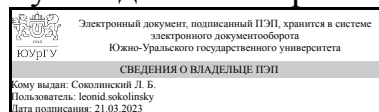


УТВЕРЖДАЮ:
Руководитель направления



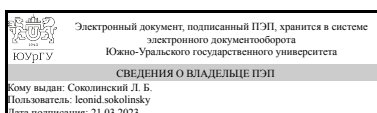
Л. Б. Соколинский

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**

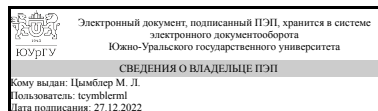
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 09.04.04 Программная инженерия, утверждённым приказом Минобрнауки от 19.09.2017 № 932

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-91 Способен использовать методы научных исследований и математического моделирования в области проектирования и управления системами искусственного интеллекта	Знает: приемы методологического обоснования научного исследования, методы организации библиотек искусственного интеллекта Умеет: проводить методологическое обоснование научного исследования, в том числе посредством создания и использования библиотек искусственного интеллекта Имеет практический опыт: проведения научного исследования по тематике искусственного интеллекта

ПК-3 Способен разрабатывать и применять методы и алгоритмы машинного обучения для решения задач	Знает: определение базовых задач машинного обучения (поиск шаблонов, классификация, кластеризация и поиск аномалий) и основные алгоритмы их решения Умеет: определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области
---	--

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
1.О.08 Разработка систем искусственного интеллекта на языке Python, 1.О.09 Машинное обучение, 1.О.02 Методология научного познания, 1.О.11 Глубокие нейронные сети	Не предусмотрены

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

Дисциплина	Требования
1.О.09 Машинное обучение	Знает: классы методов и алгоритмов машинного обучения, фундаментальные научные принципы и методы исследований, функциональность современных инструментальных средств и систем программирования в области создания моделей и методов машинного обучения Умеет: ставить задачи и разрабатывать новые методы и алгоритмы машинного обучения, адаптировать с целью практического применения фундаментальные и новые научные принципы и методы исследований Имеет практический опыт: решать основные классы задач методами и алгоритмами машинного обучения, применения классических методов исследования для решения профессиональных задач
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, принципы построения моделей глубоких нейронных сетей и глубокого машинного обучения (с подкреплением и без) Умеет: осуществлять формализацию задачи, построение математической модели, подготовку обучающего набора данных, подбор топологии и создание искусственной нейронной сети в соответствии с поставленной задачей, применять современные инструментальные средства и системы программирования для разработки и обучения моделей искусственных нейронных

	сетей Имеет практический опыт: формулирования и решения задач в области машинного обучения с использованием нейросетевого подхода
1.О.02 Методология научного познания	<p>Знает: технологии организации совместной работы, способы представления информации коллективу, методологию проведения исследовательского эксперимента, этапы проведения исследовательского эксперимента, логические методы и приемы научного исследования; методологические принципы современной науки, направления, концепции, источники знания и приемы работы с ними; основные особенности научного метода познания; программно-целевые методы решения научных проблем; основы моделирования управленческих решений; динамические оптимизационные модели; математические модели оптимального управления для непрерывных и дискретных процессов, их сравнительный анализ; многокритериальные методы принятия решений в профессиональной деятельности Умеет: разрабатывать аналитические обзоры состояния области прикладной информатики и информационных технологий по профильной направленности ООП магистратуры, планировать работу по проведению исследовательского эксперимента, строить план эксперимента, выделять факторы, влияющие на оценку результатов эксперимента, создавать условия повторяемости результатов эксперимента, применять логические методы и приемы научного исследования; методологические принципы современной науки, концепции, источники знания и приемы работы с ними; основные метода научного познания; программно-целевые методы решения научных проблем Имеет практический опыт: навыками изучения и релевантного поиска источников в заданной области, составления разнообразных аналитических отчетов, построения интеллектуальных карт предметной области; создания общих документов различных типов, репозитория для хранения данных и программ</p>
1.О.08 Разработка систем искусственного интеллекта на языке Python	<p>Знает: основные инструменты языка Python для сбора данных, необходимых для разработки программного обеспечение с применением алгоритмов машинного обучения, основные веб-фреймворки на Python, подходы многопоточного и асинхронного программирования, принципы промышленной разработки интеллектуальных систем на языке Python Умеет: подбирать наиболее подходящие инструменты сбора, анализа, обработки и визуализации данных в Python, применять конкретные специализированные фреймворки языка Python</p>

