

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

1

Рабочая программа дисциплины: разработана в соответствии с Федеральным государственным образовательным стандартом высшего образования (ФГОС ВО 3++) по направлению подготовки 09.04.02 «Информационные системы и технологии», утвержденного приказом МИНОБРНАУКИ РОССИИ от 19 сентября 2017 года № 917 на основании учебного плана принятого УМС НГТУ
протокол № 17 от 28.05.2024
протокол № 6 от 17.12.2024 г

Рабочая программа одобрена на заседании кафедры протокол от 31.03.2025 № 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 № 09.04.02-т-16
Начальник МО _____ Е.Г. Севрюкова

Заведующая отделом комплектования НТБ

(подпись)

Н.И. Кабанина

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.....	8
5. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	11
5.1 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности.....	11
5.2 ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ КОНТРОЛЯ УСПЕВАЕМОСТИ, ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ	18
6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	20
7. ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ	21
7.1 ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ	21
7.2 ПЕРЕЧЕНЬ СВОБОДНО РАСПРОСТРАНЯЕМОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ	21
7.3 ПЕРЕЧЕНЬ СОВРЕМЕННЫХ ПРОФЕССИОНАЛЬНЫХ БАЗ ДАННЫХ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ.....	21
8. ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОВЗ.....	22
9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ, НЕОБХОДИМОЕ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ.....	22
10. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ	24
10.1 ОБЩИЕ МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ, ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ.....	24
10.2 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ЗАНЯТИЙ ЛЕКЦИОННОГО ТИПА.....	25
10.3 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ НА ЛАБОРАТОРНЫХ РАБОТАХ.....	25
10.4 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ НА ПРАКТИЧЕСКИХ ЗАНЯТИЯХ	25
10.5 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ НА КУРСОВОЙ РАБОТЕ	25
10.6 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО САМОСТОЯТЕЛЬНОЙ РАБОТЕ ОБУЧАЮЩИХСЯ	25
11. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ	26
11.1 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта в ходе контроля текущей успеваемости.....	26
11.2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта в ходе промежуточной аттестации по дисциплине.....	26

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1.1 Цель освоения дисциплины

Целями изучения дисциплины являются изучение методов и подходов к анализу данных различного объема, включая предварительную обработку данных и статистический анализ, освоение различных моделей машинного обучения, предназначенных для решения задач кластеризации, классификации и регрессии и применение их для решения прикладных задач из различных сфер человеческой деятельности.

1.2 Задачи освоения дисциплины (модуля)

Дисциплина «Предиктивная аналитика и анализ данных» способствует подготовке студентов к решению следующих профессиональных задач:

1. Разработка программного кода для эффективной обработки распределенных данных большого объема в сочетании с распределенным хранением данных в Hadoop кластере.
2. Построение и применение на практике описательных и прогнозных моделей интеллектуального анализа больших данных и машинного обучения.
3. Использование программных средств визуализации и интерактивного исследования больших данных.

1. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Учебная дисциплина «Предиктивная аналитика и анализ данных» Б1.В.ОД.5 включена в обязательный перечень дисциплин вариативной части (формируемой участниками образовательных отношений), определяющий направленность образовательной программы. Дисциплина реализуется в соответствии с требованиями ФГОС, ОП ВО и УП, по данному направлению подготовки.

Дисциплина базируется на дисциплинах математического блока программы магистратуры по направлению «Информационные системы и технологии».

Дисциплина «Предиктивная аналитика и анализ данных» является основополагающей для дисциплин «Стенографические методы защиты информации», «Технологии больших данных», научно-исследовательской работы и преддипломной практики.

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

Дисциплина «Предиктивная аналитика и анализ данных» формирует компетенции ПКС-1 совместно с дисциплинами и практиками, указанными в таблице 3.1.

Дисциплинарная часть компетенции ПКС-1 «Способен использовать методы научных исследований в профессиональной деятельности»: способен применять методы статистического анализа при обработке данных.

