

МИНОБРНАУКИ РОССИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Нижегородский государственный технический университет
им. Р.Е. Алексеева» (НГТУ)

Институт радиоэлектроники и информационных технологий (ИРИТ)
(Полное и сокращенное название института, реализующего данное направление)

УТВЕРЖДАЮ:

Директор института:

_____ Мякинников А.В.

подпись

ФИО

“ 22 ” _____ 04 _____ 2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.ОД.15 Анализ больших данных

(индекс и наименование дисциплины по учебному плану)

для подготовки бакалавров

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность: Интеллектуальные системы обработки информации и управления

Форма обучения: очная, заочная

Год начала подготовки 2025

Выпускающая кафедра ИСУ

Кафедра-разработчик ИСУ

Объем дисциплины 252/7
часов/з.е

Промежуточная аттестация зачёт с оценкой, экзамен

Разработчик: Багиров М.Б., старший преподаватель

Рабочая программа дисциплины: разработана в соответствии с Федеральным государственным образовательным стандартом высшего образования (ФГОС ВО 3++) по направлению подготовки 09.03.01 Информатика и вычислительная техника, утвержденного приказом МИНОБРНАУКИ РОССИИ от 19 сентября 2017 года № 929 на основании учебного плана принятого УМС НГТУ, протокол №17 от 28.05.2024 г., протокол №6 от 17.12.2024 г.

Рабочая программа одобрена на заседании кафедры разработчика протокол от 30.03.2025 №9

Зав. кафедрой к.т.н, доцент Тимофеева О.П. _____
(подпись)

Программа рекомендована к утверждению ученым советом института ИРИТ, протокол от 22.04.2025 №3

Рабочая программа зарегистрирована в УМУ, регистрационный № 09.03.01-и-39
Начальник МО _____ Е.Г. Севрюкова

Заведующая отделом комплектования НТБ _____ Н.И. Кабанина
(подпись)

Содержание

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	4
1.1 Цель освоения дисциплины.....	4
1.2 Задачи освоения дисциплины (модуля)	4
2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ.....	4
3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	5
4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ.....	7
4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ ПО СЕМЕСТРАМ	7
4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ.....	9
5. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	13
5.1 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности	19
5.2 ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ КОНТРОЛЯ УСПЕВАЕМОСТИ, ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ	20
6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	22
7. ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ	24
7.1 ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ	24
7.2 ПЕРЕЧЕНЬ СВОБОДНО РАСПРОСТРАНЯЕМОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ	24
7.3 ПЕРЕЧЕНЬ СОВРЕМЕННЫХ ПРОФЕССИОНАЛЬНЫХ БАЗ ДАННЫХ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ.....	25
8. ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОВЗ.....	25
9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ, НЕОБХОДИМОЕ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ.....	25
10. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ	27
10.1 ОБЩИЕ МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ, ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ.....	27
10.2 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ЗАНЯТИЙ ЛЕКЦИОННОГО ТИПА.....	28
10.3 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ НА ЛАБОРАТОРНЫХ РАБОТАХ.....	28
10.4 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ НА ПРАКТИЧЕСКИХ ЗАНЯТИЯХ	28
10.5 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ НА КУРСОВОЙ РАБОТЕ	29
10.6 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО САМОСТОЯТЕЛЬНОЙ РАБОТЕ ОБУЧАЮЩИХСЯ	29
11. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ	29
11.1 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта в ходе контроля текущей успеваемости.....	29
11.2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта в ходе промежуточной аттестации по дисциплине.....	29

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1 Цель освоения дисциплины

Целями изучения дисциплины являются формирование профессиональных компетенций в области анализа больших данных, включая современные методы сбора, обработки, хранения информации и построения ML-моделей, а также освоение полного цикла работы с данными: от предобработки и разметки до развертывания промышленных решений и оценки их эффективности.

1.2 Задачи освоения дисциплины (модуля)

Дисциплина «Анализ больших данных» способствует подготовке студентов к решению следующих профессиональных задач:

1. Организация полного цикла работы с данными: от проектирования систем сбора и хранения больших объемов структурированных и неструктурированных данных до создания и сопровождения датасетов для компьютерного зрения, обработки естественного языка и аудиоаналитики.
2. Разработка и оптимизация алгоритмов машинного обучения: построение и обучение моделей для работы с высокоразмерными и разреженными данными.
3. Контейнеризация и оркестрация ML-сервисов в промышленных средах.
4. Оценка и сопровождение работающих ML-систем.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Учебная дисциплина «Анализ больших данных» Б1.В.ОД.15 включена в обязательный перечень требованиями ФГОС, ОП ВО и УП, по данному направлению подготовки.

Дисциплина базируется на дисциплинах математического блока и блока программирования программы бакалавриата по направлению «Информатика и вычислительная техника»:

- Скриптовые языки программирования,
- Теория принятия решений,
- Интеллектуальный анализ данных.

Дисциплина «Анализ больших данных» является основополагающей для преддипломной практики и выполнения выпускной квалификационной работы.

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

Дисциплина «Анализ больших данных» формирует компетенции ПКС-3 совместно с дисциплинами и практиками, указанными в таблице 3.1.

Дисциплинарная часть компетенции ПКС-3 «Способен использовать формальные и интеллектуальные методы обработки и анализа информации при решении профессиональных задач»: способен понимать и применять на практике методы интеллектуального анализа и обработки больших данных.

Наименование дисциплин, формирующих компетенцию совместно	Семестры формирования дисциплины							
	1	2	3	4	5	6	7	8
<i>ПКС-3: Способен использовать формальные и интеллектуальные методы обработки и анализа информации при решении профессиональных задач</i>								
Математическое программирование								
Теория принятия решений								
Интеллектуальный анализ данных								
Математическое моделирование в АСО и У								
Анализ больших данных								
Численные методы в АСО и У								
Вычислительная математика								
Математическая логика и теория алгоритмов								
Теоретические основы алгоритмизации								
Теория графов и дискретная математика								
Дискретные структуры								
Машинное обучение								
Практика по получению профессиональных умений и опыта профессиональной деятельности								
Выполнение и защита ВКР								

Таблица 3.1 – Формирование компетенций дисциплинам

Таблица 3.2 – Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине			Оценочные средства	
					Текущего контроля	Промежуточной аттестации
ПКС-3. Способен использовать формальные и интеллектуальные методы обработки и анализа информации при решении профессиональных задач	ИПКС-3.2. Использует интеллектуальные методы обработки и анализа информации при решении профессиональных задач	Знать: современные методы обработки больших данных	Уметь: применять современные методы обработки больших данных для решения профессиональных задач	Владеть: - навыками использования современных методов обработки больших данных; - навыками оценки предложенных способов решения профессиональных задач с точки зрения соответствия цели проекта	Выполнение и сдача 6 лабораторных работ	Вопросы для зачета с оценкой – 20 билетов, вопросы для экзамена – 20 билетов

Освоение дисциплины причастно к ТФ А/03.6 (ПС 06.042 «Специалист по большим данным»), решает задачу анализа и обработки больших данных.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1 Распределение трудоёмкости дисциплины по видам работ по семестрам

Общая трудоёмкость дисциплины составляет 7 зач. ед. 252 часа, распределение часов по видам работ и семестрам представлено в таблице 4.1 и таблице 4.2.

