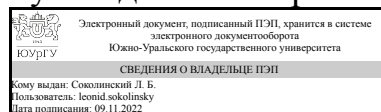


УТВЕРЖДАЮ:
Руководитель направления



Л. Б. Соколинский

РАБОЧАЯ ПРОГРАММА

дисциплины 1.О.20 Интеллектуальный анализ больших данных
для направления 02.04.02 Фундаментальная информатика и информационные
технологии

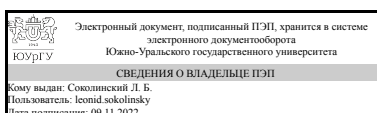
уровень Магистратура

форма обучения очная

кафедра-разработчик Системное программирование

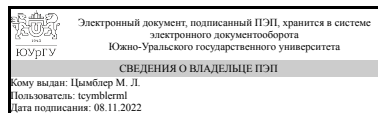
Рабочая программа составлена в соответствии с ФГОС ВО по направлению
подготовки 02.04.02 Фундаментальная информатика и информационные технологии,
утверждённым приказом Минобрнауки от 23.08.2017 № 811

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-2 Способен применять компьютерные/суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности	Знает: современные методы проектирования, разработки, отладки и тестирования приложений интеллектуального анализа данных Умеет: применять современные инструментальные средства для разработки приложений интеллектуального анализа данных Имеет практический опыт: применения современного программного инструментария для разработки приложений интеллектуального анализа данных

<p>ОПК-3 Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования</p>	<p>Знает: методы подготовки данных и оценки эффективности моделей интеллектуального анализа данных Умеет: применять методы подготовки данных и оценки эффективности аналитических моделей для разработки приложений интеллектуального анализа данных Имеет практический опыт: применения программных средств для подготовки данных и оценки эффективности моделей интеллектуального анализа данных</p>
<p>ПК-2 Способен разрабатывать системы хранения и обработки больших данных, в том числе на основе методов искусственного интеллекта</p>	<p>Знает: определения, технологический цикл и основные методы решения базовых задач интеллектуального анализа данных (поиск шаблонов, классификация, кластеризация, поиск аномалий) Умеет: выполнять проектирование приложений интеллектуального анализа данных Имеет практический опыт: разработки приложений интеллектуального анализа данных</p>

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
<p>1.О.04 Языки разметки, 1.О.18 Основы машинного обучения, ФД.01 Технологии интернета вещей, 1.О.03 Криптография и защита информации, 1.О.13 Язык Python для анализа данных, 1.О.10 Нейронные сети, ФД.02 Методы искусственного интеллекта, 1.О.14 Программирование корпоративных информационных систем на языке Java, 1.О.21 Системы управления предприятием</p>	<p>1.О.11 Технологии параллельного программирования, 1.О.07 Современные методы DevOps, 1.О.08 Анализ информационных технологий, 1.Ф.03 Анализ и прогнозирование временных рядов методами искусственного интеллекта, 1.О.17 Квантовые вычисления</p>

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

Дисциплина	Требования
<p>1.О.13 Язык Python для анализа данных</p>	<p>Знает: программное обеспечение для решения задач анализа данных, основные инструменты (программные библиотеки и язык программирования) для обработки и анализа данных, основные инструменты (программные библиотеки и язык программирования) для выполнения операций обработки и анализа данных Умеет: устанавливать программное обеспечение (среды разработок, программные библиотеки, соответствующий backend), просматривать версию и состав используемого программного обеспечения, задавать требуемый backend для решения поставленной задачи,</p>

