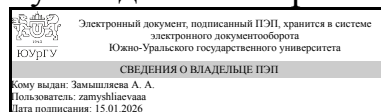


ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Руководитель направления



А. А. Замышляева

РАБОЧАЯ ПРОГРАММА

дисциплины 1.О.17 Сбор, анализ и предобработка данных в машинном обучении
для направления 09.03.03 Прикладная информатика

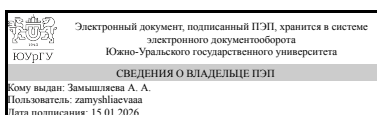
уровень Бакалавриат

форма обучения очная

кафедра-разработчик Центр ОП топ-уровня в сфере ИИ "ВиртУм"

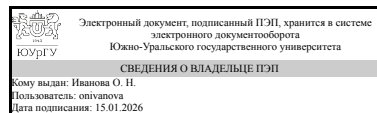
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 09.03.03 Прикладная информатика, утвержденным приказом Минобрнауки от 19.09.2017 № 922

Зав.кафедрой разработчика,
д.физ.-мат.н., проф.



А. А. Замышляева

Разработчик программы,
к.пед.н., доцент



О. Н. Иванова

1. Цели и задачи дисциплины

Цель дисциплины заключается в формировании у студентов глубоких знаний и практических навыков, необходимых для сбора, предобработки и предварительного анализа данных и эффективного применения статистических методов при подготовке датасетов в различных сферах. Дисциплина направлена на развитие системного и критического мышления, способности к декомпозиции сложных задач и качественной подготовке нестандартных наборов данных. Задачи дисциплины: 1) углубленное изучение этапов аналитического проекта, от формулировки исследовательских вопросов до оценки результатов, а также освоение методов сбора и систематизации данных; 2) применение методов проверки статистической значимости и методов предобработки данных для машинного обучения, включая методы работы с аномальными значениями, а также трансформации данных, в том числе нормализации и стандартизации; 3) изучение подходов к синтетическим табличным, текстовым, графическим, видео и аудио данным, их генерации и анонимизации, обработке и работе с категориями переменных; 4) развитие целостного представления о современных процедурах и инструментах анализа данных и навыков экспертизы данных.

Краткое содержание дисциплины

Задачи аналитика данных. Уточнение и декомпозиция исследовательских задач. Стадии проекта. Методы сбора и систематизации фактов. Модель. Обучающая, валидационная и тестовая выборки. Генеральная совокупность. Мощность и размер выборки. Объект и признак. Разведочный анализ данных, анализ тенденций. Причины ошибок в данных. Системное и критическое мышление. Элементы структурированных данных. Оценки центрального положения и вариабельности. Обследование распределения данных. Обследование двоичных и категориальных данных. Выбор признаков. Дискретизация. Корреляция. Исследование двух и более переменных (многофакторный анализ). Проверка значимости. А/В тестирование. Проверка статистических гипотез. Повторный отбор. Статистическая значимость и р-значения. Проверка на основе t-статистики. Множественное тестирование. Степени свободы. ANOVA. Проверка на основе хи-квадрат. Мощность и размер выборки. Методы предобработки данных для машинного обучения: выявление и обработка пропусков. Конверсия. Куки. Методы поиска и удаления дубликатов. Декомпозиция таблиц. Категоризация по числовым диапазонам. Категоризация на основе нескольких значений в строке. Разреженные, отсутствующие и неточные данные. Добыча сложных и истинных сэмплов. Проверка чувствительности. Преобразование данных: нормализация, стандартизация, методы работы с категориальными переменными (одноразрядное кодирование, создание новых признаков, ранжирование). Конвейер преобразований. Обработка аномальных значений. Удаление выбросов. Замена значений на среднее или медианное. Корректировка значений на основе правил бизнеса. Модулярная предобработка. Функциональная предобработка. Позиционная предобработка. Методы сбора и обработки данных из сети Интернет. Веб-скрапинг, API, базы данных. Генерация синтетических данных. Аугментация. Рандомизаторы. Метод Монте-Карло. Метод главных компонент (применение). Генерация с помощью нейронных сетей (вариационные автоэнкодеры, генеративно-состязательные модели, авторегрессионные модели). Частичные и полные синтетические данные. Генерация