Таблица 4.1 – Распределение трудоёмкости дисциплины по видам работ по семестрам для студентов очного обучения

Вид учебной работы	Трудоёмкость в час		
	Всего час.	В т.ч. по семестрам	
		7 сем	8 сем
Формат изучения дисциплины	с использованием элементов электронного обучения		
Общая трудоёмкость дисциплины по учебному плану	252	144	108
1. Контактная работа:	124	72	52
1.1 Аудиторная работа, в том числе:	118	68	50
занятия лекционного типа (Л)	64	34	30
занятия семинарского типа (ПЗ-семинары, практ. занятия и др)	-	-	-
лабораторные работы (ЛР)	54	34	20
1.2 Внеаудиторная, в том числе	6	4	2
курсовая работа (проект) (КР/КП) (консультация, защита)	-	-	-
текущий контроль, консультации по дисциплине	4	2	2
контактная работа на промежуточном контроле (КРА)	2	2	-
2. Самостоятельная работа (СРС)	92	36	56
реферат/эссе (подготовка)	-	-	-
расчётно-графическая работа (РГР) (подготовка)	-	-	-
контрольная работа	-	-	-
курсовая работа/проект (КР/КП) (подготовка)	-	-	-
самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиум и т.д.)	74	18	56
Подготовка к дифференцированному зачету	18	18	-
Подготовка к экзамену (контроль)	36	-	36

Таблица 4.2 – Распределение трудоёмкости дисциплины по видам работ по семестрам для студентов заочного обучения

Вид учебной работы	Трудоёмкость в час		
	Всего час.	В т.ч. по семестрам	
		8 сем	9 сем
Формат изучения дисциплины	с использованием элементов электронного обучения		
Общая трудоёмкость дисциплины по учебному плану	252	104	148
1. Контактная работа:	38	20	18
1.1 Аудиторная работа, в том числе:	32	16	16
занятия лекционного типа (Л)	12	6	6
занятия семинарского типа (ПЗ-семинары, практ. занятия и др)	-	-	-
лабораторные работы (ЛР)	20	10	10
1.2 Внеаудиторная, в том числе	6	4	2
курсовая работа (проект) (КР/КП) (консультация, защита)	-	-	-
текущий контроль, консультации по дисциплине	4	2	2

контактная работа на промежуточном контроле (КРА)	2	2	-
2. Самостоятельная работа (СРС)	201	84	117
реферат/эссе (подготовка)	-	-	-
расчётно-графическая работа (РГР) (подготовка)	-	-	-
контрольная работа	-	-	-
курсовая работа/проект (КР/КП) (подготовка)	-	-	-
самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к лабораторным и практическим занятиям, коллоквиум и т.д.)	201	84	117
Подготовка к дифференцированному зачету	-	-	-
Подготовка к экзамену (контроль)	13	4	9

4.2 Содержание дисциплины, структурированное по темам

Таблица 4.3 – Содержание дисциплины, структурированное по темам для студентов очного обучения

Планируемые (контролируемые) результаты освоения: код УК; ОПК; ПК и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы (час)					Вид СРС	Наименование используемых активных и интерактивных образовательных технологий	Реализация в рамках Практической подготовки (трудоемкость в часах)	Наименование разработанного Электронного курса (трудоемкость в часах)
		Контактная работа				Самостоятельная работа студента в (час)				
		Лекции (час)	Лабораторные работы (час)	Практические занятия (час)	КСР					
7 семестр										
Раздел 1. Основы работы с данными: сбор, хранение и разметка										
ПКС-3, ИПКС-3.2	Тема 1.1. Подготовка данных к обучению.	2								
	Тема 1.2. Форматы хранения табличных данных.	1				2	Подготовка к лекциям [6.1.6 – 6.1.8, 6.2.1]			
	Тема 1.3. Форматы хранения разметки для задач компьютерного зрения.	1	1		0,5					
	Тема 1.4. Инструменты для разметки данных.	1	2			2	Подготовка к лабораторной работе. [6.1.6 – 6.1.8, 6.2.1, 6.2.5]			
	Лабораторная работа. Сбор и разметка датасета для реализации задачи компьютерного зрения.	1,5	4							
	Итого по 1 разделу	6,5	7	-	0,5	4				
Раздел 2. Обучение моделей на больших данных										

ПКС-3, ИПКС-3.2	Тема 2.1. Обучение классических моделей на больших данных.	1				1	Подготовка к лабораторной работе. [6.1.1, 6.1.2, 6.2.1, 6.2.7]	Разбор конкретных ситуаций		
	Тема 2.2. Проблемы широких данных.	1						Разбор конкретных ситуаций		
	Тема 2.3. Снижение размерности.	2	2		0,5	2	Подготовка к лабораторной работе. [6.1.1, 6.1.2, 6.2.1, 6.2.7]	Разбор конкретных ситуаций		
	Тема 2.4. Хэширование признаков.	1				2	Подготовка к лабораторной работе. [6.1.1, 6.1.2, 6.2.1, 6.2.3]	Разбор конкретных ситуаций		
	Тема 2.5. Нативная работа с категориальными признаками.	0,5			0,5			Разбор конкретных ситуаций		
	Лабораторная работа. Сбор и разметка датасета для реализации задачи по автоматизированной обработке аудиопотока.	2	4							
	Итого по 2 разделу	7,5	6	-	1	5				
Раздел 3. Основы рекомендательных систем и коллаборативной фильтрации										
ПКС-3, ИПКС-3.2	Тема 3.1. Рекомендательные системы.	2				1	Подготовка к лекциям [6.1.4 – 6.1.6, 6.2.1, 6.2.4]	Разбор конкретных ситуаций		
	Тема 3.2. Коллаборативная фильтрация.	1						Разбор конкретных ситуаций		
	Тема 3.3. Похожесть товаров.	1						Разбор конкретных ситуаций		
	Тема 3.4. Матричные факторизации.	3	3		1,5	2	Подготовка к лабораторной работе. [6.1.1, 6.1.2, 6.2.1, 6.2.3]	Разбор конкретных ситуаций		
	Лабораторная работа. Обнаружение и распознавание объектов на видеопотоке.	2	6							

	Итого по 3 разделу	9	9	-	1,5	3				
Раздел 4. Обработка естественного языка (NLP)										
ПКС-3, ИПКС-3.2	Тема 4.1. Задачи автоматической обработки текста.	4	2		1	2	Подготовка к лабораторной работе. [6.1.2 – 6.1.6, 6.2.1, 6.2.2]	Разбор конкретных ситуаций		
	Тема 4.2. Векторные модели представления текста.	2	2			1	Подготовка к лабораторной работе. [6.1.3 – 6.1.5, 6.1.8, 6.2.1]	Разбор конкретных ситуаций		
	Тема 4.3. Дистрибутивные и контекстуализированные модели.	2				0,5	Подготовка к лекциям [6.1.2 – 6.1.4, 6.1.6, 6.2.1]			
	Лабораторная работа. Разработка ПО с графическим интерфейсом для автоматизированной обработки изображений с использованием алгоритмов компьютерного зрения.	3	8			2,5	Подготовка к лабораторной работе. [6.1.8 – 6.1.11, 6.2.1]			
	Итого по 4 разделу	11	12	-	1	6				
	Подготовка к дифференцируемому зачету	-	-	-	-	18				
8 семестр										
Раздел 5. Масштабирование нейронных сетей										
ПКС-3, ИПКС-3.2	Тема 5.1. Обучение нейронных сетей на больших данных.	2				3	Подготовка к лекциям [6.1.4 – 6.1.7, 6.2.1, 6.2.2]			
	Тема 5.2. Масштабирование обучения нейронных сетей.	4				4	Подготовка к лекциям [6.1.4 – 6.1.7, 6.2.1, 6.2.2]			
	Тема 5.3. Обучение на кластере.	2	1		0,5	3	Подготовка к лабораторной работе. [6.1.2 – 6.1.4, 6.2.1, 6.4.4]			

