оценка зависимости признаков с использованием корреляционного и	1	

регрессионного анализа в ИИ. 1. Степень линейной зависимости количественных признаков-корреляционный анализ. 2. Виды коэффициентов корреляции. Оценка на значимость.		
Практическое занятие. Практическое занятие №9. Применение ИИ-инструментов для анализа временных рядов	2	
Лекция. Лекция №10. Непараметрические методы в анализе данных на основе AI-алгоритмов. 1. Ранговый коэффициент Спирмена. 2. Ранговый коэффициент Кендалла. 3. Коэффициент конкордации Кендалла.	1	
Практическое занятие. Практическое занятие №10. Методы ранжирования и непараметрического анализа с поддержкой статистического ИИ.	2	
Лекция. Лекция №11. Индексный метод анализа данных с ИИ-алгоритмами. 1. Количественные индексы. 2. Качественные индексы. 3. Агрегатные индексы средних величин.	1	
Практическое занятие. Практическое занятие №11. Анализ данных на основе индексного метода и алгоритмов машинного обучения.	2	
Задания для самостоятельной работы, в том числе выполнение Изучение лекционного материала и подготовка к текущему контролю. Подготовка к практическим занятиям. Изучение исследований и разработок, представленных в научных публикациях на конференциях уровня А* (International Conference on Machine Learning, International Conference on Data Engineering, IEEE International Conference on Data Mining), статьях, опубликованных в журналах 1 уровня Белого списка научных журналов Минобрнауки России (МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПРИНЯТИЕ РЕШЕНИЙ)(УММ на основе публикаций в научных журналах 1 или 2 уровня Белого списка научных журналов Минобрнауки России, публикаций на конференциях уровня А*/А). Изучение дополнительного материала на онлайн-курсах Предварительная обработка данных в машинном обучении: инструкция, инструменты, полезные ресурсы для начинающих: https://habr.com/ru/companies/skillfactory/articles/848858/ , Разведочный анализ (EDA): https://habr.com/ru/companies/otus/articles/752434/ , Искусственный интеллект (Вводный курс) https://stepik.org/course/241239/promo?search=8085244381#toc	20	
Анализ многомерных данных	35	BD-1, BD-2, BD-4, FC-5, ML-2, ПК-4
Лекция. Лекция №12. Типы задач машинного обучения. 1. Классификация объектов исследования.	1	

2. Регрессионный анализ. 3. Кластеризация многомерных данных. 4. Обработка естественного языка.	
Практическое занятие. Практическое занятие №12. Анализ взаимосвязей между многомерными признаками с использованием ИИ.	2
Лекция. Лекция №13. Кластерный анализ и алгоритмы метод k-средних с оптимизацией и автоматизацией на базе ИИ. 1. Иерархические агломеративные методы. 2. Меры сходства, расстояния (или меры близости между объектами). 3. Расстояние между группами объектов и меры близости двух групп объектов.	1
Практическое занятие. Практическое занятие №13. Методы иерархической кластеризации данных.	2
Лекция. Лекция №14. Кластерный анализ. Метод k-средних. Функционалы качества разбиения. 1. Алгоритм метода k-средних. 2. Анализ качества кластеризации- функционалы.	1
Практическое занятие. Практическое занятие №14. Дивизимные алгоритмы формирования выборок данных.	4
Лекция. Лекция №15. Снижение размерности анализируемого признакового пространства. 1. Факторный анализ. 2. Метод главных компонент.	2
Практическое занятие. Практическое занятие №15. Построение интегральных показателей анализируемого многомерного пространства.	4
Лекция. Лекция №16. Анализ степени зависимости многомерных данных. 1. Корреляционный анализ многомерных данных. 2. Математическая модель множественного корреляционного анализа. 3. Разбор примера.	2
Практическое занятие. Практическое занятие №16. Анализ степени зависимости и множественные корреляционные модели с поддержкой ИИ.	2

Задания для самостоятельной работы, в том числе выполнение Изучение лекционного материала и подготовка к текущему контролю, Подготовка к практическим занятиям. Изучение исследований и разработок, представленных в научных публикациях на конференциях уровня А* (International Conference on Machine Learning, International Conference on Data Engineering, IEEE International Conference on Data Mining), статьях, опубликованных в журналах 1 уровня Белого списка научных журналов Минобрнауки России (МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ, ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПРИНЯТИЕ РЕШЕНИЙ)(УММ на основе публикаций в научных журналах 1 или 2 уровня Белого списка научных журналов Минобрнауки России, публикаций на конференциях уровня А*/А). Изучение дополнительного материала на онлайн-курсах Предварительная обработка данных в машинном обучении: инструкция, инструменты, полезные ресурсы для начинающих: https://habr.com/ru/companies/skillfactory/articles/848858/ , Разведочный анализ (EDA): https://habr.com/ru/companies/otus/articles/752434/ , Искусственный интеллект (Вводный курс) https://stepik.org/course/241239/promo?search=8085244381#toc	14
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.