Таблица 3.1- Формирование компетенций дисциплинам

Наименование дисциплин, формирующих компетенцию совместно	Семестры, формирования дисциплины Компетенции берутся из Учебного плана по направлению подготовки бакалавра /специалиста/магистра»			
	1	2	3	4
<i>ПКС-1 (Способен использовать методы научных исследований в профессиональной деятельности)</i>				
<i>Предиктивная аналитика и анализ данных</i>				
<i>Стенографические методы защиты информации</i>				
<i>Технологии больших данных</i>				
<i>Научно-исследовательская работа</i>				
<i>Преддипломная практика</i>				
<i>Выполнение и защита ВКР</i>				

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-1. Способен использовать методы научных исследований в профессиональной деятельности	ИПКС-1.1. Использует теоретические методы научных исследований в профессиональной деятельности	Знать: – программную платформу аналитики данных	Уметь: – реализовывать на практике алгоритмы предиктивной аналитики	Владеть: - методами предиктивной аналитики	Выполнение и сдача 3 лабораторных работ, выполнение четырех тестирований промежуточного контроля	Вопросы для устного собеседования – 20 вопросов

Освоение дисциплины причастно к трудовым функциям:

ТФ D/03.7 (ПС 06.022 «Системный аналитик»), в результате освоения дисциплины студент способен применять методологии разработки программного обеспечения, разрешать проблемные ситуации, предсказывать риски

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1 Распределение трудоёмкости дисциплины по видам работ по семестрам

Общая трудоёмкость дисциплины составляет 3 зач.ед. 108 часа, распределение часов по видам работ семестрам представлено в таблице 4.1.

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 2 сем
Формат изучения дисциплины	с использованием элементов электронного обучения	
Общая трудоёмкость дисциплины по учебному плану	108	108
1. Контактная работа:	55	55
1.1 Аудиторная работа, в том числе:	51	51
занятия лекционного типа (Л)	17	17
занятия семинарского типа (ПЗ-семинары, практ. Занятия и др)	-	-
лабораторные работы (ЛР)	34	34
1.2 Внеаудиторная, в том числе	4	4
курсовая работа (проект) (КР/КП) (консультация, защита)	-	-
текущий контроль, консультации по дисциплине	4	4
контактная работа на промежуточном контроле (КРА)	-	-
2. Самостоятельная работа (СРС)	53	53
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.)	35	35
Подготовка к зачету	18	18

4.2 Содержание дисциплины, структурированное по темам

Таблица 4.2-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 образовательны х технологий	Реализация в рамках Практической подготовки (трудоемкость в часах)	Наименование разработанного Электронного курса (трудоемкость в часах)	
		Контактная работа				Сам осто ятел ьна я рабо та студ енто в (час)					
		Лек ции (час)	Лаб орат орн ые рабо ты (час)	Пра кти ческ ие зая ния (час)	КСР						
2 семестр											
Раздел 1. Введение в алгоритмы предиктивной аналитики											
ПКС-1, ИПКС-1.1 ПКС-3, ИПКС-3.1	Тема 1.1. Предиктивная аналитика, ее задачи и место в технологиях обработки данных. Понятие основ анализа данных.	1				2	Подготовка к лекциям [6.1.1, 6.1.2, 6.1.4]				
	Тема 1.2. Введение в машинное обучение. Подготовка данных, применение алгоритмов машинного обучения.	2				3	Подготовка к лекциям [6.1.1, 6.1.2, 6.1.4]				
	Тема 1.3. Байесовская классификация.	2				3	Подготовка к лекциям [6.1.1, 6.1.2, 6.1.4]				
	Лабораторная работа. Реализация наивного байесовского классификатора.		11		1	10	Подготовка к лекциям [6.1.1, 6.1.2, 6.1.4]				
	Итого по 1 разделу	5	11		1	18					

Раздел 2. Пакеты для работы с данными									
ПКС-1, ИПКС-1.1 ПКС-3, ИПКС-3.1	Тема 2.1. Алгоритмы классификации данных. Ошибки классификации. Экспериментальная оценка качества алгоритмов классификации.	2				3	Подготовка к лекциям [6.1.1,6.1.2]	Разбор конкретных ситуаций	
	Тема 2.2. Корреляционный и регрессионный анализ данных.	2				3	Подготовка к лекциям [6.1.1, 6.1.2, 6.1.4]	Разбор конкретных ситуаций	
	Лабораторная работа. Разведочный и регрессионный анализ данных.		11		1	10	Подготовка к лабораторной работе. [6.1.1, 6.1.2, 6.1.4]	Разбор конкретных ситуаций	
	Итого по 2 разделу	4	11		1	19			
Раздел 3. Математические основы больших данных									
ПКС-1, ИПКС-1.1 ПКС-3, ИПКС-3.1	Тема 3.1. Введение в нейронные сети. Искусственная модель нейрона. Классификация нейронных сетей. Обучение нейросети.	2				3	Подготовка к лекциям [6.1.1 – 6.1.3]	Разбор конкретных ситуаций	
	Тема 3.2. Рекуррентные нейронные сети (RNN).	2	6			3	Подготовка к лекциям [6.1.1 – 6.1.4]	Разбор конкретных ситуаций	
	Тема 3.3. Метод градиентного спуска. Метод обратного распространения ошибки. Настройка обучения нейронных сетей.	4				3	Подготовка к лекциям [6.1.1, 6.1.2, 6.1.4]	Разбор конкретных ситуаций	