	Лабораторная работа. Сравнительная реализация метода обратного распространения ошибки и анализ его эффективности.		4			4	Подготовка к лабораторной работе. [6.1.2 – 6.1.4, 6.2.1, 6.4.4]			
	Итого по 5 разделу	8	5	-	0,5	14				
Раздел 6. Развертывание ML-систем										
ПКС-3, ИПКС-3.2	Тема 6.1. Проектирование и реализация систем машинного обучения.	2				2	Подготовка к лекциям [6.1.5 – 6.1.7, 6.2.1, 6.2.4]			
	Тема 6.2. Развертывание ML-сервиса.	2	1			2	Подготовка к лабораторной работе. [6.1.2 – 6.1.4, 6.1.6, 6.2.1]			
	Тема 6.3. Docker.	4	1		1	6	Подготовка к лабораторной работе. [6.1.2 – 6.1.4, 6.1.6, 6.2.1]			
	Итого по 6 разделу	8	2	-	1	10				
Раздел 7. Статистические методы оценки эффективности моделей										
ПКС-3, ИПКС-3.2	Тема 7.1. Аналитика.	2				1	Подготовка к лекциям [6.1.7 – 6.1.9, 6.2.1]			
	Тема 7.2. Разновидности метрик.	1				2	Подготовка к лекциям [6.1.4 – 6.1.9, 6.2.1]			
	Тема 7.3. Проверка гипотез.	1				2	Подготовка к лекциям [6.1.4 – 6.1.9, 6.2.1]			
	Тема 7.4. Статистические тесты.	1	2		0,5	1	Подготовка к лабораторной работе. [6.1.8 – 6.1.10, 6.1.6, 6.2.1, 6.2.5, 6.4.3, 6.4.6]			
	Тема 7.5. Онлайн-метрики.	2	2			2	Подготовка к лабораторной работе. [6.1.8 – 6.1.10, 6.1.6, 6.2.1, 6.2.5, 6.4.3, 6.4.6]			

	Лабораторная работа. Исследование сходимости статистических оценок для равномерного распределения на больших выборках.	1	5			4	Подготовка к лабораторной работе. [6.1.2 – 6.1.4, 6.2.1, 6.4.4]			
	Итого по 7 разделу	8	9	-	0,5	12				
Раздел 8. Процессы внедрения и обновления МЛ-систем										
ПКС-3, ИПКС-3.2	Тема 8.1. Процесс ввода изменений в эксплуатацию.	2				4	Подготовка к лекциям [6.1.4 – 6.1.9, 6.2.1]			
	Тема 8.2. Требования к вводу в эксплуатацию.	1	1			7	Подготовка к лабораторной работе. [6.1.5, 6.1.6, 6.2.1, 6.2.3]			
	Тема 8.3. Прием изменений и ввод в эксплуатацию, техническая сторона.	2	2			4	Подготовка к лабораторной работе. [6.1.5, 6.1.6, 6.2.1, 6.2.3]			
	Тема 8.4. Типы и стратегии внедрения обновлений.	1	1			5	Подготовка к лабораторной работе. [6.1.2 – 6.1.7, 6.2.1, 6.2.3]			
	Итого по 8 разделу	6	4	-	-	20				
	Подготовка к экзамену (контроль)	-	-	-	-	36				
	Итого	64	54	-	6	92				

Таблица 4.3 – Содержание дисциплины, структурированное по темам для студентов заочного обучения

Планируемые (контролируемые) результаты освоения: код УК; ОПК; ПК и индикаторы достижения компетенций	Наименование разделов, тем	Виды учебной работы (час)				КСП	Самостоятельная работа студента в (час)	Вид СРС	Наименование используемых активных и интерактивных образовательных технологий	Реализация в рамках Практической подготовки (трудоемкость в часах)	Наименование разработанного Электронного курса (трудоемкость в часах)
		Контактная работа									
		Лекции (час)	Лабораторные работы (час)	Практические занятия (час)							
8 семестр											

Раздел 1. Основы работы с данными: сбор, хранение и разметка										
ПКС-3, ИПКС-3.2	Тема 1.1. Подготовка данных к обучению.	0,5				6	Подготовка к лекциям [6.1.6 – 6.1.8, 6.2.1]			
	Тема 1.2. Форматы хранения табличных данных.	0,5	0,5							
	Тема 1.3. Форматы хранения разметки для задач компьютерного зрения.	0,5	0,5		0,5	4	Подготовка к лабораторной работе. [6.1.6 – 6.1.8, 6.2.1, 6.2.5]			
	Тема 1.4. Инструменты для разметки данных.	0,5	0,5		0,5	5	Подготовка к лабораторной работе. [6.1.6 – 6.1.8, 6.2.1, 6.2.5]			
	Лабораторная работа. Сбор и разметка датасета для реализации задачи компьютерного зрения.		1,5			6	Подготовка к лабораторной работе. [6.1.6 – 6.1.8, 6.2.1, 6.2.5]			
	Итого по 1 разделу	2	3	-	1	21				
Раздел 2. Обучение моделей на больших данных										
ПКС-3, ИПКС-3.2	Тема 2.1. Обучение классических моделей на больших данных.	0,5			0,5	3	Подготовка к лекциям. [6.1.1, 6.1.2, 6.2.1, 6.2.7]	Разбор конкретных ситуаций		
	Тема 2.2. Проблемы широких данных.		0,5			7	Подготовка к лабораторной работе. [6.1.1, 6.1.2, 6.2.1, 6.2.7]	Разбор конкретных ситуаций		
	Тема 2.3. Снижение размерности.	0,5			0,5	4	Подготовка к лекциям [6.1.1 – 6.1.3, 6.2.1, 6.2.3]	Разбор конкретных ситуаций		
	Тема 2.4. Хэширование признаков.	0,5				2	Подготовка к лекциям [6.1.1 – 6.1.3, 6.2.1, 6.2.3]	Разбор конкретных ситуаций		
	Тема 2.5. Нативная работа с категориальными признаками.	0,5				8	Подготовка к лекциям [6.1.1 – 6.1.4, 6.2.1, 6.2.5]	Разбор конкретных ситуаций		

	Лабораторная работа. Сбор и разметка датасета для реализации задачи по автоматизированной обработке аудиопотока.		1,5			9	Подготовка к лабораторной работе. [6.1.7, 6.1.9, 6.2.1, 6.2.2]			
	Итого по 2 разделу	2	2	-	1	33				
Раздел 3. Основы рекомендательных систем и коллаборативной фильтрации										
ПКС-3, ИПКС-3.2	Тема 3.1. Рекомендательные системы.	0,5						Разбор конкретных ситуаций		
	Тема 3.2. Коллаборативная фильтрация.					3	Подготовка к лекциям [6.1.4 – 6.1.6, 6.2.1, 6.2.4]	Разбор конкретных ситуаций		
	Тема 3.3. Похожесть товаров.					2	Подготовка к лекциям [6.1.1, 6.1.2, 6.2.1, 6.2.3]	Разбор конкретных ситуаций		
	Тема 3.4. Матричные факторизации.	0,5			1,5			Разбор конкретных ситуаций		
	Лабораторная работа. Обнаружение и распознавание объектов на видеопотоке.		1			8	Подготовка к лабораторной работе. [6.1.7, 6.1.8, 6.1.10, 6.2.1]			
	Итого по 3 разделу	1	1	-	1,5	13				
Раздел 4. Обработка естественного языка (NLP)										
ПКС-3, ИПКС-3.2	Тема 4.1. Задачи автоматической обработки текста.	0,5			0,5	1	Подготовка к лекциям [6.1.1 – 6.1.3, 6.2.1, 6.2.3]	Разбор конкретных ситуаций		
	Тема 4.2. Векторные модели представления текста.	0,5				4	Подготовка к лекциям [6.1.3 – 6.1.5, 6.1.8, 6.2.1]	Разбор конкретных ситуаций		
	Тема 4.3. Дистрибутивные и контекстуализированные модели.					3	Подготовка к лекциям [6.1.1 – 6.1.3, 6.2.1, 6.2.3]			

