ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Заведующий выпускающей кафедрой

ЮУрГУ
Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Соколинский Л. Б.
Пользователь: leonid.sokolinsky
Дата подписания: 02.09.2024

Л. Б. Соколинский

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.М0.05.01 Технологии распределенной обработки данных
для направления 02.04.02 Фундаментальная информатика и информационные технологии
уровень Магистратура
магистерская программа Машинное обучение и анализ больших данных
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.

ЮУрГУ
Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Соколинский Л. Б.
Пользователь: leonid.sokolinsky
Дата подписания: 02.09.2024

Л. Б. Соколинский

Разработчик программы,
к.физ.-мат.н., доцент

ЮУрГУ
Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Иванова Е. В.
Пользователь: aksenovaev
Дата подписания: 29.08.2024

Е. В. Иванова

Челябинск

1. Цели и задачи дисциплины

Целью курса является изучение студентами задач, связанных с распределенным хранением и обработкой больших данных. При изучении этого курса должны быть решены следующие задачи: изучить понятие и проблематику больших данных, способы распределенного хранения больших данных, способы распределенной обработки больших данных, хранение и обработка больших данных с помощью современных программных решений, машинное обучение на больших данных.

Краткое содержание дисциплины

Понятие больших данных. Распределенная обработка больших данных. SQL, NoSQL и NewSQL-решения. Экосистема Hadoop: HDFS, MapReduce, Pig, Apache Hive, Apache Spark и машинное обучение, Hadoop YARN, Zookeeper, Apache Kafka. Классификация NoSQL-решений: хранилища "ключ-значения", документо-ориентированные хранилища, хранение в виде семейства столбцов, графовые СУБД. Теорема CAP. Согласованность данных в базе данных. Структуры для хранения больших данных. Секционирование данных. Репликация данных.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-1 Способен разрабатывать системы хранения и обработки больших данных, в том числе на основе методов искусственного интеллекта	Знает: основные положения и концепции в области хранения и обработки больших данных Умеет: анализировать типовые решения в области хранения и обработки больших данных, реализовывать техническое сопровождение информационных систем и баз данных, используемых для решения задач в области хранения и обработки больших данных, программировать системы хранения и обработки больших данных Имеет практический опыт: интеграции различных типов программного обеспечения в области хранения и обработки больших данных

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Методы и системы обработки больших данных, Интеллектуальный анализ данных, Глубокие нейронные сети, Анализ естественного языка методами искусственного интеллекта	Не предусмотрены

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

Дисциплина	Требования
Интеллектуальный анализ данных	Знает: определения, технологический цикл и основные методы решения базовых задач интеллектуального анализа данных (поиск шаблонов, классификация, кластеризация, поиск аномалий), современные методы проектирования, разработки, отладки и тестирования приложений интеллектуального анализа данных, методы подготовки данных и оценки эффективности моделей интеллектуального анализа данных Умеет: выполнять проектирование приложений интеллектуального анализа данных, применять современные инструментальные средства для разработки приложений интеллектуального анализа данных, применять методы подготовки данных и оценки эффективности аналитических моделей для разработки приложений интеллектуального анализа данных Имеет практический опыт: разработки приложений интеллектуального анализа данных, применения современного программного инструментария для разработки приложений интеллектуального анализа данных, применения программных средств для подготовки данных и оценки эффективности моделей интеллектуального анализа данных
Анализ естественного языка методами искусственного интеллекта	Знает: типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения для решения задач обработки естественного языка Умеет: применять типовые решения, библиотеки программных модулей, шаблоны, классы объектов при проектировании программного обеспечения Имеет практический опыт: проектирования и реализации приложений для решения задач обработки естественного языка с использованием методов машинного обучения и нейронных сетей
Методы и системы обработки больших данных	Знает: фундаментальные знания в области разработки систем управления большими данными Умеет: осуществлять первичный сбор и анализ материала в области разработки систем управления большими данными Имеет практический опыт: анализа и оптимизации найденных решений в области разработки систем управления большими данными
Глубокие нейронные сети	Знает: классы задач обработки больших данных на основе методов искусственных нейронных сетей, специализированные библиотеки для создания искусственных нейронных сетей, математическую модель нейрона, технологии создания искусственных нейронных сетей, методы оптимизации, регуляризации и нормализации параметров нейронной сети и