табличных данных. Генерация персональных данных, анонимизация. Маркирования, псевдомаркировка. Предварительная обработка текстов. Удаление шума и несоответствий. Нормализация, токенизация, удаление стоп-слов, лемматизация и стемминг, векторизация. Использование библиотек NLTK, TextBlob и CoreNLP. Шкалирование и категориальные переменные. Шкалирование переменных. Доминантные переменные. Категориальные данные и расстояние Говера. Проблемы смешанных данных. Применение преобразования Фурье и вейвлет-анализа. Введение в уменьшение размерности данных с помощью методов PCA, t-SNE и UMAP. Методы балансировки классов для работы с несбалансированными данными различных степеней. Понижающий и повышающий отбор. Понижающая и повышающая перевесовка. Методы проведения экспертиз для выявления наиболее значимых факторов. Понятия парциальной и комплексной обработки. Интерпретация результатов предварительного анализа данных

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-8 [BD-1] Способен осуществлять поиск, сбор, очистку и предварительный анализ данных	Знает: - [И-1, БУ] основные числовые характеристики статистических данных и методы их нахождения Умеет: - [И-1, БУ] вычислять числовые характеристики статистических данных, применять методы визуализации данных
ПК-9 [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	Знает: - [И-1, БУ] основные требования к наборам данных для решения задач машинного обучения Умеет: - [И-2, ПУ] подбирать инструментарий разметки данных под условия задачи Имеет практический опыт: - [И-1, БУ] проверки данных на корректность, разметки данных

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
1.О.09 Введение в проектную деятельность	1.О.25 Анализ временных рядов, 1.О.22 Машинное обучение, Производственная практика (проектно-технологическая, стажировка) (4 семестр)

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Дисциплина	Требования
1.О.09 Введение в проектную деятельность	Знает: понятие информационной культуры и роль информационно-коммуникационных технологий (ИКТ) в проектной деятельности Умеет: -[И-2, БУ] подготавливать данные для дальнейшей обработки и анализа, организовывать

	совместную работу членов команды посредством ИКТ (совместные онлайн-документы, сервисы группового общения); соблюдать правила информационной безопасности при хранении и передаче конфиденциальных данных Имеет практический опыт: групповой коммуникации и взаимодействия внутри коллектива через дистанционные средства связи
--	---

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 3 з.е., 108 ч., 70,25 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам
		в часах
		Номер семестра
		3
Общая трудоёмкость дисциплины	108	108
<i>Аудиторные занятия:</i>	64	64
Лекции (Л)	32	32
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	0	0
Лабораторные работы (ЛР)	32	32
<i>Самостоятельная работа (СРС)</i>	37,75	37,75
Подготовка к сдаче зачета	8	8
Изучение научных статей	12	12
Решение кейса от индустриального партнера	14	14
Подготовка к контрольным работам	3,75	3.75
Консультации и промежуточная аттестация	6,25	6,25
Вид контроля (зачет, диф.зачет, экзамен)	-	зачет

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Введение в аналитику данных	12	8	0	4
2	Разведочный анализ данных	12	4	0	8
3	Предобработка данных	20	10	0	10
4	Синтетические данные. Понижение размерности. Отбор признаков. Инструменты предварительного анализа данных	20	10	0	10

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Введение в аналитику данных: цели дисциплины, ключевые понятия, этапы аналитического проекта	2

2	1	Формулировка исследовательских вопросов и декомпозиция задач: от общего вопроса к конкретным гипотезам и метрикам	2
3	1	Стратегии сбора и систематизации данных: источники, методология сбора, документация и качество данных	2
4	1	Архитектура данных и выбор выборки: генеральная совокупность, размер и мощность выборки, обучающие/валидационные/тестовые наборы	2
5	2	Разведочный анализ данных (EDA): распределения, тенденции, пропуски, аномалии	2
6	2	Ошибки в данных и их влияние: источники ошибок, способы обнаружения и коррекции	2
7	3	Введение в предобработку данных для машинного обучения: методы работы с пропусками и дубликатами, нормализация и стандартизация	2
8	3	Преобразование данных: методы масштабирования, кодирование категориальных признаков, создание признаков	2
9	3	Преобразование данных: методы замены, удаления, анализа чувствительности	2
10	3	Обработка категориальных переменных и декомпозиция таблиц: одноразрядное кодирование, ранжирование, хэширование	2
11	3	Предобработка текста: нормализация, токенизация, удаление стоп-слов, лемматизация/стемминг, векторизация	2
12	4	Генерация и анонимизация синтетических данных: подходы к табличным, текстовым, графическим и мультимедийным данным	2
13	4	Введение в методы отбора признаков: корреляции, дискретизация, многофакторный анализ, принципы отбора	2
14	4	Методы снижения размерности: PCA, t-SNE, UMAP и их применение в анализе данных	2
15	4	Статистическая проверка гипотез: t-статистика, ANOVA, хи-квадрат, множественные тестирования	2
16	4	Инструменты и пайплайны анализа данных: конвейеры преобразований, сквозные процедуры, обзор современных инструментов	2