Содержание дисциплины включает изучение исследований и разработок, представленных в научных публикациях на конференциях уровня А*, статьях, опубликованных в журналах 1 уровня Белого списка научных журналов Минобрнауки России (УММ на основе публикаций в научных журналах 1 или 2 уровня Белого списка научных журналов Минобрнауки России, публикаций на конференциях уровня А*/А).

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.

Предусмотрено участие ИТ-практиков в различных форматах обучения (лекции, семинары с

ИТ-практиками, конкурсы, хакатоны, кейс-чемпионаты, мастер-классы, командные проектные задания и др.) и в качестве наставников.

В программе предусмотрено изучение кейсов (кейсы для изучения, практической и проектной работы) и реальных задач (проектов по разработке, внедрению, развитию цифровых технологий, ИТ-разработок и др.) из практического опыта Индустриальных партнеров в объеме не менее 1 кейса на 1 з.е.

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 (на платформе LMS Moodle (elearning.volgatech.net), интегрированного с федеральной платформой ГИС СЦОС online.edu.ru открытый образовательный портал (moored.net) и ресурсы открытых образовательных платформ Яндекс-Практикум, Сберуниверситет.

Периодичность проведения, формы текущего контроля успеваемости, система оценивания хода освоения дисциплин представлены в рабочей программе.

Формой промежуточной аттестации по дисциплине является зачет.

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.	Макшанов, А. В. Технологии интеллектуального анализа данных [Электронный ресурс] : учебное пособие / А. В. Макшанов, А. Е. Журавлев. 2-е изд., стер. Санкт-Петербург: Лань, 2022. - 212 с. ISBN 978-5-8114-4493-9.	https://e.lanbook.com/book/206711
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.	Кулаичев, Алексей Павлович. Методы и средства комплексного анализа данных [Текст] : учеб. пособие для вузов по дисциплинам "Прикладная статистика" и "Информатика" / А. П. Кулаичев. 4-е изд., перераб. и доп. М.: Формула ИНФРА-М, 2006. - 511 с. ISBN 5-8199-0234-3. Экземпляры: всего 19.	19

4.	Вуколов, Эдуард Александрович. Основы статистического анализа [Текст] : практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL : [учеб. пособие по специальности "Менеджмент орг."] / Э. А. Вуколов. 2-е изд., испр. и доп. М.: ФОРУМ, 2012. - 463 с. ISBN 978-5-91134-231-9. Экземпляры: всего 14.	14
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

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.	429 (III)	ПК RAMEC GALE/i7-3770/B75M2x4DDR3/GTX650/500S ATA3/монит.LCD PHILIPS 23,6"клав.,м (8), Принтер HP LaserJet Professional P1102 (1), Проектор VIEWSONIC PJD6550LW белый (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
2.	521 (I)	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
3.	522 (I)	Анализатор спектра NS-30A (1), Антенна M102 в компл. с кабелем ВЧ TNCm-SMAm (1), Блок питания лаборат. НУ 3003 D-3 (1), Внешний HDD WD 2TB 3.0 , 3.5"USB (1), Внешний накопитель 1 Seagate Original USB 3.0 4 Tb (1), Внешний накопитель флешка USB TRANSCEND Jetflash 780 64 Gb (1), Гигабитный управляемый коммутатор на 16 портов (1), Измеритель CN -801 HP (1), Кондиционер AEG ACS-09HR (1), Многофункциональный измерительный прибор (1), Монитор 20 "Beng FP 202W (2), Монитор LCD Samsung 17" SM 713N (1), МФУ Canon i-SENSYS MF 4018 (1), МФУ 1 Лазерный Canon i-Sensys MF226 (1), Набор ВЧ переходников (1), Ноутбук Dell Latitude E6520 Intel Core I5 Processor 2520M 15,6" (2), Ноутбук TOSHIBA Satellite L655-1H2-RU (1), Паяльная станция AOYUE 968 (1), Переключатель ZX80-DR230 (1), Персональный компьютер 3 Atlant A2X4/4G(3)/512Mb/монитор Pyama 2209/3Y (1), ПК RAMEC GALE LCD LG 23"/Intel i5 4590/MSI B85M- E45/2x4DDR3/GT740 2Gb/500Gb/клав (28), Преобразователь SP-200-24-AC-DC в кожухе 199x99x50мм (1), Приемо- передающая программно- конфигурируемая радиоплатформа G32 (1), Принтер Canon LBP 2900 лазерный с кабелем (1), Проектор мультимедийный Hitachi CP-EX250 (1), Проектор мультимедийный Hitachi CP-EX251N (1), Сист. блок Pen D 945 3.4 DDR 2 1024*2/FDD 3.5/250 Gb/DVD- RW/кл+мышь+коврик (1), Систем.блок CPU Intel Core i7- 6700/ASRod Z-170/32Gb/GTX 1070/200Gb/Wi-Fi +клав, (1), Станок сверлильный 350 Вт (1), Универсальная приёмо-передающая платформа для проектирования	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

	СВЧ-систем компл.мг (1), Усилитель LZY-22 (1), Усилитель ZHL-3A-S (1), Комплект учебной мебели (1)	
--	---	--

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.Определить по данному дискретному вариационному ряду средний курс продажи одной акции.			
Сделка	Количество проданных акций, шт ()	Курс продажи, руб. (/span>)	
1	500	1080	
2	300	1050	
3	1100	1145	
2.Вычислить медиану по данным таблицы, в которой приведена информация об успеваемости по			

математике 100 учащихся 7-х классов (успеваемость оценивается по 12-бальной шкале).