	процесса ее обучения Умеет: применять современные инструментальные средства и системы программирования для разработки и обучения моделей искусственных нейронных сетей , осуществлять формализацию задачи, построение математической модели, подготовку обучающего набора данных, подбор топологии и создание искусственной нейронной сети в соответствии с поставленной задачей Имеет практический опыт: создания и обучения искусственных нейронных сетей с применением специализированных библиотек, формулирования и решения задач в области машинного обучения с использованием нейросетевого подхода

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 3 з.е., 108 ч., 54,25 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах
		Номер семестра
		3
Общая трудоёмкость дисциплины	108	108
Аудиторные занятия:	48	48
Лекции (Л)	32	32
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	16	16
Лабораторные работы (ЛР)	0	0
Самостоятельная работа (СРС)	53,75	53,75
Изучение тем и проблем, не выносимых на лекции и практические занятия	40	40
Подготовка к зачету	13,75	13.75
Консультации и промежуточная аттестация	6,25	6,25
Вид контроля (зачет, диф.зачет, экзамен)	-	зачет

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Введение в большие данные. Экосистема Hadoop	30	18	12	0
2	NoSQL-решения	10	6	4	0
3	Распределенное хранение и обработка больших данных	8	8	0	0

5.1. Лекции

№	№	Наименование или краткое содержание лекционного занятия	Кол-

лекции	раздела		во часов
1	1	Введение в большие данные. Система хранения больших данных. Введение в распределенную обработку больших данных. SQL, NoSQL и NewSQL-решения.	2
2	1	Введение в платформу Hadoop. Экосистема Hadoop. Примеры систем на базе Hadoop.	2
3	1	Распределенная файловая система Hadoop (HDFS).	2
4	1	Технология MapReduce.	2
5	1	Введение в Pig и СУБД Apache Hive.	2
6	1	Apache Spark.	2
7	1	Разработка приложений для Apache Spark.	2
8	1	Координация распределенных сервисов с Zookeeper.	2
9	1	Брокер сообщений Apache Kafka.	2
10	2	Классификация NoSQL-систем. Хранилища "ключ-значения".	2
11	2	Документо-ориентированные хранилища.	2
12	2	Хранение в виде семейства столбцов. Введение в NoSQL-систему HBase на базе Hadoop. Графовые СУБД. Язык запросов Cypher. Фрейморк Pregel. Другие виды NoSQL-систем.	2
13	3	Распределенная обработка больших данных. Теорема CAP. Согласованность. Виды согласованности: строгая (Strong Consistency), конечная (Eventual Consistency), согласованное префиксное чтение (Consistent Prefix), с ограниченным устареванием (Bounded Staleness), монотонные чтения (Monotonic Reads, Session guarantee), чтение своих записей (Read My Writes). Структуры для хранения больших данных. Хеш-индексы. SS-таблицы. LSM-деревья. B-деревья.	2
14	3	Секционирование. Виды секционирования: по диапазонам значений ключа, по хешу ключа. Добавление/удаление секций, методы перебалансировки.	2
15	3	Репликация. Виды репликации: синхронная, асинхронная, полусинхронная репликация. Репликация с одним ведущим узлом. Добавление узлов в систему. Обработка сбоев узлов. Журнал репликации. Задержка репликации. Репликация с несколькими ведущими узлами.	2
16	3	Репликация без ведущего узла. Чтение и запись по кворуму. Обработка конкурентных записей. Векторы версий. Цепная репликация.	2

5.2. Практические занятия, семинары

№ занятия	№ раздела	Наименование или краткое содержание практического занятия, семинара	Кол-во часов
1	1	Установка платформы Hadoop. Работа с HDFS	3
2	1	Разработка MapReduce-приложения	3
3	1	Обработка данных с помощью Apache Hive	3
4	1	Анализ данных с помощью Apache Spark	3
5	2	Разработка базы данных с использованием NoSQL-решения	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
Изучение тем и проблем, не выносимых на лекции и практические занятия	[Осн. лит., 3], Гл. 8-9, с. 273–342; [Доп. лит., 5], Гл. 1-4, с. 5–47.	3	40
Подготовка к зачету	[Осн. лит., 1], Гл. 25, с. 458–466; [Осн. лит., 2], Гл. 4, с. 171–186; [Доп. лит., 4], Ч.3, гл. 13, с. 331–355.	3	13,75