5.2. Практические занятия, семинары

Не предусмотрены

5.3. Лабораторные работы

№ занятия	№ раздела	Наименование или краткое содержание лабораторной работы	Кол-во часов
1	1	Практика сбора данных: источники, методы протоколирования, создание мини-набора данных	2
2	1	Построение генеральной совокупности и определение объема выборки: мощность тестов, размер выборки	2
3	2	Разведочный анализ данных: визуализация distributions, выявление трендов и пропусков	2
4	2	Обнаружение и устранение ошибок в данных: диагностика ошибок, корректирующие процедуры	2
5	2	Предобработка пропусков: имитация пропусков, заполнение пропусками, оценка влияния на итоговые результаты	2
6	2	Очистка дубликатов и дублирование записей: алгоритмы обнаружения и удаления дубликатов	2

7	3	Нормализация и стандартизация: применение различных методов масштабирования	2
8	3	Работа с категориальными переменными: кодирование, ранжирование, создание признаков	2
9	3	Обработка текста: подготовка корпуса, токенизация, удаление шума	2
10	3	Введение в синтетические данные: генерация, анонимизация, оценка достоверности	2
11	3	Генерация табличных и текстовых данных с помощью простых методов: простые генераторы и их параметры	2
12	4	Применение методов Монте-Карло для оценки неопределенности	2
13	4	Введение в PCA и другие методы снижения размерности: реализация и интерпретация результатов	2
14	4	Балансировка классов: методы подготовки сбалансированных наборов	2
15	4	Прогнозирование и оценка качества: разделение на обучающую/валидационную/тестовую выборки, метрики	2
16	4	Решение кейсов от индустриального партнера	2

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Подготовка к сдаче зачета	Основная и доп. литература, отчет. и зарубежные журналы, имеющиеся в библиотеке – см. раздел 8 настоящей программы	3	8
Изучение научных статей	См. методические указания в разделе СРС	3	12
Решение кейса от индустриального партнера	Основная и доп. литература, отчет. и зарубежные журналы, имеющиеся в библиотеке – см. раздел 8 настоящей программы.	3	14
Подготовка к контрольным работам	Основная и доп. литература, отчет. и зарубежные журналы, имеющиеся в библиотеке – см. раздел 8 настоящей программы.	3	3,75

6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учитывается в ПА
1	3	Текущий	Лабораторная	5	5	Лабораторная работа оценивается по шкале	зачет

		контроль	работа 01			от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	
2	3	Текущий контроль	Лабораторная работа 02	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет
3	3	Текущий контроль	Лабораторная работа 03	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет

4	3	Текущий контроль	Лабораторная работа 4	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет
5	3	Текущий контроль	Лабораторная работа 05	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет
6	3	Текущий контроль	Лабораторная работа 06	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно	зачет

						ответил на все вопросы.	
7	3	Текущий контроль	Лабораторная работа 07	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет
8	3	Текущий контроль	Лабораторная работа 08	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет
9	3	Текущий контроль	Лабораторная работа 09	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью,	зачет

						ошибок и замечаний нет, студент правильно ответил на все вопросы.	
12	3	Текущий контроль	Лабораторная работа	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет
13	3	Текущий контроль	Лабораторная работа	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет
14	3	Текущий контроль	Лабораторная работа 14	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя	зачет