	для сбора, обработки и анализа данных для решения различных задач анализа данных, подбирать наиболее подходящие фреймворки и библиотеки для разработки веб-сервисов сбора, анализа и обработки данных Имеет практический опыт: сбора данных в различных форматах; предварительной обработки данных (приведение типов/форматов, заполнение пропусков фильтрация и т.п.); анализа и визуализации данных, анализа готовых информационных наборов данных; разработки и развертывания разработанного программного обеспечения для сбора и анализа данных в условиях решения реальных задач, разработки оригинальных программных сервисов сбора, анализа и обработки данных на Python
--	---

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

№ занятия	№ раздела	Наименование или краткое содержание практического занятия, семинара	Кол-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	Учи-тыва-ется 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, правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 15 минут.	экзамен
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	экзамен

					<p>(комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно);</p> <ul style="list-style-type: none"> – код компилируется без синтаксических ошибок и предупреждений компилятора; – при выполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты. <p>2. Корректная визуализация полученных результатов</p> <ul style="list-style-type: none"> – представленный студентом код корректно строит диаграммы/графики, указанные в задании. <p>3. Понимание разработанного решения</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. <p>4. Понимание полученных результатов</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. <p>5. Готовность отчета</p> <ul style="list-style-type: none"> – 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. 		
3	2	Текущий контроль	Практическое задание "Поиск шаблонов. Поиск ассоциативных правил"	3,5	10	<p>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):</p> <p>1. Корректное решение поставленной задачи</p> <ul style="list-style-type: none"> – представленный студентом программный код четко соответствует 	экзамен

					<p>поставленной задаче;</p> <ul style="list-style-type: none"> – 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. <p>2. Корректная визуализация полученных результатов</p> <ul style="list-style-type: none"> – представленный студентом код корректно строит диаграммы/графики, указанные в задании. <p>3. Понимание разработанного решения</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. <p>4. Понимание полученных результатов</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. <p>5. Готовность отчета</p> <ul style="list-style-type: none"> – 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	<p>Контрольный опрос проводится в виде компьютерного теста по окончании изучения темы "Поиск шаблонов".</p> <p>Прохождение компьютерного теста оценивается от 0 до 10 баллов. Тест состоит из 10 равнозначных вопросов, правильный ответ на один вопрос</p>	экзамен

						дает один балл. Время на прохождение теста – не менее 15 минут.	
5	2	Текущий контроль	Практическое задание "Классификация. Байесовская классификация"	3,5	10	<p>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):</p> <p>1. Корректное решение поставленной задачи</p> <ul style="list-style-type: none"> – 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. <p>2. Корректная визуализация полученных результатов</p> <ul style="list-style-type: none"> – представленный студентом код корректно строит диаграммы/графики, указанные в задании. <p>3. Понимание разработанного решения</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. <p>4. Понимание полученных результатов</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. <p>5. Готовность отчета</p> <ul style="list-style-type: none"> – студентом подготовлен отчет о выполнении задания, представляющий собой связный и структурированный документ со следующей информацией: 	экзамен

					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	<p>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):</p> <ol style="list-style-type: none"> 1. Корректное решение поставленной задачи <ul style="list-style-type: none"> – 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 <ul style="list-style-type: none"> – представленный студентом код корректно строит диаграммы/графики, указанные в задании. 3. Понимание разработанного решения <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. 4. Понимание полученных результатов <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. 5. Готовность отчета <ul style="list-style-type: none"> – студентом подготовлен отчет о 	экзамен

						<p>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.</p>	
7	2	Текущий контроль	<p>Практическое задание "Классификация. Бэггинг"</p>	3,5	10	<p>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):</p> <p>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.</p> <p>2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании.</p> <p>3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы.</p> <p>4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла</p>	экзамен