	Лабораторная работа. Разработка ПО с графическим интерфейсом для автоматизированной обработки изображений с использованием алгоритмов компьютерного зрения.		4			9	Подготовка к лабораторной работе. [6.1.8 – 6.1.11, 6.2.1]				
	Итого по 4 разделу	1	4	-	0,5	17					
	Подготовка к дифференцируемому зачету (контроль)	-	-	-	-	4					
9 семестр											
Раздел 5. Масштабирование нейронных сетей											
ПКС-3, ИПКС-3.2	Тема 5.1. Обучение нейронных сетей на больших данных.	0,5			0,5	9	Подготовка к лекциям [6.1.4 – 6.1.7, 6.2.1, 6.2.2]				
	Тема 5.2. Масштабирование обучения нейронных сетей.	0,5				4	Подготовка к лекциям [6.1.4 – 6.1.7, 6.2.1, 6.2.2]				
	Тема 5.3. Обучение на кластере.	0,5			0,5	5	Подготовка к лекциям [6.1.2 – 6.1.4, 6.2.1]				
	Лабораторная работа. Сравнительная реализация метода обратного распространения ошибки и анализ его эффективности.		2				17	Подготовка к лабораторной работе. [6.1.2 – 6.1.4, 6.2.1, 6.4.4]			
	Итого по 5 разделу	1,5	2	-	1	35					
Раздел 6. Развертывание ML-систем											
ПКС-3, ИПКС-3.2	Тема 6.1. Проектирование и реализация систем машинного обучения.	0,5				6	Подготовка к лекциям [6.1.5 – 6.1.7, 6.2.1, 6.2.4]				

	Тема 6.2. Развертывание сервиса. ML-	0,5				4	Подготовка к лекциям [6.1.2 – 6.1.4, 6.1.6, 6.2.1]				
	Тема 6.3. Docker.	1	1		0,5	13	Подготовка к лабораторной работе. [6.1.2 – 6.1.4, 6.1.6, 6.2.1]				
	Итого по 6 разделу	2	1	-	0,5	23					
Раздел 7. Статистические методы оценки эффективности моделей											
ПКС-3, ИПКС-3.2	Тема 7.1. Аналитика.	0,5				6	Подготовка к лекциям [6.1.7 – 6.1.9, 6.2.1]				
	Тема 7.2. Разновидности метрик.	0,5			0,5	7	Подготовка к лекциям [6.1.4 – 6.1.9, 6.2.1]				
	Тема 7.3. Проверка гипотез.					8	Подготовка к лекциям [6.1.4 – 6.1.9, 6.2.1]				
	Тема 7.4. Статистические тесты.	0,5				4	Подготовка к лекциям [6.1.8 – 6.1.10, 6.1.6, 6.2.1, 6.2.5, 6.4.3, 6.4.6]				
	Тема 7.5. Онлайн-метрики.		1			2	Подготовка к лабораторной работе. [6.1.8 – 6.1.10, 6.1.6, 6.2.1, 6.2.5, 6.4.3, 6.4.6]				
	Лабораторная работа. Исследование сходимости статистических оценок для равномерного распределения на больших выборках.		4				10	Подготовка к лабораторной работе. [6.1.2 – 6.1.4, 6.2.1, 6.4.4]			
	Итого по 7 разделу	1,5	5	-	0,5	37					
Раздел 8. Процессы внедрения и обновления ML-систем											
ПКС-3, ИПКС-3.2	Тема 8.1. Процесс ввода изменений в эксплуатацию.	0,5				4	Подготовка к лекциям [6.1.4 – 6.1.9, 6.2.1]				
	Тема 8.2. Требования к вводу в эксплуатацию.		1			6	Подготовка к лабораторной работе. [6.1.5, 6.1.6, 6.2.1, 6.2.3, 6.4.4]				

	Тема 8.3. Прием изменений и ввод в эксплуатацию, техническая сторона.	0,5				6	Подготовка к лекциям [6.1.5, 6.1.6, 6.2.1, 6.2.3]			
	Тема 8.4. Типы и стратегии внедрения обновлений.		1			6	Подготовка к лабораторной работе. [6.1.2 – 6.1.7, 6.2.1, 6.2.3, 6.4.4]			
	Итого по 8 разделу	1	2	-	-	22				
	Подготовка к экзамену (контроль)	-	-	-	-	9				
	Итого	12	20	-	6	201				

5. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ.

5.1 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности

1. Примерный перечень вопросов при защите лабораторных работ:

- Что такое анализ данных? Какие технологии используются при обработке данных?
- Какие инструменты разметки применяли и в чем их преимущества?
- Чем отличаются структурированные данные от неструктурированных?
- Какие матричные преобразования используются в машинном обучении?
- Какие основные задачи решает машинное обучение?
- Какие функции активации вам известны?
- Как вы обеспечивали репрезентативность собранного датасета?
- Опишите основные способы обучения нейронных сетей? В чем их отличия от алгоритмов машинного обучения?
- Назовите основные подходы для работы с изображениями?
- Какие задачи решает компьютерное зрение?
- В чем преимущества и недостатки вашей реализации?
- Какие проблемы возникли при сборе/разметке и как их решали?
- Какие методы применяли для контроля качества разметки?
- Что такое обработка естественного языка, назовите основные алгоритмы?
- Как решали проблему перекрытия объектов?
- Какие закономерности выявили в поведении оценок?
- Какие методы постобработки результатов применяли?
- Как сравнивали свою реализацию с библиотечными решениями?
- Реализация алгоритма обратного распространения ошибки.
- Методы ускорения сходимости.

2. Примерный перечень вопросов для зачета с оценкой:

- Основные понятия: большие данные, наука о данных, анализ данных, машинное обучение. Основные этапы извлечения знаний из данных. Примеры задач анализа данных.
- Инструменты для разметки данных: преимущества и недостатки.
- Обработка данных: описательные статистики, поиск аномалий (включая гистограммы, ящичковые диаграммы, ядерные оценки плотности), анализ выбросов и шумов. Нормализация и стандартизация данных.
- Сравнение форматов хранения табличных данных по производительности и особенностям использования.
- Матричные преобразования, векторы, функции регрессии.
- Машинное обучение: основные понятия, задачи, которые можно решить с помощью машинного обучения. Алгоритмы машинного обучения с учителем и без учителя (перечислить).
- Машинное обучение: этапы моделирования. Отбор признаков на основе модели. Оценка качества построенных моделей.
- Машинное обучение: ансамбли алгоритмов. Способы реализации ансамблей на языке python.
- Специфические проблемы при обучении классических моделей на больших данных.
- Проблема широких данных и подходы к ее решению.
- Метод хэширования признаков.