	<p>подбирать наиболее подходящие инструменты для анализа имеющихся данных и выявления закономерностей, применять специализированные библиотеки языка Python для сбора, обработки и анализа данных Имеет практический опыт: установки и инсталляции программного обеспечения, используемого для решения задач в области сбора, обработки и анализа данных, анализа готовых информационных наборов данных, сбора данных в различных форматах (csv, json, xml), предварительной подготовки данных (приведение типов/форматов, заполнение пропусков фильтрация, и т.п.); анализа и визуализации данных</p>
<p>1.О.21 Системы управления предприятием</p>	<p>Знает: функциональные возможности систем класса ERP, среду разработки системы SAP ERP, методологию разработки, внедрения и эксплуатации систем управления предприятием, классификацию современных систем управления предприятием, задачи, решаемые с помощью систем управления предприятием Умеет: применять полученные знания для решения задач профессиональной деятельности с помощью систем управления предприятием, применять методологию разработки, внедрения и эксплуатации систем управления предприятием, находить, формулировать и решать актуальные проблемы с помощью систем управления предприятием Имеет практический опыт: решения задач профессиональной деятельности с помощью систем управления предприятием, управления проектами, связанными с разработкой, внедрением и эксплуатацией систем управления предприятием, решения актуальных проблем с помощью систем управления предприятием</p>
<p>1.О.10 Нейронные сети</p>	<p>Знает: технологии создания искусственных нейронных сетей с применением высокоуровневого языка программирования Python, математические основы и технологии создания и обучения искусственных нейронных сетей, математическую модель нейрона, основы линейной алгебры, технологии создания искусственных нейронных сетей, методы оптимизации, регуляризации и нормализации параметров нейронной сети и процесса ее обучения, методы оптимизации, регуляризации и нормализации параметров нейронной сети и процесса ее обучения Умеет: создавать и обучать глубокие и сверточные искусственные нейронные сети на Python, применять современные методы машинного обучения на основе нейронных сетей, осуществлять формализацию задачи, построение математической модели, подготовку обучающего</p>

	<p>набора данных, подбор топологии и создание искусственной нейронной сети в соответствии с поставленной задачей, создавать глубокие и сверточные искусственные нейронные сети с применением специализированных библиотек на разных вычислительных платформах (CPU/GPU/TPU) Имеет практический опыт: анализа и оптимизации полученных решений на основе нейросетевого подхода, формулирования и решения задач в области машинного обучения с использованием нейросетевого подхода, решения задач распознавания образов на разных вычислительных платформах (CPU/GPU/TPU)</p>
<p>1.О.04 Языки разметки</p>	<p>Знает: основы языков XSL, XSD и DTD, основные направления применения стандарта XML в управлении IT-проектами, корпоративными информационными системами и высоконагруженными Web-системами, способы разработки языков разметки на основе XML Умеет: создавать спецификацию XML-документа с помощью языков XSD и DTD, преобразовывать XML-документ в HTML с помощью XSL шаблона, осуществлять импорт-экспорт данных для XML-формата, на основе анализа исходных данных формировать XML-документ Имеет практический опыт: владения навыками по валидации и отладке XSD, DTD и XSL-документов, владения инструментами импорта-экспорта данных для XML-формата, владения методами валидации и отладки XML-документа</p>
<p>ФД.01 Технологии интернета вещей</p>	<p>Знает: принципы организации киберфизических систем, существующие технологии в интернете вещей, отечественные и зарубежные достижения в области программно-аппаратных комплексов интернета вещей Умеет: анализировать существующие IoT-технологии и применять их в конкретных условиях, определять сервисы, функции и выбирать технологии их реализации при разработке киберфизических программно-аппаратных компонентов Имеет практический опыт: владения специальной терминологией, навыками программирования конечных устройств, навыками разработки моделей и алгоритмов для взаимодействия с программными и аппаратными компонентами, самостоятельного проектирования и реализации компонентов интернета вещей</p>
<p>1.О.03 Криптография и защита информации</p>	<p>Знает: основные подходы к математической формализации различных аспектов безопасности информационных систем и реализации средств защиты информации, основные требования информационной безопасности, основные алгоритмы шифрования данных, базовые понятия для математического обеспечения информационной безопасности Умеет:</p>

