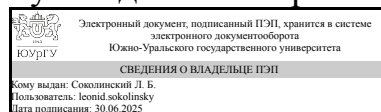


ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Руководитель направления



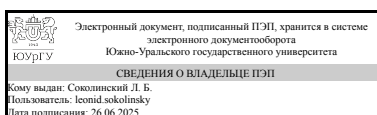
Л. Б. Соколинский

РАБОЧАЯ ПРОГРАММА

дисциплины 1.О.12 Интеллектуальный анализ данных
для направления 09.04.04 Программная инженерия
уровень Магистратура
форма обучения очная
кафедра-разработчик Системное программирование

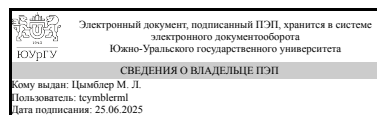
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 09.04.04 Программная инженерия, утверждённым приказом Минобрнауки от 19.09.2017 № 932

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



Л. Б. Соколинский

Разработчик программы,
д.физ.-мат.н., доц., профессор



М. Л. Цымблер

1. Цели и задачи дисциплины

Целью дисциплины является ознакомление магистрантов с основными методами и алгоритмами интеллектуального анализа данных.

Краткое содержание дисциплины

Введение в дисциплину. Феномен Больших данных. Понятие интеллектуального анализа данных. Технологический цикл анализа данных. Основные задачи интеллектуального анализа данных: поиск шаблонов, классификация, кластеризация, поиск аномалий. Поиск шаблонов. Понятия транзакции, частого набора, шаблона, поддержки, достоверности. Основные алгоритмы поиска частых наборов: Apriori, Eclat, FP-Growth. Выбор полезных шаблонов на основе мер support, confidence, lift и др. Компактное представление частых наборов: максимально частые и замкнутые наборы, иерархии наборов. Фрагментация и сэмплинг для поиска частых наборов. Классификация. Процесс классификации: обучение модели, оценка модели, применение модели. Деревья решений. Меры оценки доли примесей в узле дерева решений: индекс Джини, энтропия; алгоритмы классификации ID3, C4.5, CART. Байесовская классификация. Классификация по ближайшим соседям. Оценка качества классификации: меры Accuracy, Precision, Recall, F1. Ансамблевая классификация: бэггинг, бустинг, случайный лес. Кластеризация. Задачи кластеризации данных и подходы к ее решению. Разделительная кластеризация: алгоритмы k-means, k-medoids и др. Иерархическая кластеризация: дендрограммы, агломеративный и дивизимный подход. Меры схожести кластеров: Single linkage, Complete linkage, Group average и др. Плотностная кластеризация: алгоритм DBSCAN. Нечеткая кластеризация: алгоритм Fuzzy C-Means. Меры качества кластеризации: критерий Хопкинса, кросс-валидация, метод локтя, силуэтный коэффициент и др. Поиск аномалий. Понятия аномалии (выброса), шума, новизны в данных. Виды аномалий: точечные, глобальные, контекстные, смешанные. Статистические методы поиска аномалий: z-значимость, правило трех сигм, гистограммы. Поиск аномалий на основе расстояния. Поиск аномалий на основе плотности: метод вложенных циклов, метод решеток. Поиск аномалий с помощью разделительной и плотностной кластеризации. Поиск аномалий на основе классификации: метод One Class SVM, метод изолирующего леса.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ОПК-7 Способен применять при решении профессиональных задач методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе, в глобальных компьютерных сетях	Знает: определения базовых задач интеллектуального анализа данных (поиск шаблонов, классификация, кластеризация и поиск аномалий) и принципы подготовки данных для выполнения их интеллектуального анализа Умеет: выполнять проектирование и реализацию приложений и подпрограмм, выполняющих подготовку данных для выполнения их интеллектуального анализа Имеет практический опыт: реализации