- Суть коллаборативной фильтрации, ее виды.
- Нейронные сети: базовые понятия, отличия нейросетей от классических методов машинного обучения.
- Рекомендательные системы. Основные подходы к построению рекомендательных систем. Примеры рекомендательных систем.
- Измерение схожести товаров в рекомендательных системах.
- Принцип работы матричных факторизаций в рекомендательных системах.
- Основные задачи автоматической обработки текста.
- Специфические проблемы обработки естественного языка на больших данных.
- Отличия между дистрибутивными и контекстуализированными моделями.
- Векторные модели представления текста.

3. Примерный перечень вопросов для экзамена:

- Понятие большие данные.
- Основные подходы к масштабированию обучения нейронных сетей для больших данных.
- «Жизненный цикл» проекта по аналитике больших данных.
- Корреляция. Регрессионный анализ.
- Примеры использования корреляции в области больших данных.
- Обработка текстов с помощью глубоких нейросетей.
- Специфические проблемы при обучении глубоких нейронных сетей на распределенных кластерах.
- Рекуррентные нейронные сети для обработки текстов.
- Одномерные сверточные нейронные сети для обработки текста.
- Сравнение эффективности синхронного и асинхронного SGD при распределенном обучении.
- Полносвязные и рекуррентные нейронные сети для прогнозирования временных рядов.
- Прямые и сверточные нейронные сети для обработки аудио сигналов.
- Стратегии развертывания ML-сервисов.
- Парадигма MapReduce. Ее реализация Hadoop.
- Проблема переобучения. Регуляризация.
- Организация A/B тестирования для ML-моделей в production.
- Интерпретация онлайн-метрик при оценке бизнес-эффективности модели.
- Требования к ML-моделям перед вводом в промышленную эксплуатацию.
- Методика проверки статистических гипотез при анализе качества ML-моделей.
- Принципы контейнеризации ML-приложений.

Комплект оценочных средств является неотъемлемой частью ФОС и хранится на кафедре «Информатика и системы управления».

5.2 Описание показателей и критериев контроля успеваемости, описание шкал оценивания

Для оценки знаний, умений, навыков и формирования компетенции по дисциплине применяется **традиционная** система, при которой успеваемость студентов оценивается по четырехбалльной шкале «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Таблица 5.1 – Критерии оценивания результата обучения по дисциплине и шкала оценивания

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Критерии оценивания результатов обучения			
		Оценка «неудовлетворительно» / «не зачтено» 0-59% от max рейтинговой оценки контроля	Оценка «удовлетворительно» / «зачтено» 60-74% от max рейтинговой оценки контроля	Оценка «хорошо» / «зачтено» 75-89% от max рейтинговой оценки контроля	Оценка «отлично» / «зачтено» 90-100% от max рейтинговой оценки контроля
ПКС-3. Способен использовать формальные и интеллектуальные методы обработки и анализа информации при решении профессиональных задач	ИПКС-3.2. Использует интеллектуальные методы обработки и анализа информации при решении профессиональных задач	Изложение учебного материала бессистемное, неполное, отсутствует понимание особенностей и методов работы с данными, не способен применять инструменты вычислений на Python	Имеет частичное понятие об основных особенностях и методах работы с наборами данных, частичное применение инструментов вычислений на Python при решении отдельных задач, имеет трудности в применении различных методов обучения нейронных сетей	Знает особенности и методы работы с наборами данных, применяет на практике инструменты вычислений на Python, владеет основами работы с инструментами распределенной обработки данных	Имеет глубокие системные знания в области интеллектуального анализа данных; применяет на практике инструменты вычислений на Python и методы обучения нейронных сетей при решении отдельных задач; успешно применяет инструменты распределенной обработки данных

Таблица 5.2 – Критерии оценивания

Оценка	Критерии оценивания
Высокий уровень «5» (отлично)	оценку « отлично » заслуживает студент, освоивший знания, умения, компетенции и теоретический материал без пробелов; выполнивший все задания, предусмотренные учебным планом на высоком качественном уровне; практические навыки профессионального применения освоенных знаний сформированы.
Средний уровень «4» (хорошо)	оценку « хорошо » заслуживает студент, практически полностью освоивший знания, умения, компетенции и теоретический материал, учебные задания не оценены максимальным числом баллов, в основном сформулировал практические навыки.
Пороговый уровень «3» (удовлетворительно)	оценку « удовлетворительно » заслуживает студент, частично с пробелами освоивший знания, умения, компетенции и теоретический материал, многие учебные задания либо не выполнил, либо они оценены числом баллов близким к минимальному, некоторые практические навыки не сформированы.
Минимальный уровень «2» (неудовлетворительно)	оценку « неудовлетворительно » заслуживает студент, не освоивший знания, умения, компетенции и теоретический материал, учебные задания не выполнил, практические навыки не сформированы.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1 Учебная литература

- 6.1.1. Алпатов, А. В. Работа с массивами и табличными данными на языке Python [Электронный ресурс]: учебное пособие / А. В. Алпатов; Министерство науки и высшего образования Российской Федерации, ВПИ (филиал) ФГБОУ ВО ВолгГТУ. — Волжский, 2024. — Режим доступа: <http://lib.volpi.ru>
- 6.1.2. Дружинин, Д. В. Высокопроизводительные вычисления и облачные технологии: учебное пособие / Д. В. Дружинин. — Томск: ТГУ, 2020. — 94 с. — URL: <https://e.lanbook.com/book/202355>
- 6.1.3. Волкова, С. С. Введение в анализ данных. Логические и метрические модели: Учебное пособие / С. С. Волкова. — Вологда: Вологодский государственный университет, 2023. — 79 с. — URL: <https://elibrary.ru/item.asp?id=54385367>
- 6.1.4. Донская, А. Р. Основы математического моделирования и анализа данных: учебное пособие / А. Р. Донская, А. В. Зубков, Ю. А. Орлова. — Волгоград: Волгоградский государственный технический университет, 2024. — 92 с. — URL: <https://elibrary.ru/item.asp?id=67132379>
- 6.1.5. Лимановская, О. В. Основы машинного обучения: учебное пособие для студентов вуза / О. В. Лимановская, Т. И. Алферьева; Министерство науки и высшего образования Российской Федерации, Уральский федеральный университет им. первого Президента России Б. Н. Ельцина. — Екатеринбург: Уральский федеральный университет имени первого Президента России Б.Н. Ельцина, 2020. — 88 с. — URL: https://elar.urfu.ru/bitstream/10995/88687/1/978-5-7996-3015-7_2020.pdf
- 6.1.6. Волкова, С. С. Введение в машинное обучение. Линейные модели: учебное пособие / С. С. Волкова. — Вологда: Вологод, 2023. — 76 с. — ISBN 978-5-907606-46-3. — URL: <https://elibrary.ru/item.asp?id=50732254>
- 6.1.7. Кугаевских, А. В. Классические методы машинного обучения / А. В. Кугаевских, Д. И. Муромцев, О. В. Кирсанова; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики. — Санкт-Петербург, 2022. — 53 с. — URL: <https://books.ifmo.ru/file/pdf/3075.pdf>
- 6.1.8. Ермагамбетов Р. Т., Киселев Е. С. Современные системы хранения и обработки больших данных: hadoop и apache spark // Форум молодых ученых. 2018. — №8