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

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

6.1. Контрольные мероприятия (КМ)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учи-тыва-ется в ПА
1	3	Текущий контроль	Письменный опрос 1	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 1. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
2	3	Текущий контроль	Письменный опрос 2	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 2. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
3	3	Текущий контроль	Письменный опрос 3	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 3. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
4	3	Текущий	Письменный опрос	2	5	Письменный опрос проводится в виде	зачет

		контроль	4			электронного теста в конце лекции 4. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	
5	3	Текущий контроль	Письменный опрос 5	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 5. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
6	3	Текущий контроль	Письменный опрос 6	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 6. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
7	3	Текущий контроль	Письменный опрос 7	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 7. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
8	3	Текущий контроль	Письменный опрос 8	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 8. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
9	3	Текущий контроль	Письменный опрос 9	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 9. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за	зачет

						вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	
10	3	Текущий контроль	Письменный опрос 10	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 10. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
11	3	Текущий контроль	Письменный опрос 11	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 11. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
12	3	Текущий контроль	Письменный опрос 12	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 12. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
13	3	Текущий контроль	Письменный опрос 13	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 13. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
14	3	Текущий контроль	Письменный опрос 14	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 14. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
15	3	Текущий	Письменный опрос	2	5	Письменный опрос проводится в виде	зачет

		контроль	15			электронного теста в конце лекции 15. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	
16	3	Текущий контроль	Письменный опрос 16	2	5	Письменный опрос проводится в виде электронного теста в конце лекции 16. Тест содержит 5 случайных равноценных вопросов, за каждый из которых можно получить максимум 1 балл. Студент получает 1 балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. Время, отведенное на опрос, 10 минут.	зачет
17	3	Текущий контроль	Практическое задание 1. Установка платформы Hadoop. Работа с HDFS	12	5	5 баллов: задание выполнено полностью. Даны ответы на вопросы. 3 балла: выполнены только п.1 и п.2. Даны ответы на вопросы. 0 баллов: иначе	зачет
18	3	Текущий контроль	Практическое задание 2. Разработка MapReduce-приложения	14	5	5 баллов: задание выполнено полностью. Даны ответы на вопросы. 3 балла: выполнен только п.1. Даны ответы на вопросы. 0 баллов: иначе	зачет
19	3	Текущий контроль	Практическое задание 3. Обработка данных с помощью Apache Hive	14	5	5 баллов: задание выполнено полностью. Даны ответы на вопросы. 3 балла: выполнены только п.1 и п.2. Даны ответы на вопросы. 0 баллов: иначе	зачет
20	3	Текущий контроль	Практическое задание 4. Анализ данных с помощью Apache Spark	14	5	5 баллов: задание выполнено полностью. Даны ответы на вопросы. 3 балла: выполнены только п.1 и п.2. Даны ответы на вопросы. 0 баллов: иначе	зачет
21	3	Текущий контроль	Практическое задание 5. Разработка базы данных с использованием NoSQL-решения	14	5	5 баллов: задание выполнено полностью. Даны ответы на вопросы. 3 балла: выполнены только п.1 и п.2. Даны ответы на вопросы. 0 баллов: иначе	зачет
22	3	Проме-жуточная аттестация	Мероприятия промежуточной аттестации (компьютерное тестирование)	-	20	Промежуточная аттестация проводится в виде компьютерного тестирования. Тестирование проводится в системе edu.susu.ru. Тест состоит из 20 случайных равноценных вопросов, позволяющих оценить сформированность компетенций. За каждый вопрос студент может получить максимум 1 балл. Студент получает 1	зачет

						балл за вопрос, если ответ полностью верный, 0 баллов - иначе. Оценка студента за тест - это сумма баллов за каждый вопрос. На ответы отводится 45 минут.	