	Лабораторная работа. Анализ текстовой информации на основе рекуррентной нейронной сети.		12		2	10	Подготовка к лабораторной работе.[6.1.1 6.1.4,6.1.6]	к			
	Итого по 3 разделу	8	12		2	16		—			
	Итого за семестр	17	34	-	4	53					

4. 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?
 - Чем отличаются структурированные данные от неструктурированных?
 - Опишите процедуру сбора и хранения больших данных?
 - Какие бывают структуры данных в библиотеке pandas?
 - С помощью чего визуализируются данные в python, назовите основные библиотеки?
 - Что такое задача регрессии?
 - Что такое градиентный спуск?
 - Какие матричные преобразования используются в машинном обучении?
 - Какие основные задачи решает машинное обучение?
 - Что такое ансамбли алгоритмов и когда они применяются?
 - Какие функции активации вам известны?
 - Опишите основные способы обучения нейронных сетей? В чем их отличия от алгоритмов машинного обучения?
 - Какие методы применяются для работы с данными?
 - Назовите основные подходы для работы с изображениями? Какие задачи решает компьютерное зрение?
 - Опишите задачу сегментации?
 - Что такое обработка естественного языка, назовите основные алгоритмы?

2. Примерный тест по теме «Ошибки классификации. Экспериментальная оценка качества алгоритмов».

Вопросы для оценки компетенций ПСК-1:

1. Что такое ошибка классификации?
 - а) Количество правильно классифицированных объектов.
 - б) Количество объектов, для которых алгоритм выдал неверный класс.
 - в) Сумма вероятностей принадлежности объекта ко всем классам.
 - г) Разность между предсказанными и истинными значениями.
2. Какая метрика оценивает долю верно классифицированных объектов от общего числа объектов?
 - а) Precision (точность).
 - б) Recall (полнота).
 - в) Accuracy (достоверность/правильность).
 - г) F1-score.
3. Что такое “False Positive” (FP)?
 - а) Объект, который был верно классифицирован как положительный.
 - б) Объект, который был верно классифицирован как отрицательный.
 - в) Объект, который был ошибочно классифицирован как положительный, хотя на самом деле он отрицательный.
 - г) Объект, который был ошибочно классифицирован как отрицательный, хотя на самом деле он положительный.

4. Какая метрика показывает, какую долю объектов, действительно принадлежащих к положительному классу, алгоритм смог правильно определить?
- Precision (точность).
 - Recall (полнота).
 - Accuracy (достоверность/правильность).
 - F1-score.
5. Какая метрика показывает, какую долю объектов, классифицированных как положительные, действительно являются положительными?
- Precision (точность).
 - Recall (полнота).
 - Accuracy (достоверность/правильность).
 - F1-score.
6. Что такое F1-score?
- Среднее арифметическое между Precision и Recall.
 - Среднее геометрическое между Precision и Recall.
 - Гармоническое среднее между Precision и Recall.
 - Сумма Precision и Recall.
7. В каком случае F1-score наиболее полезна?
- Когда классы сбалансированы.
 - Когда необходимо максимизировать количество правильно классифицированных объектов.
 - Когда есть дисбаланс классов и важно учитывать как Precision, так и Recall.
 - Когда важна только Accuracy.
8. Что такое ROC-кривая?
- График, показывающий зависимость Accuracy от порога классификации.
 - График, показывающий зависимость Precision от Recall.
 - График, показывающий зависимость True Positive Rate (TPR) от False Positive Rate (FPR) для различных порогов классификации.
 - График, показывающий распределение объектов по классам.
9. Что такое порог классификации?
- Значение, разделяющее объекты на два класса.
 - Параметр, определяющий скорость обучения модели.
 - Функция, отображающая входные данные в пространство признаков.
 - Метод уменьшения переобучения модели.
10. Значение AUC, равное 0.5, говорит о том, что:
- Классификатор идеально классифицирует объекты.
 - Классификатор работает хуже, чем случайное угадывание.
 - Классификатор работает так же, как случайное угадывание.
 - Классификатор не может быть использован для данной задачи.