- (24). — URL: <https://cyberleninka.ru/article/n/sovremennye-sistemy-hraneniya-i-obrabotki-bolshih-dannyh-hadoop-i-apache-spark>
- 6.1.9. Некратюк А. А., Сафарьян О. А. Использование метода MapReduce в big data // Молодой исследователь Дона. 2020. — №3 (24). — URL: <https://cyberleninka.ru/article/n/ispolzovanie-metoda-mapreduce-v-big-data>
- 6.1.10. Emad Saddam Lake Data Warehouse Architecture for Big Data Solutions / Saddam Emad. — Climate Change Information Center and Renewable Energy and Expert System Agricultural Research Center (ARC) Giza, Egypt. — International Journal of Advanced Computer Science and Applications, Vol. 11, No. 8, 2020. — URL: https://thesai.org/Downloads/Volume11No8/Paper_54Lake_Data_Warehouse_Architecture.pdf
- 6.1.11. Lidong Wang Machine Learning in Big Data / Lidong Wang, Chery Ann Alexander. — Department of Engineering Technology, Mississippi Valley State University, USA, 2016. — 10 с. — URL: <https://ijmems.in/assets/ijmems-16-009-vol.-1,-no.-2,-52-61,-2016.pdf>

6.2 Справочно-библиографическая литература

- 6.2.1. Соколова, Э. С. Интеллектуальный анализ больших данных / Э. С. Соколова, М. Б. Багиров. — Нижегородский государственный технический университет им. Р.Е. Алексеева, 2024. — 129 с.
- 6.2.2. Ризаев, И. С. Интеллектуальный анализ данных: учебное пособие / И. С. Ризаев, Э. Г. Тахавова. — Казань: Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ, 2020. — 116 с. — URL: <https://elibrary.ru/item.asp?id=53761279>
- 6.2.3. Запечников, С. В. Основы интеллектуального анализа данных и машинного обучения: Конспект лекций: учебное пособие / С. В. Запечников. — Москва: НИЯУ МИФИ, 2022. — 136 с. — URL: <https://e.lanbook.com/book/355580>
- 6.2.4. Андреев, А. Е. Системы обработки больших данных: Учебно-методическое пособие / А. Е. Андреев, В. А. Егунов, П. Д. Кравченя. — Волгоград: Волгоградский государственный технический университет, 2022. — 96 с. — URL: <https://elibrary.ru/item.asp?id=49814723>
- 6.2.5. Колмогорова, С. С. Основы искусственного интеллекта: учебное пособие для студентов / С. С. Колмогорова. — Санкт-Петербург: СПбГЛТУ, 2022. — 108 с. — URL: <https://e.lanbook.com/book/257804>
- 6.2.6. Apache Spark - A Unified engine for large-scale data analytics, 2018. — Официальная документация. — URL: <https://spark.apache.org/docs/latest/#apache-spark-a-unified-engine-for-large-scale-data-analytics>
- 6.2.7. Apache Airflow, 2025. — Официальная документация. — URL: <https://airflow.apache.org/docs/>

6.3 Перечень журналов по профилю дисциплины:

- 6.3.1. «Искусственный интеллект и принятие решений» (<http://www.aidt.ru>)
- 6.3.2. «Мир больших данных (Big data)»

6.4 Методические указания, рекомендации и другие материалы к занятиям

- 6.4.1. Анализ больших данных: метод. рекомендации по подготовке лабораторной работы 1 по дисциплине «Анализ больших данных» для студентов высших учебных заведений направления 09.03.01 «Информатика и вычислительная техника» очной формы обучения / М.Б. Багиров, Э.С. Соколова, П.А. Шагалова; Нижегород. гос. техн. ун-т им. Р.Е. Алексеева. — Нижний Новгород, 2023. — 15 с.

- 6.4.2. Анализ больших данных: метод. рекомендации по подготовке лабораторной работы 2 по дисциплине «Анализ больших данных» для студентов высших учебных заведений направления 09.03.01 «Информатика и вычислительная техника» очной формы обучения / М.Б. Багиров, Э.С. Соколова, П.А. Шагалова; Нижегород. гос. техн. ун-т им. Р.Е. Алексеева. — Нижний Новгород, 2023. — 21 с.
- 6.4.3. Анализ больших данных: метод. рекомендации по подготовке лабораторной работы 3 по дисциплине «Анализ больших данных» для студентов высших учебных заведений направления 09.03.01 «Информатика и вычислительная техника» очной формы обучения / М.Б. Багиров, Э.С. Соколова; Нижегород. гос. техн. ун-т им. Р.Е. Алексеева. — Нижний Новгород, 2023. — 17 с.
- 6.4.4. Анализ больших данных: метод. рекомендации по подготовке лабораторной работы 4 по дисциплине «Анализ больших данных» для студентов высших учебных заведений направления 09.03.01 «Информатика и вычислительная техника» очной формы обучения / М.Б. Багиров, Э.С. Соколова; Нижегород. гос. техн. ун-т им. Р.Е. Алексеева. — Нижний Новгород, 2023. — 21 с.
- 6.4.5. Большие данные: Методические указания по контактной и самостоятельной работе. — Краснодар: Кубанский государственный аграрный университет имени И.Т. Трубилина, 2020. — 35 с. — URL: <https://elibrary.ru/item.asp?id=44227811>
- 6.4.6. Практический курс классического машинного обучения с использованием моделей математического программирования: учебно-методическое пособие / П. Ф. Чернавин, Н. П. Чернавин, Ф. П. Чернавин; М-во науки и высшего образования Екатеринбург: Изд-во Урал. ун-та, 2023 — 124 с. — URL: https://elar.urfu.ru/bitstream/10995/128482/1/978-5-7996-3767-5_2023.pdf

Электронные варианты всех методических указаний отправляются на электронные адреса групп.

7. ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Учебный процесс по дисциплине обеспечен необходимым комплектом свободно распространяемого программного обеспечения (состав по дисциплине определен в настоящей РПД и подлежит обновлению при необходимости).

7.1 Перечень информационных справочных систем

Таблица 7.1 – Перечень электронных библиотечных систем

№	Наименование ЭБС	Ссылка к ЭБС
1	Консультант студента	http://www.studentlibrary.ru/
2	Лань	https://e.lanbook.com/
3	Юрайт	https://biblio-online.ru/
4	TNT-ebook	https://www.tnt-ebook.ru/

7.2 Перечень свободно распространяемого программного обеспечения

Таблица 7.2 – Программное обеспечение, используемое студентами очного обучения

Программное обеспечение, используемое в университете на договорной основе	Программное обеспечение свободного распространения
-	Adobe Acrobat Reader (https://acrobat.adobe.com/ru/ru/acrobat/pdf-reader.html)
	Linux (https://www.linux.com/)
	OpenOffice (FreeWare) https://www.openoffice.org/ru/

	JDK 8 и выше (https://adoptopenjdk.net/)
	Фреймворк Java Spring 5 (https://spring.io/projects/spring-framework)
	Eclipse (https://www.eclipse.org/)
	IntelliJ Idea (https://www.jetbrains.com/ru-ru/idea/)
	git (https://git-scm.com/), github (https://github.com/)
	Maven (https://maven.apache.org/), Gradle (https://gradle.org/)
	Редактор блок-схем (https://app.diagrams.net/)

7.3 Перечень современных профессиональных баз данных и информационных справочных систем

В таблице 7.4 указан перечень профессиональных баз данных и информационных справочных систем, к которым обеспечен доступ (удаленный доступ). Данный перечень подлежит обновлению в соответствии с требованиями ФГОС ВО.

В данном разделе могут быть приведены ресурсы (ссылки на сайты), на которых можно найти полезную для курса информацию, в т.ч. статистические или справочные данные, учебные материалы, онлайн курсы и т.д.