Количество баллов	Число учащихся
1	3
2	4
3	4
4	9
5	11
6	12
7	18
8	14
9	9
10	8
11	6
12	2

3. В результате обобщения итогов выборочного бюджетного обследования населения города N-ска построен вариационный ряд, отражающий распределение жителей города по величине среднедушевого дохода (табл., гр. 1—3). Построить гистограмму, полигон и кумуляту для данного вариационного ряда.

Среднедушевой денежный доход в среднем за месяц, тыс. руб.	Число жителей чел.	в % к итогу)
1	2	3
До 5	26	0,9
5-10	463	16,5
10-15	690	24,6
15-20	528	18,8
20-25	434	15,4
25-30	350	12,5
30 и более	318	11,3
Итого	2809	100,0

4. На основе вариационного ряда темпов роста объема производства предприятий металлургии требуется вычислить:

Темп роста объема производства	102-104	104-106	106-108	108-110	110-112
Количество предприятий	7	14	13	9	7

а) среднее значение темпов роста объема производства предприятий металлургии;

б) дисперсию и средние квадратическое отклонения;

в) коэффициент вариации;

г) коэффициенты асимметрии, эксцесса.

5. По иерархическому агломеративному алгоритму провести классификацию $n=10$ десяти страховых компаний по наличию активов, тыс. руб. и чистой прибыли организаций, тыс. руб. данные представлены в таблице.

Страховые компании	Активы, тыс. руб	Чистая прибыль, тыс. руб.
Группа "Росэнерго"	650550	7984
Группа "УралСиб"	8881517	84578
Мегарусс	1246162	32763
Мегарусс-Д	1214055	647
Медэкспресс	294176	312
Москва	852515	7632
Росмедстрах	362580	10851
РОСНО	8103652	34213
СОГАЗ	8718438	320862
Согласие	7091756	8103

За расстояние между объектами принять взвешенное Евклидово расстояние, а расстояние между кластерами измерять по принципу «ближнего соседа» и «дальнего соседа». Построить дендограммы. Сравнить полученные классификации. Сделать вывод. Выполнить классификацию в пакете Statistica.

6. Построить модель для предсказания цены квартиры на основе площади, расположения, возраста дома.

7. Спрогнозировать объем продаж товара на следующий месяц с учетом сезонности и рекламных кампаний.

8. Сформировать кластеры по типам заболеваний на основе статистики симптомов.

9. Построить модель для определения категории товара (одежда, электроника, книги) на основе описания и характеристик.

Перечень вопросов для проведения промежуточной аттестации

1. Роль аналитики в современном обществе.
2. Структуры данных.
3. Типы данных и переменных.
4. Требования к наборам данных.
5. Классификация различных типов данных: одномерные, двумерные, многомерные.
6. Классификация методов разведочного анализа данных.
7. Анализ структуры и характеристики данных.
8. Выявление аномалий и выбросов в данных.
9. Идентификация связи и корреляция между переменными.
10. Подготовка данных для дальнейшего анализа.
11. Статистические ряды распределения и их виды: Атрибутивный ряд распределения, вариационный ряд распределения, дискретные и интервальные вариационные ряды.
12. Графическое изображение. Гистограммы.
13. Статистические характеристики рядов распределения.
14. Средние величины: степенные и структурные, простые и взвешенные.
15. Обобщающие показатели: медиана, ранг, мода, перцентиль.
16. Изменчивость: стандартное отклонение, дисперсия, размах, коэффициент вариации.
17. Анализ статистической взаимосвязи социально-экономических явлений.
18. Диаграмма рассеяния, коэффициент корреляции (парный, частный, множественный), непараметрическая статистика.
19. Ранговый коэффициент корреляции Спирмена.
20. Коэффициент корреляции Кендалла.
21. Коэффициент конкордации (согласованности) Кендалла.
22. Типы задач машинного обучения.
23. Регрессионный анализ данных.
24. Обработка естественного языка.
25. Основные методы многомерного анализа данных.
26. Методы классификации.
27. Классификация без обучения.
28. Кластерный анализ.
29. Кластер, группа, класс, таксон.
30. Метрики, евклидово расстояние, меры сходства, дендограмма.
31. Иерархические агломеративные методы классификации.
32. Итеративные кластерные процедуры - агломеративные методы, дивизимные методы.

- 33. Итеративные методы - метод К-средних.
- 34. Критерии качества классификации.
- 35. Факторный анализ в снижении размерности признакового пространства.
- 36. Сущность и алгоритм метода главных компонент.