	приложений и подпрограмм, выполняющих подготовку данных для выполнения их интеллектуального анализа, с помощью современных языков и инструментов программирования
ПК-2 Способен разрабатывать и проводить экспериментальную проверку и тестирование программных систем по обеспечению требуемых критериев эффективности и качества функционирования	Знает: определения метрик оценки качества решений базовых задач интеллектуального анализа данных (поддержка и достоверность шаблона, точность и полнота классификации и др.) Умеет: выполнять проектирование и реализацию приложений и подпрограмм, выполняющих оценку и визуализацию качества решений базовых задач интеллектуального анализа данных Имеет практический опыт: реализации приложений и подпрограмм, выполняющих оценку и визуализацию качества решений базовых задач интеллектуального анализа данных, с помощью современных языков и инструментов программирования
ПК-3 Способен разрабатывать и применять методы и алгоритмы машинного обучения для решения задач	Знает: основные методы и алгоритмы решения базовых задач интеллектуального анализа данных (алгоритм Априори поиска шаблонов, классификация с помощью деревьев решений и ансамблей, кластеризация k-средних и др.) Умеет: выполнять проектирование и реализацию приложений и подпрограмм, выполняющих решение базовых задач интеллектуального анализа данных Имеет практический опыт: реализации приложений и подпрограмм, выполняющих решение базовых задач интеллектуального анализа данных, с помощью современных языков и инструментов программирования

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
1.О.09 Машинное обучение, 1.О.11 Глубокие нейронные сети, 1.О.08 Разработка систем искусственного интеллекта на языке Python	1.О.10 Технологии параллельного программирования

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

Дисциплина	Требования
1.О.11 Глубокие нейронные сети	Знает: математическую модель нейрона, технологии создания искусственных нейронных сетей, методы оптимизации, регуляризации и нормализации параметров нейронной сети и

	процесса ее обучения Умеет: осуществлять формализацию задачи, построение математической модели, подготовку обучающего набора данных, подбор топологии и создание искусственной нейронной сети в соответствии с поставленной задачей Имеет практический опыт: формулирования и решения задач в области машинного обучения с использованием нейросетевого подхода
1.О.08 Разработка систем искусственного интеллекта на языке Python	Знает: основные веб-фреймворки на Python, подходы многопоточного и асинхронного программирования, принципы разработки интеллектуальных систем на языке Python, синтаксис языка Python, основные библиотеки языка Python, применяемые для обработки данных , виды программных ошибок Умеет: применять конкретные специализированные фреймворки языка Python для сбора, обработки и анализа данных , подбирать наиболее подходящие фреймворки и библиотеки для разработки веб-сервисов сбора, анализа и обработки данных, разрабатывать программы на языке Python, в том числе с использованием современных интеллектуальных технологий, осуществлять тестирование программ Имеет практический опыт: анализа готовых информационных наборов данных, разработки и развертывания разработанного программного обеспечения для сбора и анализа данных , разработки программных сервисов сбора, анализа и обработки данных на Python, разработки программ на языке Python, поиска и устранения ошибок в программе
1.О.09 Машинное обучение	Знает: основы работы с различными источниками информации, включая базы данных, веб-сервисы и API (программируемые интерфейсы приложений), стратегические подходы к анализу сложных проблем, интегративные методы, заимствованные из социальной, экономической и других профессиональной сфер, классы методов и алгоритмов машинного обучения, функциональность современных инструментальных средств и систем программирования в области создания моделей и методов машинного обучения Умеет: эффективно использовать современные программные средства и инструменты для обработки информации, включая навыки работы с языками программирования для ИИ, а также знание библиотек и фреймворков, предназначенных для работы с данными и другие инструменты для манипуляций с данными, самостоятельно идентифицировать и формулировать нестандартные задачи, используя полученные математические и

	естественнонаучные знания для их решения, разрабатывать стратегические подходы к анализу сложных проблем, применяя интегративные методы, заимствованные из социальной, экономической и профессиональной сфер, включая умение адаптировать знания из одной дисциплины для эффективного применения в другой, что способствует генерации инновационных решений в междисциплинарном контексте, ставить задачи и разрабатывать новые методы и алгоритмы машинного обучения, самостоятельно разрабатывать и реализовывать модели машинного обучения, включая этапы предобработки данных, выбора наиболее эффективных методов, настройки гиперпараметров и оценки полученных результатов; уметь использовать инструменты для анализа и визуализации данных, чтобы обосновать выбранные подходы и улучшить понимание итоговых моделей Имеет практический опыт: извлечения, обработки и визуализации данных с использованием современных подходов и инструментов, проведения исследования и экспериментов, анализа полученных результатов, а также применения теоретических знаний для разработки решений, успешного применения методов количественного и качественного анализа, участия в проектах, решать основные классы задач методами и алгоритмами машинного обучения, использования систем программирования для создания и оптимизации алгоритмов машинного обучения, работы с языками программирования для ИИ
--	---