Таблица 7.4 – Перечень современных профессиональных баз данных и информационных справочных систем

№	Наименование профессиональной базы данных, информационно-справочной системы	Доступ к ресурсу (удаленный доступ с указанием ссылки/доступ из локальной сети университета)
1	База данных стандартов и регламентов РОССТАНДАРТ	https://www.rst.gov.ru/portal/gost/home/standarts
2	Перечень профессиональных баз данных и информационных справочных систем	https://cyberpedia.su/21x47c0.html
3	Каталог паттернов проектирования	https://refactoring.guru/ru/design-patterns/catalog

8. ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОВЗ

В таблице 8.1 указан перечень образовательных ресурсов, имеющих формы, адаптированные к ограничениям их здоровья, а также сведения о наличии специальных технических средств обучения коллективного и индивидуального пользования. При заполнении таблицы может быть использована информация, размещенная в подразделе «Доступная среда» специализированного раздела сайта НГТУ «Сведения об образовательной организации» <https://www.nntu.ru/sveden/accenv/>

Таблица 8.1 – Образовательные ресурсы для инвалидов и лиц с ОВЗ

№	Перечень образовательных ресурсов, приспособленных для использования инвалидами и лицами с ОВЗ	Сведения о наличии специальных технических средств обучения коллективного и индивидуального пользования
1	ЭБС «Консультант студента»	озвучка книг и увеличение шрифта
2	ЭБС «Лань»	специальное мобильное приложение - синтезатор речи, который воспроизводит тексты книг и меню навигации
3	ЭБС «Юрайт»	версия для слабовидящих

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ, НЕОБХОДИМОЕ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ

Для контактной и самостоятельной работы обучающихся выделены помещения, оснащённые компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду организации:

- зал электронно-информационных ресурсов (ауд. 2210 – 11 компьютеров, ауд. 6119 – 9 компьютеров);
- читальный зал открытого доступа (ауд. 6162 – 2 компьютера);
- ауд. 2303, 2202, оборудованные Wi-Fi.

Перечень материально-технического обеспечения, необходимого для реализации программы бакалавриата для студентов очного обучения, включает в себя компьютерные классы.

1. Ауд. 4408 кафедры «Информатика и системы управления» – лаборатория Информационных технологий

Компьютеры, оснащенные необходимым оборудованием, техническими и электронными средствами обучения и контроля знаний студентов:

- 8 рабочих мест на базе тонких клиентов DellWise,
- мультимедийный проектор BenQ PB6240,
- ноутбук Lenovo V130-151KB,
- стенд для изучения автоматических систем управления на базе блока MyRio с FPGA под управлением LabView.

Пакеты ПО (лицензионное):

- Dr.Web (с/н ZNFC-CR5D-5U3U-JKGP от 20.05.2024).

Пакеты ПО (распространяемое по свободной лицензии):

- Apache OpenOffice;
- Linux Ubuntu 20.04 (<https://releases.ubuntu.com/20.04/>)
- git (<https://git-scm.com/>)
- Microsoft Visual Studio 2017 Community Edition (<https://visualstudio.microsoft.com/ru/vs/community/>)

Также, для самостоятельной работы обучающихся выделены помещения, оснащённые компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду организации:

Таблица 9.1 – Оснащенность аудиторий и помещений для самостоятельной работы студентов по дисциплине

№	Наименование аудиторий и помещений для самостоятельной работы	Оснащенность аудиторий помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	1	2	3

1	6421 учебная аудитория для проведения занятий лекционного и семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации; г. Нижний Новгород, Казанское ш., 12	1. Доска меловая – 1 шт. 3. Экран – 1 шт. 4. Мультимедийный проектор Epson X12 – 1 шт. 5. Компьютер PC MB Asus на чипсете Nvidia/AMDAthlonXII CPU 2.8Ggz/ RAM 4 Ggb/SVGAStandartGraphics +Ge-FORCE Nvidia GT210/HDD 250 Ggb, SATAinterface, монитор 19”, с выходом на проектор. 6. Рабочее место студента - 74 7. Рабочее место для преподавателя – 1 шт.	1. Windows 7 32 bit корпоративная; VL 49477S2 2. Adobe Acrobat Reader DC-Russian (беспл.) 3. Microsoft Office Professional Plus 2007 (лицензия № 42470655); 4. Dr.Web (с/н ZNFC-CR5D-5U3U-JKGP от 20.05.2024)
2	6543 компьютерный класс - помещение для СРС, курсового проектирования (выполнения курсовых работ), г. Нижний Новгород, Казанское ш., 12)	1. Рабочие места студента, оснащенные ПК на базе Intel Core i5 с мониторами – 8 шт. 2. Рабочие места студента, оснащенные ПК на базеCore 2 Duo с мониторами –2 шт. 3. Рабочее место преподавателя, оснащенное ПК на базе Intel Core i5 с монитором – 1 шт. 4. Проектор Ассер, проекционный экран – 1 шт. ПК подключены к сети «Интернет» и обеспечивают доступ в электронную информационно-образовательную среду университета 5. Принтер HP LaserJet 1200 – 1 шт.	1. Microsoft Windows 7 MSDN реквизиты договора - подписка DreamSpark Premium, договор № 0509/KMP от 15.10.18 2. Бесплатное ПО: Пакет программ Open Office, True Conf, Браузер Google Chrome, Браузер Mozilla Firefox, Браузер Opera, McAfee Security Scan, Adobe Acrobat Reader DC, AutoCAD2013

10. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

10.1 Общие методические рекомендации для обучающихся по освоению дисциплины, образовательные технологии

Дисциплина реализуется посредством проведения контактной работы с обучающимися (включая проведение текущего контроля успеваемости), самостоятельной работы обучающихся и промежуточной аттестации.

При преподавании дисциплины «Анализ больших данных», используются современные образовательные технологии, позволяющие повысить активность студентов при освоении материала курса и предоставить им возможность эффективно реализовать часы самостоятельной работы.

Весь лекционный материал курса сопровождается компьютерными презентациями, в которых наглядно преподносятся материал различных разделов курса и что дает возможность обсудить материал со студентами во время чтения лекций, активировать их деятельность при освоении материала. Электронные материалы лекций в период дистанционного обучения отправляются по электронной почте на адреса групп и могут быть получены до чтения лекций и проработаны студентами в ходе самостоятельной работы.

На лекциях, практических занятиях реализуются интерактивные технологии, приветствуются вопросы и обсуждения, используется лично-ориентированный подход,

технология работы в малых группах, что позволяет студентам проявить себя, получить навыки самостоятельного изучения материала, выровнять уровень знаний в группе.

Все вопросы, возникшие при самостоятельной работе над домашним заданием, подробно разбираются на практических занятиях и лекциях. Проводятся индивидуальные и групповые консультации с использованием современных информационных технологий: электронная почта, мессенджеры, Zoom.

Иницируется активность студентов, поощряется задание любых вопросов по материалу, практикуется индивидуальный ответ на вопросы студента, рекомендуются методы успешного самостоятельного усвоения материала в зависимости от уровня его базовой подготовки.

Для оценки знаний, умений, навыков и уровня сформированности компетенции применяется традиционная система контроля и оценки успеваемости студентов в процессе текущего контроля.

Промежуточная аттестация проводится в форме экзамена с учетом текущей успеваемости.

Результат обучения считается сформированным на повышенном уровне, если теоретическое содержание курса освоено полностью. При устных собеседованиях студент исчерпывающе, последовательно, четко и логически излагает учебный материал; свободно справляется с заданиями, вопросами, использует в ответе дополнительный материал. Все предусмотренные рабочей учебной программой задания выполнены в соответствии с установленными требованиями, студент способен анализировать полученные результаты, проявляет самостоятельность при выполнении заданий.