Регламент проведения промежуточной аттестации в форме компьютерного тестирования

Кол-во заданий в банке вопросов	Кол-во заданий, предъявляемых студенту	Время на тестирование, мин.
не менее 20 заданий	10 заданий	20 минут

Полный фонд оценочных средств для проведения промежуточной аттестации в форме компьютерного тестирования размещен в банке вопросов данного курса дисциплины в СДО eLearning Server 5G ЭИОС НГТУ.

В ходе подготовки к промежуточной аттестации обучающимся предоставляется возможность пройти тест самопроверки. Тест для самопроверки по дисциплине размещен в eLearning Server 5G ЭИОС НГТУ в свободном для студентов доступе.

Примерный тест по теме «Классификация и кластеризация данных»

Вопросы для оценки компетенций ПСК-1:

1. Какой принцип лежит в основе наивного байесовского классификатора?
 - а) Минимизация расстояния между объектами в разных классах.
 - б) Максимизация вероятности принадлежности объекта к определенному классу на основе теоремы Байеса.
 - в) Поиск оптимальной разделяющей гиперплоскости между классами.
 - г) Создание иерархической структуры для группировки объектов.
2. Какое основное допущение делает наивный байесовский классификатор?
 - а) Независимость объектов друг от друга.
 - б) Независимость признаков друг от друга внутри каждого класса.
 - в) Равномерное распределение объектов по классам.
 - г) Линейную зависимость между признаками и классами.
3. Какой тип данных наилучшим образом подходит для наивного байесовского классификатора?
 - а) Числовые данные с нормальным распределением.
 - б) Категориальные данные.
 - в) Временные ряды.
 - г) Графовые данные.
4. Что такое “классификация” в машинном обучении?
 - а) Группировка объектов по их признакам.
 - б) Предсказание категориальной переменной.
 - в) Поиск аномалий в данных.
 - г) Уменьшение размерности данных.
5. Что такое “кластеризация” в машинном обучении?
 - а) Предсказание числовой переменной.
 - б) Разделение данных на группы на основе их схожести без предварительной информации о классах.
 - в) Обучение модели на размеченных данных.
 - г) Оценка качества модели.
6. Какой из алгоритмов является алгоритмом классификации?
 - а) K-Means.
 - б) DBSCAN.
 - в) Support Vector Machine (SVM).
 - г) Hierarchical Clustering.

7. Какой из алгоритмов является алгоритмом кластеризации?
- Decision Tree.
 - Logistic Regression.
 - K-Nearest Neighbors (KNN).
 - DBSCAN.
8. Какой алгоритм кластеризации требует предварительного задания количества кластеров?
- DBSCAN.
 - K-Means.
 - Hierarchical Clustering (в некоторых вариантах).
 - Affinity Propagation.
9. Какой алгоритм кластеризации основан на плотности данных?
- K-Means.
 - Hierarchical Clustering.
 - DBSCAN.
 - Gaussian Mixture Models (GMM).
10. Какой из методов можно использовать для оценки качества кластеризации без знания истинных меток классов?
- Accuracy.
 - Precision.
 - Silhouette Score.
 - Recall.

Регламент проведения промежуточной аттестации в форме компьютерного тестирования

Кол-во заданий в банке вопросов	Кол-во заданий, предъявляемых студенту	Время на тестирование, мин.
не менее 20 заданий	10 заданий	20 минут

Полный фонд оценочных средств для проведения промежуточной аттестации в форме компьютерного тестирования размещен в банке вопросов данного курса дисциплины в СДО eLearning Server 5G ЭИОС НГТУ.

В ходе подготовки к промежуточной аттестации обучающимся предоставляется возможность пройти тест самопроверки. Тест для самопроверки по дисциплине размещен в eLearning Server 5G ЭИОС НГТУ в свободном для студентов доступе.