					<p>полученных результатов, включая их визуализацию.</p> <p>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.</p>	
8	2	Текущий контроль	<p>Практическое задание "Классификация. Случайный лес"</p>	3,5	<p>10</p> <p>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):</p> <p>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.</p> <p>2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании.</p> <p>3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы.</p> <p>4. Понимание полученных</p>	экзамен

					<p>результатов</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. <p>5. Готовность отчета</p> <ul style="list-style-type: none"> – 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	<p>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):</p> <p>1. Корректное решение поставленной задачи</p> <ul style="list-style-type: none"> – 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. <p>2. Корректная визуализация полученных результатов</p> <ul style="list-style-type: none"> – представленный студентом код корректно строит диаграммы/графики, указанные в задании. <p>3. Понимание разработанного решения</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко 	экзамен

					<p>ответить на контрольные вопросы преподавателя, касающиеся разработанной программы.</p> <p>4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию.</p> <p>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.</p>		
10	2	Текущий контроль	Контрольный опрос по теме "Классификация"	1	10	<p>Контрольный опрос проводится в виде компьютерного теста по окончании изучения темы "Классификация".</p> <p>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.</p>	экзамен
11	2	Текущий контроль	Практическое задание "Кластеризация. Разделительная кластеризация"	3,5	10	<p>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):</p> <p>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</p>	экзамен

					<p>предупреждений компилятора; – при исполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты.</p> <p>2. Корректная визуализация полученных результатов – представленный студентом код корректно строит диаграммы/графики, указанные в задании.</p> <p>3. Понимание разработанного решения – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы.</p> <p>4. Понимание полученных результатов – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию.</p> <p>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.</p>		
12	2	Текущий контроль	Практическое задание "Кластеризация. Плотностная кластеризация"	3,5	10	<p>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):</p> <p>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</p>	экзамен

					<p>входных и выходных параметров соответственно);</p> <ul style="list-style-type: none"> – 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. <p>2. Корректная визуализация полученных результатов</p> <ul style="list-style-type: none"> – представленный студентом код корректно строит диаграммы/графики, указанные в задании. <p>3. Понимание разработанного решения</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. <p>4. Понимание полученных результатов</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. <p>5. Готовность отчета</p> <ul style="list-style-type: none"> – 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. 		
13	2	Текущий контроль	Практическое задание "Кластеризация. Иерархическая кластеризация"	3,5	10	<p>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):</p> <p>1. Корректное решение поставленной задачи</p> <ul style="list-style-type: none"> – представленный студентом программный код четко соответствует поставленной задаче; – код документирован: файлы и 	экзамен

					<p>подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно);</p> <ul style="list-style-type: none"> – 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. <p>2. Корректная визуализация полученных результатов</p> <ul style="list-style-type: none"> – представленный студентом код корректно строит диаграммы/графики, указанные в задании. <p>3. Понимание разработанного решения</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. <p>4. Понимание полученных результатов</p> <ul style="list-style-type: none"> – студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. <p>5. Готовность отчета</p> <ul style="list-style-type: none"> – 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. 		
14	2	Текущий контроль	Практическое задание "Кластеризация. Качество кластеризации"	3	10	<p>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):</p> <ol style="list-style-type: none"> 1. Корректное решение поставленной задачи 	экзамен

					<p>– представленный студентом программный код четко соответствует поставленной задаче;</p> <p>– код документирован: файлы и подпрограммы исходного кода содержат спецификации (комментарии с описанием семантики кода в файле и описание семантики входных и выходных параметров соответственно);</p> <p>– код компилируется без синтаксических ошибок и предупреждений компилятора;</p> <p>– при выполнении кода на заданном наборе данных и с различными значениями параметров, указанными в задании, выдаются корректные результаты.</p> <p>2. Корректная визуализация полученных результатов</p> <p>– представленный студентом код корректно строит диаграммы/графики, указанные в задании.</p> <p>3. Понимание разработанного решения</p> <p>– студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы.</p> <p>4. Понимание полученных результатов</p> <p>– студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию.</p> <p>5. Готовность отчета</p> <p>– 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.</p>		
15	2	Текущий контроль	Контрольный опрос по теме "Кластеризация"	1	10	<p>Контрольный опрос проводится в виде компьютерного теста по окончании изучения темы "Кластеризация".</p> <p>Прохождение компьютерного теста оценивается от 0 до 10 баллов. Тест</p>	экзамен