	<p>применять математические методы и алгоритмы защиты информации при решении профессиональных задач в области информационной безопасности, применять математические методы защиты информации, кодировать информацию с помощью основных алгоритмов шифрования Имеет практический опыт: самостоятельного формулирования задач и политик безопасности, построения систем защиты, использования основных алгоритмов шифрования для защиты данных и информационной безопасности</p>
<p>1.О.14 Программирование корпоративных информационных систем на языке Java</p>	<p>Знает: основы объектно-ориентированного языка, основные понятия, виды и характеристики современного программного обеспечения технологии Java, подходы к тестированию программ на Java, основные понятия, виды и характеристики современного программного обеспечения технологии Java Умеет: применять подходы объектно-ориентированного программирования при разработке программного обеспечения, проектировать и разрабатывать локальные приложения на языке Java, разрабатывать документацию с помощью Javadoc, использовать специализированные среды разработки Java Имеет практический опыт: проектирования классов, ООП-архитектуры, создания программных проектов в специализированных средах разработки Java, разработки тестов для веб-сайта с помощью библиотеки Selenium, создания программных проектов в специализированных средах разработки Java</p>
<p>1.О.18 Основы машинного обучения</p>	<p>Знает: математические основы, принципы создания, обучения и валидации моделей машинного обучения, технологию создания моделей машинного обучения с помощью библиотек языка Python, методы оптимизации, регуляризации, нормализации и валидации моделей машинного обучения Умеет: применять современные методы машинного обучения, создавать и обучать модели машинного обучения с помощью библиотек языка Python Имеет практический опыт: анализа и оптимизации полученных решений на основе машинного обучения, решения задач машинного обучения с помощью библиотек языка Python</p>
<p>ФД.02 Методы искусственного интеллекта</p>	<p>Знает: математические основы и технологии машинного обучения, современные интегрированные среды разработки ПО на языках высокого уровня и специализированные библиотеки искусственного интеллекта Умеет: применять современные методы машинного обучения на основе нейронных сетей, создавать и обучать глубокие и сверточные искусственные нейронные сети с применением</p>

	специализированных библиотек Имеет практический опыт: анализа и оптимизации полученных решений на основе нейросетевого подхода, решения задач в области машинного обучения и компьютерного зрения
--	---

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 4 з.е., 144 ч., 82,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		2	
Общая трудоёмкость дисциплины	144	144	
<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>	61,5	61,5	
Индивидуальное задание	61,5	61,5	
Консультации и промежуточная аттестация	18,5	18,5	
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	Семестр	Кол-во

	ресурс		часов
Индивидуальное задание	Tan P.-N., Steinbach M., Karpatne A., Kumar V. Introduction to Data Mining. 2nd Edition. Pearson, 2019. 839 p. Chapter 10. Avoiding False Discoveries, p. 750-808.	2	61,5

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, 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 (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	экзамен

					<p>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>		
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 – 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);</p>	экзамен

					<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>		
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 дает один балл. Время на прохождение теста – не менее 15 минут.</p>	экзамен
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</p>	экзамен

					<p>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: 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	Практическое	3,5	10	Выполнение задания оценивается от 0	экзамен

		контроль	задание "Классификация. Деревья решений"		<p>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 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	<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 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.		
8	2	Текущий контроль	Практическое задание "Классификация. Случайный лес"	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>	
9	2	Текущий контроль	<p>Практическое задание "Классификация. Бустинг"</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> <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 	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>	
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 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</p>	экзамен

					<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>		
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</p> <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 исполнении кода на заданном наборе данных и с различными значениями параметров, указанными</p>	экзамен

					<p>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>		
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 – 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>		
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> <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. 		
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 состоит из 10 равнозначных вопросов, правильный ответ на один вопрос дает один балл. Время на прохождение теста – не менее 15 минут.</p>	экзамен
16	2	Текущий контроль	Практическое задание "Поиск аномалий. Поиск"	3,5	10	<p>Выполнение задания оценивается от 0 до 10 баллов. Максимальная оценка выставляется при полном выполнении</p>	экзамен

			точечных аномалий"		<p>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: 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	<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: формулировка задания; гиперссылка на каталог репозитория с исходными 	экзамен

						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 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.</p>	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	10	11	12	13	14	15	16	17	18	19
ОПК-2	Знает: современные методы проектирования, разработки, отладки и тестирования приложений интеллектуального анализа данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ОПК-2	Умеет: применять современные инструментальные средства для разработки приложений интеллектуального анализа данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ОПК-2	Имеет практический опыт: применения современного программного инструментария для разработки приложений интеллектуального анализа данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ОПК-3	Знает: методы подготовки данных и оценки эффективности моделей интеллектуального анализа данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ОПК-3	Умеет: применять методы подготовки данных и оценки эффективности аналитических моделей для разработки приложений интеллектуального анализа данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ОПК-3	Имеет практический опыт: применения программных средств для подготовки данных и оценки эффективности моделей интеллектуального анализа данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ПК-2	Знает: определения, технологический цикл и основные методы решения базовых задач интеллектуального анализа данных (поиск шаблонов, классификация, кластеризация, поиск аномалий)	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ПК-2	Умеет: выполнять проектирование приложений интеллектуального анализа данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ПК-2	Имеет практический опыт: разработки приложений интеллектуального анализа 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. 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

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