Результат обучения считается сформированным на пороговом уровне, если теоретическое содержание курса освоено полностью. При устных собеседованиях студент последовательно, четко и логически излагает учебный материал; справляется с заданиями, вопросами, требующих применения знаний; все предусмотренные рабочей учебной программой задания выполнены в соответствии с установленными требованиями, студент способен анализировать полученные результаты; проявляет самостоятельность при выполнении заданий.

Результат обучения считается несформированным, если студент при выполнении заданий не демонстрирует знаний учебного материала, допускает ошибки, неуверенно, с большими затруднениями выполняет задания, не демонстрирует необходимых умений, качество выполненных заданий не соответствует установленным требованиям, качество их выполнения оценено числом баллов ниже трех по оценочной системе, что соответствует пороговому уровню.

10.2 Методические указания для занятий лекционного типа

Лекционный курс предполагает систематизированное изложение основных вопросов тематического плана. В ходе лекционных занятий раскрываются базовые вопросы в рамках каждой темы дисциплины (Таблицы 4.4, 4.5, 4.6). Обозначаются ключевые аспекты тем, а также делаются акценты на наиболее сложные и важные положения изучаемого материала. Материалы лекций являются опорной основой для подготовки обучающихся к лабораторным занятиям и выполнения заданий самостоятельной работы, а также к мероприятиям текущего контроля успеваемости и промежуточной аттестации по дисциплине.

10.3 Методические указания по освоению дисциплины на лабораторных работах

Подготовку к каждой лабораторной работе студент должен начать с ознакомления с планом занятия, который отражает содержание предложенной темы. Каждая выполненная работа с оформленным отчетом подлежит защите у преподавателя.

При оценивании лабораторных работ учитывается следующее:

- качество выполнения практической части работы и степень соответствия результатов работы заданным требованиям;
- качество оформления отчета по работе;
- качество устных ответов на контрольные вопросы при защите работы.

10.4 Методические указания по освоению дисциплины на практических занятиях

Практические занятия по дисциплине не предусмотрены.

10.5 Методические указания по освоению дисциплины на курсовой работе

Курсовая работа не предусмотрена учебным планом.

10.6 Методические указания по самостоятельной работе обучающихся

Самостоятельная работа обеспечивает подготовку обучающегося к аудиторным занятиям и мероприятиям текущего контроля и промежуточной аттестации по изучаемой дисциплине. Результаты этой подготовки проявляются в активности обучающегося на занятиях и в качестве выполненных практических заданий и других форм текущего контроля.

При выполнении заданий для самостоятельной работы рекомендуется проработка материалов лекций по каждой пройденной теме, а также изучение рекомендуемой литературы, представленной в Разделе 6.

В процессе самостоятельной работы при изучении дисциплины студенты могут работать на компьютере в специализированных аудиториях для самостоятельной работы, указанных в Разделе 9. В аудиториях имеется доступ через информационно-телекоммуникационную сеть «Интернет» к электронной информационно-образовательной среде университета (ЭИОС) и электронной библиотечной системе (ЭБС), где в электронном виде располагаются учебные и учебно-методические материалы, которые могут быть использованы для самостоятельной работы при изучении дисциплины.

11. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

11.1 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта в ходе контроля текущей успеваемости

Для текущего контроля знаний студентов по дисциплине проводится комплексная оценка знаний, включающая

- защиту лабораторных работ.

11.1.1. Типовые задания для лабораторных работ

Типовые задания для лабораторных работ приведены в учебно-методических указаниях по проведению лабораторных работ.

11.2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта в ходе промежуточной аттестации по дисциплине

11.2.1. Защита курсового проекта/ работы

Курсовая работа не предусмотрена учебным планом.

11.2.2. Зачет с оценкой для студентов очной формы обучения в 7 семестре, для студентов заочной формы – в 8 семестре. Проводится в виде устного собеседования по типовым вопросам.

Типовые вопросы для промежуточной аттестации в форме зачета с оценкой для студентов очной и заочной формы обучения:

1. Основные понятия: большие данные, наука о данных, анализ данных, машинное обучение. Основные этапы извлечения знаний из данных. Примеры задач анализа данных.
2. Инструменты для разметки данных: преимущества и недостатки.
3. Обработка данных: описательные статистики, поиск аномалий (включая гистограммы, ящичковые диаграммы, ядерные оценки плотности), анализ выбросов и шумов. Нормализация и стандартизация данных.
4. Сравнение форматов хранения табличных данных по производительности и особенностям использования.
5. Матричные преобразования, векторы, функции регрессии.
6. Машинное обучение: основные понятия, задачи, которые можно решить с помощью машинного обучения. Алгоритмы машинного обучения с учителем и без учителя (перечислить).
7. Машинное обучение: этапы моделирования. Отбор признаков на основе модели. Оценка качества построенных моделей.
8. Машинное обучение: ансамбли алгоритмов. Способы реализации ансамблей на языке python.
9. Специфические проблемы при обучении классических моделей на больших данных.
10. Проблема широких данных и подходы к ее решению.
11. Метод хэширования признаков.
12. Суть коллаборативной фильтрации, ее виды.
13. Нейронные сети: базовые понятия, отличия нейросетей от классических методов машинного обучения.
14. Рекомендательные системы. Основные подходы к построению рекомендательных систем. Примеры рекомендательных систем.
15. Измерение схожести товаров в рекомендательных системах.
16. Принцип работы матричных факторизаций в рекомендательных системах.
17. Основные задачи автоматической обработки текста.
18. Специфические проблемы обработки естественного языка на больших данных.
19. Отличия между дистрибутивными и контекстуализированными моделями.
20. Векторные модели представления текста.

11.2.3. Экзамен для студентов очной формы обучения в 8 семестре, для студентов заочной формы – в 9 семестре. Проводится в виде устного собеседования по типовым вопросам.

Типовые вопросы для промежуточной аттестации в форме экзамена для студентов очной и заочной формы обучения:

1. Понятие большие данные.
2. Основные подходы к масштабированию обучения нейронных сетей для больших данных.

3. «Жизненный цикл» проекта по аналитике больших данных.
4. Корреляция. Регрессионный анализ.
5. Примеры использования корреляции в области больших данных.
6. Обработка текстов с помощью глубоких нейросетей.
7. Специфические проблемы при обучении глубоких нейронных сетей на распределенных кластерах.
8. Рекуррентные нейронные сети для обработки текстов.
9. Одномерные сверточные нейронные сети для обработки текста.
10. Сравнение эффективности синхронного и асинхронного SGD при распределенном обучении.
11. Полносвязные и рекуррентные нейронные сети для прогнозирования временных рядов.
12. Прямые и сверточные нейронные сети для обработки аудио сигналов.
13. Стратегии развертывания ML-сервисов.
14. Парадигма MapReduce. Ее реализация Hadoop.
15. Проблема переобучения. Регуляризация.
16. Организация A/B тестирования для ML-моделей в production.
17. Интерпретация онлайн-метрик при оценке бизнес-эффективности модели.
18. Требования к ML-моделям перед вводом в промышленную эксплуатацию.
19. Методика проверки статистических гипотез при анализе качества ML-моделей.
20. Принципы контейнеризации ML-приложений.

В полном объеме оценочные средства имеются на кафедре «Информатика и системы управления». Оценочные средства могут быть получены по требованию.