**Примерный тест по теме «Рекуррентные нейронные сети»
Вопросы для оценки компетенций ПСК-1:**

- Что такое нейронная сеть?
 - Программа для работы с базами данных.
 - Модель машинного обучения, вдохновленная структурой человеческого мозга.
 - Алгоритм сжатия данных.
 - Инструмент для анализа социальных сетей.
- Что является основным элементом нейронной сети?
 - Байесовская сеть.
 - Нейрон (персептрон).
 - Дерево решений.
 - Кластер.

3. Что такое функция активации в нейронной сети?

- а) Функция, определяющая скорость обучения сети.
- б) Функция, применяемая к выходным данным нейрона для внесения нелинейности.
- в) Функция, вычисляющая ошибку сети.
- г) Функция, используемая для нормализации данных.

4. Что такое “backpropagation” (обратное распространение ошибки)?

- а) Алгоритм классификации данных.
- б) Алгоритм кластеризации данных.
- в) Алгоритм обучения нейронной сети, основанный на распространении ошибки от выхода к входу.
- г) Метод оптимизации гиперпараметров.

5. Что такое рекуррентная нейронная сеть (RNN)?

- а) Тип нейронной сети, используемый для обработки изображений.
- б) Тип нейронной сети, предназначенный для обработки последовательностей данных.
- в) Тип нейронной сети, используемый для кластеризации данных.
- г) Тип нейронной сети, не имеющий скрытых слоев.

6. Какая основная особенность RNN, отличающая их от обычных нейронных сетей?

- а) Наличие обратных связей, позволяющих сохранять информацию о предыдущих шагах последовательности.
- б) Использование только линейных функций активации.
- в) Отсутствие скрытых слоев.
- г) Ограничение на количество входных данных.

7. Что такое проблема исчезающего градиента (vanishing gradient problem)?

- а) Проблема, возникающая при обучении RNN, когда градиенты становятся очень малыми, что затрудняет обучение.
- б) Проблема, возникающая из-за недостатка данных для обучения.
- в) Проблема, связанная с переобучением модели.
- г) Проблема, связанная с большой вычислительной сложностью обучения.

8. Какие архитектуры RNN используются для решения проблемы исчезающего градиента?

- а) CNN (Convolutional Neural Networks).
- б) GAN (Generative Adversarial Networks).
- в) LSTM (Long Short-Term Memory) и GRU (Gated Recurrent Unit).
- г) SVM (Support Vector Machines).

9. Для каких задач особенно хорошо подходят RNN?

- а) Классификация изображений.
- б) Прогнозирование временных рядов, обработка естественного языка, машинный перевод.
- в) Кластеризация данных.
- г) Линейная регрессия.

10. Что такое “токен” в контексте обработки естественного языка с использованием RNN?

- а) Гиперпараметр нейронной сети.
- б) Отдельное слово или часть слова, используемая для представления текста.
- в) Функция активации нейрона.
- г) Метод визуализации данных.

Регламент проведения промежуточной аттестации в форме компьютерного тестирования

Кол-во заданий в банке вопросов	Кол-во заданий, предъявляемых студенту	Время на тестирование, мин.
не менее 20 заданий	10 заданий	20 минут

Полный фонд оценочных средств для проведения промежуточной аттестации в форме компьютерного тестирования размещен в банке вопросов данного курса дисциплины в СДО eLearning Server 5G ЭИОС НГТУ.

В ходе подготовки к промежуточной аттестации обучающимся предоставляется возможность пройти тест самопроверки. Тест для самопроверки по дисциплине размещен в eLearning Server 5G ЭИОС НГТУ в свободном для студентов доступе.

Примерный тест по теме «Предиктивная аналитика»
Вопросы для оценки компетенций ПСК-1:

1. Что такое предиктивная аналитика?
 - а) Описание исторических данных.
 - б) Использование статистических методов и машинного обучения для прогнозирования будущих событий.
 - в) Визуализация данных в графической форме.
 - г) Инструмент для сбора данных.
2. Какая основная цель предиктивной аналитики?
 - а) Оптимизация бизнес-процессов.
 - б) Снижение издержек.
 - в) Прогнозирование будущих событий или поведения.
 - г) Повышение лояльности клиентов.
3. Какие типы моделей используются в предиктивной аналитике?
 - а) Только регрессионные модели.
 - б) Только классификационные модели.
 - в) Регрессионные, классификационные, кластеризационные и другие модели машинного обучения.
 - г) Только модели временных рядов.
4. Что такое “обучающая выборка” в машинном обучении для предиктивной аналитики?
 - а) Данные, используемые для оценки точности модели.
 - б) Данные, используемые для обучения модели.
 - в) Данные, используемые для визуализации результатов.
 - г) Данные, используемые для очистки данных.
5. Что такое “тестовая выборка” в машинном обучении для предиктивной аналитики?
 - а) Данные, используемые для обучения модели.
 - б) Данные, используемые для оценки точности обученной модели на новых данных.
 - в) Данные, используемые для предварительной обработки данных.
 - г) Данные, используемые для выбора алгоритма.
6. Что такое “переобучение” (overfitting) в контексте предиктивной аналитики?
 - а) Модель плохо работает на обучающих данных.
 - б) Модель хорошо работает на обучающих данных, но плохо обобщает на новые данные.
 - в) Модель идеально подходит для всех типов данных.
 - г) Модель недостаточно обучена.

7. Какие методы используются для предотвращения переобучения?

- а) Увеличение размера обучающей выборки.
- б) Упрощение модели.
- в) Использование регуляризации.
- г) Все ответы верны.

8. Какие метрики используются для оценки моделей классификации?

- а) RMSE (Root Mean Squared Error).
- б) MAE (Mean Absolute Error).
- в) Accuracy, Precision, Recall, F1-score.
- г) R-squared.

9. Какие метрики используются для оценки моделей регрессии?

- а) Accuracy, Precision, Recall, F1-score.
- б) Silhouette Score.
- в) RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), R-squared.
- г) Adjusted Rand Index.

10. Какие источники данных могут использоваться в предиктивной аналитике?

- а) Базы данных.
- б) Социальные сети.
- в) Датчики и устройства IoT.
- г) Все ответы верны.

Регламент проведения промежуточной аттестации в форме компьютерного тестирования

Кол-во заданий в банке вопросов	Кол-во заданий, предъявляемых студенту	Время на тестирование, мин.
не менее ²⁰ заданий	10 заданий	20 минут

Полный фонд оценочных средств для проведения промежуточной аттестации в форме компьютерного тестирования размещен в банке вопросов данного курса дисциплины в СДО eLearning Server 5G ЭИОС НГТУ.

В ходе подготовки к промежуточной аттестации обучающимся предоставляется возможность пройти тест самопроверки. Тест для самопроверки по дисциплине размещен в eLearning Server 5G ЭИОС НГТУ в свободном для студентов доступе.

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

3. Примерный перечень вопросов для зачета:

1. Основные понятия: большие данные, наука о данных, анализ данных, машинное обучение. Основные этапы извлечения знаний из данных. Примеры задач анализа данных.
2. Структурированные и неструктурированные данные. Категориальные и непрерывные переменные. Методы отбора признаков (переменных). Библиотека pandas: объекты Series и DataFrame.
3. Обработка данных: поиск пропущенных значений, основные методы обработки пропущенных значений, обработка пропущенных значений с помощью pandas, поиск и удаление дублирующихся значений в pandas.
4. Обработка данных: описательные статистики, поиск аномалий (включая гистограммы, ящичковые диаграммы, ядерные оценки плотности), анализ выбросов и шумов. Нормализация и стандартизация данных.
5. Визуализация данных: виды графиков и диаграмм. Основные инструменты визуализации данных в Python.
6. Математика в Data Science: основные понятия, градиентный спуск.
7. Матричные преобразования, векторы, функции регрессии.
8. Кластерный анализ: иерархический кластерный анализ, построение дендрограмм, методы k-средних. Кластерный анализ в Python.
9. Машинное обучение: основные понятия, задачи, которые можно решить с помощью машинного обучения. Алгоритмы машинного обучения с учителем и без учителя (перечислить). Инструменты Python, используемые в машинном обучении.
10. Машинное обучение: этапы моделирования. Отбор признаков на основе модели. Оценка качества построенных моделей.
11. Машинное обучение: ансамбли алгоритмов. Способы реализации ансамблей на языке python.
12. Задача классификации: постановка задачи, пример моделей, понятие переобучения, оценка качества классификации, тонкая настройка модели.
13. Задача регрессии: постановка задачи, пример моделей, понятие переобучения, оценка качества классификации, тонкая настройка модели.
14. «Жизненный цикл» проекта по аналитике больших данных.
15. Корреляция. Регрессионный анализ.
16. Примеры использования корреляции в области больших данных.
17. Задачи в области больших данных, решаемые методом регрессионного анализа.
18. Постановка задачи классификации.
19. Постановка задачи кластеризации.
20. Задача построения ассоциативных правил.