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 5 з.е., 180 ч., 76,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах
		Номер семестра
		2
Общая трудоёмкость дисциплины	180	180
<i>Аудиторные занятия:</i>	64	64
Лекции (Л)	32	32
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	32	32
Лабораторные работы (ЛР)	0	0
<i>Самостоятельная работа (СРС)</i>	103,5	103,5
Подготовка к контрольным опросам	103,5	103.5

Консультации и промежуточная аттестация	12,5	12,5
Вид контроля (зачет, диф.зачет, экзамен)	-	экзамен

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Введение в дисциплину	2	2	0	0
2	Поиск шаблонов	16	8	8	0
3	Классификация	16	8	8	0
4	Кластеризация	16	8	8	0
5	Поиск аномалий	14	6	8	0

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Феномен Больших данных. Понятие интеллектуального анализа данных. Технологический цикл анализа данных. Основные задачи анализа данных: поиск шаблонов, классификация, кластеризация, поиск аномалий.	2
2	2	Понятия транзакции, частого набора, шаблона, поддержки, достоверности. Основные алгоритмы поиска частых наборов: Apriori, Eclat, FP-Growth.	4
3	2	Выбор полезных шаблонов на основе мер support, confidence, lift и др. Компактное представление частых наборов: максимально частые и замкнутые наборы, иерархии наборов. Фрагментация и сэмплинг для поиска частых наборов	4
4	3	Процесс классификации: обучение модели, оценка модели, применение модели. Деревья решений. Меры оценки доли примесей в узле дерева решений: индекс Джини, энтропия; алгоритмы классификации ID3, C4.5, CART. Байесовская классификация. Классификация по ближайшим соседям. Оценка качества классификации: меры Accuracy, Precision, Recall, F1.	6
5	3	Ансамблевая классификация: бэггинг, бустинг, случайный лес.	2
6	4	Задачи кластеризации данных и подходы к ее решению. Разделительная кластеризация: алгоритмы k-means, k-medoids и др.	3
7	4	Иерархическая кластеризация: дендрограммы, агломеративный и дивизимный подход. Меры схожести кластеров: Single linkage, Complete linkage, Group average и др. Плотностная кластеризация: алгоритм DBSCAN. Нечеткая кластеризация: алгоритм Fuzzy C-Means.	3
8	4	Меры качества кластеризации: критерий Хопкинса, кросс-валидация, метод локтя, силуэтный коэффициент и др.	2
9	5	Понятия аномалии (выброса), шума, новизны в данных. Виды аномалий: точечные, глобальные, контекстные, смешанные. Статистические методы поиска аномалий: z-значимость, правило трех сигм, гистограммы. Поиск аномалий на основе расстояния. Поиск аномалий на основе плотности: метод вложенных циклов, метод решеток. Поиск аномалий с помощью разделительной и плотностной кластеризации. Поиск аномалий на основе классификации: метод One Class SVM, метод изолирующего леса.	6

5.2. Практические занятия, семинары

№	№	Наименование или краткое содержание практического занятия, семинара	Кол-
---	---	---	------

занятия	раздела		во часов
1	2	Поиск частых наборов с помощью алгоритмов Apriori, ECLAT и FP-Growth.	4
2	2	Поиск шаблонов с помощью мер support, confidence, lift.	4
3	3	Байесовская классификация. Классификация с помощью дерева решений.	4
4	3	Ансамблевая классификация с помощью бэггинга. Ансамблевая классификация с помощью случайного леса. Ансамблевая классификация с помощью бустинга.	4
5	4	Разделительная кластеризация с помощью алгоритмов k-Means и k-Medoids.	2
6	4	Плотностная кластеризация с помощью алгоритма DBSCAN.	2
7	4	Иерархическая кластеризация с помощью различных мер схожести кластеров.	2
8	4	Вычисление качества кластеризации	2
9	5	Поиск точечных аномалий с помощью z-значимости, правила трех сигм, гистограмм.	4
10	5	Поиск коллективных аномалий метода вложенных циклов, метода решеток, кластеризации.	4