						5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	
15	3	Текущий контроль	Лабораторная работа 15	5	5	Лабораторная работа оценивается по шкале от 0 до 5 баллов. 0 баллов - работа не сделана 1 балл - работа выполнена менее, чем на 25%, студент правильно отвечает на вопросы по сделанной части 2 балла - работа выполнена менее, чем на 50%, студент правильно отвечает на вопросы по сделанной части 3 балла - работа выполнена менее, чем на 75%, студент правильно отвечает на вопросы по сделанной части 4 балла - работа выполнена полностью, ошибок и замечаний нет, студент правильно отвечает не на все вопросы и замечания преподавателя 5 баллов - работа выполнена полностью, ошибок и замечаний нет, студент правильно ответил на все вопросы.	зачет
16	3	Текущий контроль	Лабораторная работа 16	9	9	Лабораторная работа оценивается по шкале от 0 до 9 баллов. Особенности процедуры выставления оценки см. в тексте задания.	зачет
17	3	Бонус	Бонусные баллы	-	15	Студент получает по 0,5 бонусного балла за посещение каждого лекционного или лабораторного занятия. За выполнение дополнительных заданий в некоторых практических работах и на лекциях также начисляются дополнительные баллы. Максимальное количество бонусных баллов, которое может получить студент, - 15.	зачет
18	3	Промежуточная аттестация	Финальный тест	-	100	В финальном тесте 25 вопросов. Каждый вопрос оценивается 0.4 баллами. Ограничение по времени на прохождение теста - 50 минут. Вопросы выбираются случайным образом из всех разделов дисциплины, по 4-5 вопросов из каждой темы.	зачет
19	3	Текущий контроль	Контрольная работа 01	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	зачет
20	3	Текущий контроль	Контрольная работа 02	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	зачет
21	3	Текущий контроль	Контрольная работа 03	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время	зачет

[illegible]

						Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	
31	3	Текущий контроль	Контрольная работа 13	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	зачет
32	3	Текущий контроль	Контрольная работа 14	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	зачет
33	3	Текущий контроль	Контрольная работа 15	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	зачет
34	3	Текущий контроль	Контрольная работа 16	1	1	Тест проводится по теме лекции. В тесте 3 вопроса. На прохождение теста дается одна попытка. Тест доступен только во время лекции. Ограничение по времени - 9 минут. Вес каждого вопроса – 0.33. Максимальное количество баллов за тест – 1 балл.	зачет

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
зачет	Зачет проводится в форме теста и выполнения практического задания по билетам. Тест проводится на портале "Электронный ЮУрГУ 2.0". Тестовая база вопросов содержит более 100 вопросов, каждому студенту на зачете система выбирает случайным образом 25 вопросов (по 3-4 вопроса из разных разделов курса). Каждый вопрос оценивается в 1 балл. Баллы суммируются с баллами, набранными студентом в течение семестра, и с баллами, полученными на зачете за выполнение практического задания по билету. Выполнение практического задания по билету необходимо в случае, если суммарное количество баллов в рейтинге, набранных студентом в течение семестра и на зачетном тесте, менее 51. Студент тянет билет. Время на выполнение задания - 20 мин. Задача должна быть формализована и решения на языке VBA. Максимальный балл за выполнение задания - 25. Баллы суммируются с набранными в течение семестра и с итоговым тестом. Если сумма остается менее 51 балла, студенту выставляется оценка "Не зачтено". Из 25 баллов: задача решена и работает на всех тестовых вариантах - 10 баллов, студент отвечает на все вопросы преподавателя по алгоритму и особенностям его реализации - 15 баллов.	В соответствии с пп. 2.5, 2.6 Положения

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ																																	
		1	2	3	4	5	6	7	8	9	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34		
ПК-8	Знает: - [И-1, БУ] основные числовые характеристики статистических данных и методы их нахождения					+		+	+	+		+	+	+	+	+	+			+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
ПК-8	Умеет: - [И-1, БУ] вычислять числовые характеристики статистических данных, применять методы визуализации данных								+	+	+	+	+			+																			
ПК-9	Знает: - [И-1, БУ] основные требования к наборам данных для решения задач машинного обучения					+	+					+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
ПК-9	Умеет: - [И-2, ПУ] подбирать инструментарий разметки данных под условия задачи					+	+			+	+	+	+	+		+																			
ПК-9	Имеет практический опыт: - [И-1, БУ] проверки данных на корректность, разметки данных								+	+		+	+	+	+		+																		

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) основная литература:

Не предусмотрена

б) дополнительная литература:

Не предусмотрена

в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:
Не предусмотрены

г) методические указания для студентов по освоению дисциплины:

1. Методические указания

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Методические указания

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Дополнительная литература	ЭБС издательства Лань	Предобработка данных для нейросетевого управления : учебное пособие / Л. И. Воронова, В. Р. Брус, В. И. Воронов, А. Н. Баширов. — Москва : МТУСИ, 2021. — 49 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/215198 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
2	Дополнительная литература	ЭБС издательства Лань	Шарден, Б. Крупномасштабное машинное обучение вместе с Python : учебное пособие / Б. Шарден, Л. Массарон, А. Боскетти ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2018. — 358 с. — ISBN 978-5-97060-506-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/105836 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
3	Дополнительная литература	ЭБС издательства Лань	Рашка, С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения : руководство / С. Рашка ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2017. — 418 с. — ISBN 978-5-97060-409-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/100905 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
4	Основная литература	ЭБС издательства Лань	Запечников, С. В. Основы интеллектуального анализа данных и машинного обучения: Конспект лекций : учебное пособие / С. В. Запечников. — Москва : НИЯУ МИФИ, 2022. — 136 с. — ISBN 978-5-7262-2856-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/355580 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
5	Основная литература	ЭБС издательства Лань	Газанова, Н. Ш. Методы искусственного интеллекта : учебно-методическое пособие / Н. Ш. Газанова, С. Н. Никольский. — Москва : РТУ МИРЭА, 2023. — 102 с. — ISBN 978-5-7339-1805-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/368756 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
6	Дополнительная	ЭБС	Прокопенко, Н. Ю. Аналитические информационные

	литература	издательства Лань	системы поддержки принятия решений : учебное пособие / Н. Ю. Прокопенко. — Нижний Новгород : ННГАСУ, 2020. — 142 с. — ISBN 978-5-528-00395-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/164866 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
7	Основная литература	ЭБС издательства Лань	Акулова, Е. А. Автоматизированная обработка измерений в сети планово-высотного обоснования и топографической съемки : учебно-методическое пособие / Е. А. Акулова. — Екатеринбург : УГГУ, 2023 — Часть 1 — 2023. — 72 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/453659 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
8	Основная литература	ЭБС издательства Лань	Золкин, А. Л. Инструментальные средства разработки интеллектуальных информационных систем : учебник для вузов / А. Л. Золкин. — Санкт-Петербург : Лань, 2025. — 140 с. — ISBN 978-5-507-51532-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/450848 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
9	Дополнительная литература	ЭБС издательства Лань	Ферлитш, Э. Шаблоны и практика глубокого обучения / Э. Ферлитш ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2022. — 538 с. — ISBN 978-5-93700-113-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/241199 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
10	Основная литература	ЭБС издательства Лань	Золкин, А. Л. Реализация принципов организации и использования средств машинного обучения и искусственного интеллекта в медицине : учебное пособие / А. Л. Золкин, В. Д. Мунистер. — Самара : , 2024. — 123 с. — ISBN 978-5-907359-23-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/429719 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
11	Основная литература	ЭБС издательства Лань	Баланов, А. Н. Цифровая трансформация. Финансовые услуги и банковское дело : учебное пособие для вузов / А. Н. Баланов. — Санкт-Петербург : Лань, 2024. — 564 с. — ISBN 978-5-507-49589-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/422564 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
12	Основная литература	ЭБС издательства Лань	Конкина, В. В. Введение в большие данные и анализ информации : учебное пособие / В. В. Конкина, А. Б. Борисенко, И. Л. Коробова. — Тамбов : ТГТУ, 2024. — 82 с. — ISBN 978-5-8265-2749-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/472331 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
13	Дополнительная литература	ЭБС издательства Лань	Предобработка данных для нейросетевого управления : учебное пособие / Л. И. Воронова, В. Р. Брус, В. И. Воронов, А. Н. Баширов. — Москва : МТУСИ, 2021. — 49 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/215198 (дата

			обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
14	Основная литература	ЭБС издательства Лань	Алетдинова, А. А. Интеллектуальный анализ больших данных : учебное пособие / А. А. Алетдинова, М. Ш. Муртазина. — Новосибирск : НГТУ, 2023. — 66 с. — ISBN 978-5-7782-4899-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/404567 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.
15	Дополнительная литература	ЭБС издательства Лань	Котиков, П. Е. Анализ данных : учебно-методическое пособие / П. Е. Котиков. — Санкт-Петербург : СПбГПМУ, 2019. — 48 с. — ISBN 978-5-907184-46-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/174498 (дата обращения: 28.11.2025). — Режим доступа: для авториз. пользователей.

Перечень используемого программного обеспечения:

1. ЗАО СПСС Русь-SPSS (Base 14, Tables, Regression Models, Advanced Models, Trends и др.)(бессрочно)
2. Python Software Foundation-Python (бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

1. -База данных ВИНТИ РАН(бессрочно)

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Зачет		Компьютер преподавателя оснащен выходом в интернет. В аудитории есть Wi-Fi сеть или компьютеры для студентов, оснащенные выходом интернет
Лекции		Компьютер преподавателя оснащен проектором и выходом в интернет
Лабораторные занятия		Компьютеры студентов оснащены выходом в интернет. Компьютер преподавателя оснащен проектором и выходом в интернет.