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
зачет	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. No 25-13/09). Процедура прохождения промежуточной аттестации осуществляется согласно Положению о текущем контроле успеваемости и промежуточной аттестации (приказ ректора от 27.02.2024 № 33-13/09). Оценка за дисциплину формируется на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля следующим образом: • Зачтено: Величина рейтинга обучающегося по дисциплине 60…100 %. • Незачтено: Величина рейтинга обучающегося по дисциплине 0…59 %. Если студент согласен с оценкой, полученной по результатам текущего контроля, то он может в день, предшествующий промежуточной аттестации дать свое согласие на автомат в личном кабинете. В случае явки студента на промежуточную аттестацию, давшего свое согласие на автомат в личном кабинете, студент имеет право пройти мероприятия текущего контроля по дисциплине на промежуточной аттестации для улучшения своего рейтинга в день ее проведения. Снижение оценки в этом случае запрещено. Если студент не дал согласия в личном кабинете, то он может согласиться с оценкой лично на промежуточной аттестации в день ее проведения. Если студент не согласен с оценкой, то он имеет право пройти контрольно-рейтинговые мероприятия на промежуточной аттестации для улучшения своего рейтинга в день ее проведения. Фиксация результатов учебной деятельности по дисциплине проводится в день промежуточной аттестации на основе согласия студента, данного им в личном кабинете. При отсутствии согласия в журнале дисциплины фиксация результатов происходит при личном присутствии студента. Если студент не дал согласие в личном кабинете и не явился на промежуточную аттестацию – ему выставляется «неявка». Промежуточная аттестация проводится в форме тестирования. Тестирование проводится в системе edu.susu.ru. Тест содержит 20 вопросов, на выполнение теста дается 45 минут. В этом случае оценка за дисциплину рассчитывается на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля и промежуточной аттестации.	В соответствии с пп. 2.5, 2.6 Положения

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ																					
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
ПК-1	Знает: основные положения и концепции в области хранения и обработки больших данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ПК-1	Умеет: анализировать типовые решения в области хранения и обработки больших данных, реализовывать техническое сопровождение информационных систем и баз данных, используемых для решения задач в области хранения и обработки больших данных, программировать системы хранения и обработки больших данных	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ПК-1	Имеет практический опыт: интеграции различных типов программного обеспечения в области хранения и обработки больших данных																		+	+	+		+

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) основная литература:

Не предусмотрена

б) дополнительная литература:

Не предусмотрена

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

Не предусмотрены

г) методические указания для студентов по освоению дисциплины:

1. Методические указания для студентов и преподавателей по освоению и организации самостоятельной работы студентов

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. Методические указания для студентов и преподавателей по освоению и организации самостоятельной работы студентов

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	Электронно-библиотечная	Осипов, Д. Л. Технологии проектирования баз данных / Д. Л. Осипов. — Москва : ДМК Пресс, 2019. — 498 с. — ISBN

		система издательства Лань	978-5-97060-737-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/131692 (дата обращения: 11.10.2021).
2	Основная литература	Электронно-библиотечная система издательства Лань	Методы и модели исследования сложных систем и обработки больших данных : монография / И. Ю. Парамонов, В. А. Смагин, Н. Е. Косых, А. Д. Хомоненко ; под редакцией В. А. Смагинаи А. Д. Хомоненко. — Санкт-Петербург : Лань, 2020. — 236 с. — ISBN 978-5-8114-4006-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/126938 (дата обращения: 11.10.2021).
3	Основная литература	Электронно-библиотечная система издательства Лань	Шарден, Б. Крупномасштабное машинное обучение вместе с Python : учебное пособие / Б. Шарден, Л. Массарон, А. Боскетти ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2018. — 358 с. — ISBN 978-5-97060-506-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/105836 (дата обращения: 11.10.2021).
4	Дополнительная литература	Электронно-библиотечная система издательства Лань	Григорьев, Ю. А. Реляционные базы данных и системы NoSQL : учебное пособие / Ю. А. Григорьев, А. Д. Плутенко, О. Ю. Плужникова. — Благовещенск : АмГУ, 2018. — 424 с. — ISBN 978-5-93493-308-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/156492 (дата обращения: 11.10.2021).
5	Дополнительная литература	Электронно-библиотечная система издательства Лань	Бутаков, Н. А. Обработка больших данных с Apache Spark : учебно-методическое пособие / Н. А. Бутаков, М. В. Петров, Д. Насонов. — Санкт-Петербург : НИУ ИТМО, 2019. — 50 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/136573 (дата обращения: 11.10.2021).

Перечень используемого программного обеспечения:

1. РСК Технологии-Система "Персональный виртуальный компьютер" (ПВК) (MS Windows, MS Office, открытое ПО)(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Практические занятия и семинары	110 (3г)	Компьютерный класс с доступом к сети Интернет
Зачет	110 (3г)	Компьютерный класс с доступом к сети Интернет

Лекции	110 (3г)	Мультимедийный проектор