5.3. Лабораторные работы

Не предусмотрены

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 контрольным опросам	Tan P.-N., Steinbach M., Karpatne A., Kumar V. Introduction to Data Mining. 2nd Edition. Pearson, 2019. 839 p.	2	103,5

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	2	Текущий контроль	Контрольный опрос по теме "Введение в дисциплину"	1	10	Контрольный опрос проводится в виде компьютерного теста по окончании изучения темы "Введение в дисциплину". Прохождение компьютерного теста оценивается от 0 до 10 баллов. Тест состоит из 10 равнозначных вопросов,	экзамен

						правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 15 минут.	
2	2	Текущий контроль	Практическое задание "Поиск шаблонов. Поиск частых наборов"	3,5	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со	экзамен

					следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.	
3	2	Текущий контроль	Практическое задание "Поиск шаблонов. Поиск ассоциативных правил"	3,5	10 Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета	экзамен

						– студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.	
4	2	Текущий контроль	Контрольный опрос по теме "Поиск шаблонов"	1	10	Контрольный опрос проводится в виде компьютерного теста по окончании изучения темы "Поиск шаблонов". Прохождение компьютерного теста оценивается от 0 до 10 баллов. Тест состоит из 10 равнозначных вопросов, правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 15 минут.	экзамен
5	2	Текущий контроль	Практическое задание "Классификация. Байесовская классификация"	3,5	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в	экзамен

					задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.	
6	2	Текущий контроль	Практическое задание "Классификация. Деревья решений"	3,5	10 Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация	экзамен

					полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.		
7	2	Текущий контроль	Практическое задание "Классификация. Бэггинг"	3,5	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными	экзамен

					значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.	
8	2	Текущий контроль	Практическое задание "Классификация. Случайный лес"	3,5	10 Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без	экзамен

					синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.		
9	2	Текущий контроль	Практическое задание "Классификация. Бустинг"	3,5	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики	экзамен

					кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.		
10	2	Текущий контроль	Контрольный опрос по теме "Классификация"	1	10	Контрольный опрос проводится в виде компьютерного теста по окончании изучения темы "Классификация". Прохождение компьютерного теста оценивается от 0 до 10 баллов. Тест состоит из 10 равнозначных вопросов, правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 15 минут.	экзамен
11	2	Текущий контроль	Практическое задание	3,5	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка	экзамен

			"Кластеризация. Разделительная кластеризация"		выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи — представленный студентом программный код четко соответствует поставленной задаче; — код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); — код компилируется без синтаксических ошибок и предупреждений компилятора; — при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов — представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения — студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов — студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета — студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации;	
--	--	--	---	--	--	--

						пояснения, раскрывающие смысл полученных результатов.	
12	2	Текущий контроль	Практическое задание "Кластеризация. Плотностная кластеризация"	3,5	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка	экзамен

						на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.	
13	2	Текущий контроль	Практическое задание "Кластеризация. Иерархическая кластеризация"	3,5	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания,	экзамен

						представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.	
14	2	Текущий контроль	Практическое задание "Кластеризация. Качество кластеризации"	3	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи — представленный студентом программный код четко соответствует поставленной задаче; — код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); — код компилируется без синтаксических ошибок и предупреждений компилятора; — при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов — представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения — студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов — студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их	экзамен

						визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.	
15	2	Текущий контроль	Контрольный опрос по теме "Кластеризация"	1	10	Контрольный опрос проводится в виде компьютерного теста по окончании изучения темы "Кластеризация". Прохождение компьютерного теста оценивается от 0 до 10 баллов. Тест состоит из 10 равнозначных вопросов, правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 15 минут.	экзамен
16	2	Текущий контроль	Практическое задание "Поиск аномалий. Поиск точечных аномалий"	3,5	10	Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. 2. Корректная визуализация полученных результатов – представленный студентом код	экзамен

					корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.	
17	2	Текущий контроль	Практическое задание "Поиск аномалий. Поиск коллективных аномалий"	3,5	10 Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении каждого из следующих пяти критериев (критерий оценивается от 0 до 2 баллов: 0 - не выполнен, 1 - выполнен частично, 2 - выполнен полностью): 1. Корректное решение поставленной задачи – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно); – код компилируется без синтаксических ошибок и предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные	экзамен