						состоит из 10 равнозначных вопросов, правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 15 минут.	
16	2	Текущий контроль	Практическое задание "Поиск аномалий. Поиск точечных аномалий"	3,5	10	<p>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):</p> <ol style="list-style-type: none"> Корректное решение поставленной задачи <ul style="list-style-type: none"> 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. Корректная визуализация полученных результатов <ul style="list-style-type: none"> представленный студентом код корректно строит диаграммы/графики, указанные в задании. Понимание разработанного решения <ul style="list-style-type: none"> студент в состоянии быстро и четко ответить на контрольные вопросы преподавателя, касающиеся разработанной программы. Понимание полученных результатов <ul style="list-style-type: none"> студент в состоянии быстро и четко ответить на вопросы преподавателя, касающиеся содержательного смысла полученных результатов, включая их визуализацию. Готовность отчета <ul style="list-style-type: none"> студентом подготовлен отчет о выполнении задания, представляющий собой связный и 	экзамен

					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 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.	экзамен

						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	<p>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). Оценка за дисциплину формируется на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля. Отлично: Величина рейтинга обучающегося по дисциплине 85...100 %.</p> <p>Хорошо: Величина рейтинга обучающегося по дисциплине 75...84 %. Удовлетворительно: Величина рейтинга обучающегося по дисциплине 60...74 %. Неудовлетворительно: Величина рейтинга обучающегося по дисциплине 0...59 %.</p> <p>Если студент не согласен с оценкой, полученной по</p>	В соответствии с пп. 2.5, 2.6 Положения

	результатам текущего контроля, студент проходит мероприятие промежуточной аттестации в виде тестирования. Тестирование проводится в системе edu.susu.ru. Тест состоит из 25 равноценных вопросов, позволяющих оценить сформированность компетенций. На ответы отводится не менее 45 мин. В этом случае оценка за дисциплину рассчитывается на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля и промежуточной аттестации. Фиксация результатов учебной деятельности по дисциплине проводится в день экзамена при личном присутствии студента.	
--	--	--

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ																		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ОПК-91	Знает: приемы методологического обоснования научного исследования, методы организации библиотек искусственного интеллекта	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ОПК-91	Умеет: проводить методологическое обоснование научного исследования, в том числе посредством создания и использования библиотек искусственного интеллекта	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ОПК-91	Имеет практический опыт: проведения научного исследования по тематике искусственного интеллекта	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ПК-3	Знает: определение базовых задач машинного обучения (поиск шаблонов, классификация, кластеризация и поиск аномалий) и основные алгоритмы их решения	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ПК-3	Умеет: определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

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. 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
3	Методические пособия для самостоятельной работы студента	Образовательная платформа Юрайт	Миркин Б.Г. Введение в анализ данных. Учебник и практикум. Москва, 2020. 174 с. https://urait.ru/bcode/432851
4	Дополнительная литература	Электронно-библиотечная система издательства Лань	Новикова О.А., Андрианова Е.Г. Анализ данных. М.: МИРЭА, 2020. 162 с. https://e.lanbook.com/book/167597
5	Дополнительная литература	Электронно-библиотечная система издательства Лань	Макшанов А.В., Журавлев А.Е. Технологии интеллектуального анализа данных. М.: Лань, 2019. 212 с. ISBN 978-5-8114-4493-9 https://e.lanbook.com/book/120063
6	Основная литература	Электронно-библиотечная система издательства Лань	Маккинни У. Python и анализ данных. М.: ДМК-Пресс, 2020. 540 с. https://e.lanbook.com/book/131721

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
Лекции		Проектор
Практические занятия и семинары		Персональный компьютер