					результаты. 2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: формулировка задания; гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами; рисунки с результатами визуализации; пояснения, раскрывающие смысл полученных результатов.		
18	2	Текущий контроль	Контрольный опрос по теме "Поиск аномалий"	1	10	Контрольный опрос проводится в виде компьютерного теста по окончании изучения темы "Поиск аномалий". Прохождение компьютерного теста оценивается от 0 до 10 баллов. Тест состоит из 10 равнозначных вопросов, правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 15 минут.	экзамен
19	2	Промежуточная аттестация	Компьютерное тестирование	-	25	Промежуточная аттестация проводится во время экзамена в виде компьютерного теста. Тест состоит из 25 равноценных вопросов (под 5 вопросов на каждую из пяти тем курса), позволяющих оценить сформированность компетенций по курсу в целом, правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 45 мин.	экзамен

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
экзамен	При оценивании результатов учебной деятельности обучающегося по дисциплине используется балльно-рейтинговая система оценивания результатов учебной деятельности обучающихся (Положение о БРС утверждено приказом ректора от 24.05.2019 г. № 179, в редакции приказа ректора от 10.03.2022 г. № 25-13/09). Процедура прохождения промежуточной аттестации осуществляется согласно Положению о текущем контроле успеваемости и промежуточной аттестации (приказ ректора от 27.02.2024 № 33-13/09). Оценка за дисциплину формируется на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля следующим образом: • Отлично: Величина рейтинга обучающегося по дисциплине 85...100 %. • Хорошо: Величина рейтинга обучающегося по дисциплине 75...84 %. • Удовлетворительно: Величина рейтинга обучающегося по дисциплине 60...74 %. • Неудовлетворительно: Величина рейтинга обучающегося по дисциплине 0...59 %. Если студент согласен с оценкой, полученной по результатам текущего контроля, то он может в день, предшествующий промежуточной аттестации дать свое согласие на автомат в личном кабинете. В случае явки студента на промежуточную аттестацию, давшего свое согласие на автомат в личном кабинете, студент имеет право пройти мероприятия текущего контроля по дисциплине на промежуточной аттестации для улучшения своего рейтинга в день ее проведения. Снижение оценки в этом случае запрещено. Если студент не дал согласия в личном кабинете, то он может согласиться с оценкой лично на промежуточной аттестации в день ее проведения. Если студент не согласен с оценкой, то он имеет право пройти контрольно-рейтинговые мероприятия на промежуточной аттестации для улучшения своего рейтинга в день ее проведения. Фиксация результатов учебной деятельности по дисциплине проводится в день промежуточной аттестации на основе согласия студента, данного им в личном кабинете. При отсутствии согласия в журнале дисциплины фиксация результатов происходит при личном присутствии студента. Если студент не дал согласие в личном кабинете и не явился на промежуточную аттестацию – ему выставляется «неявка». Промежуточная аттестация проводится в форме тестирования. Тестирование проводится в системе edu.susu.ru. Тест содержит 25 вопросов. На выполнение теста дается 45 минут. В этом случае оценка за дисциплину рассчитывается на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля и промежуточной аттестации.	В соответствии с пп. 2.5, 2.6 Положения

6.3. Паспорт фонда оценочных средств

[illegible]

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. Миркин Б.Г. Введение в анализ данных. Учебник и практикум. Москва, 2020. 174 с.

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Миркин Б.Г. Введение в анализ данных. Учебник и практикум. Москва, 2020. 174 с.

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	eLIBRARY.RU	Алексеев Д.С., Щекочихин О.В. Технологии интеллектуального анализа данных. Учебное пособие. Кострома, 2020. 140 с. https://elibrary.ru/item.asp?id=43946965
2	Дополнительная литература	eLIBRARY.RU	Жаров А.Н., Минеичева И.Г. Анализ данных. Ярославль, 2020. 148 с. https://elibrary.ru/item.asp?id=43846458

Перечень используемого программного обеспечения:

Нет

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Практические занятия и семинары		Персональный компьютер
Лекции		Проектор