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-1. Способен использовать методы научных исследований в профессиональной деятельности	ИПКС-1.1. Использует теоретические методы научных исследований в профессиональной деятельности	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.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1 Учебная литература

- 6.1.1. 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. Смагин, А. Д. Хомоненко. — 2-е изд., стер. — Санкт-Петербург : Лань, 2024. — 236 с. — ISBN 978-5-507-50398-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/425024>
- 6.1.2. Основы анализа данных и интеллектуальные системы : учебное пособие для вузов / С. Н. Косников, А. Л. Золкин, Ф. Р. Ахмадуллин [и др.]. — Санкт-Петербург : Лань, 2025. — 176 с. — ISBN 978-5-507-50239-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/440060>
- 6.1.3. Базовые возможности языка Python для анализа данных : учеб.-метод. пособие / С. В. Рындина. — Пенза : Изд-во ПГУ, 2022. — 72 с.
- 6.1.4. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — 3-е изд., перераб. и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.: ил. + CD-ROM — (Учебная литература для вузов).

6.2 Справочно-библиографическая литература

- 6.1.4. Атаян, А. М. Аналитика больших данных : учебно-методическое пособие / А. М. Атаян, М. Б. Вольфсон. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2023. — 69 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/426077>
- 6.1.5. 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>

6.3 Перечень журналов по профилю дисциплины:

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

6.4 Методические указания, рекомендации и другие материалы к занятиям

6.1.6. Методические указания для выполнения лабораторных работ по дисциплине «Предиктивная аналитика и анализ данных», для студентов направления подготовки 09.04.01 «Информатика и вычислительная техника» и 09.04.02 «Информационные системы и технологии» очной формы обучения / НГТУ; Сост.: Санников А.Н., Н.Новгород, 2025

Электронные варианты всех методических указаний находится в системе электронного курса на базе платформы iLearning 5G по ссылке: https://edu.ntnu.ru/subject/lessons/edit/subject_id/2175

6. 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	2	3
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

7. 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

8. 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. 4403 кафедры «Информатика и системы управления» - лаборатория Программирования АСО и У

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

- 10 АРМ (терминалов);
- мультимедийный проектор Vivitek H 1180,
- экран настенный LMP 100109,
- сетевая купольная PTZ-камера AXIS M5014.

Пакеты ПО (лицензионное):

- Dr.Web (с/н ZNFC-CR5D-5U3U-JKGP от 20.05.2024),
- MATLAB R2008a DVD KIT-WIN & UNIX/MAC (№ лицензии 527840, № заказа 2035235 Softline от 05.05.2008).

Пакеты ПО (распространяемое по свободной лицензии):

- ApacheOpenOffice;
- Eclipse (<https://www.eclipse.org/>)
- git (<https://git-scm.com/>)
- Microsoft Visual Studio 2017 Community Edition (<https://visualstudio.microsoft.com/ru/vs/community/>)

2. 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,	1. Доска меловая – 1 шт. 3. Экран – 1 шт. 4. Мультимедийный проектор Epson X12 – 1 шт. 5. Компьютер PC MB Asus на чипсете Nvidia/AMD AthlonXII CPU 2.8Ggz/ RAM 4 Ggb/SVGAS standartGraphics + GeForce Nvidia GT210/HDD 250Ggb,SATA interface, монитор 19", с выходом на проектор.	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)

	текущего контроля и промежуточной аттестации; г. Нижний Новгород, Казанское ш., 12	6. Рабочее место студента - 74 7. Рабочее место для преподавателя – 1 шт.	
	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 базе Core 2 Duo с мониторами – 2 шт. 3. Рабочее место преподавателя, оснащенное ПК на базе Intel Core i5 с монитором – 1 шт. 4. Проектор Acer, 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

9. 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 «Предиктивная аналитика», 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 личностно-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, Яндекс.Телемост.

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 балльно-рейтинговая система контроля и оценки успеваемости студентов в процессе текущего контроля.

Промежуточная аттестация проводится в форме дифференцированного зачета 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 практическим 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.

10. 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;
- прохождение тестирований промежуточного контроля знаний.

10.1.1. 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 для студентов очной формы обучения в 2 семестре.

Проводится в виде устного собеседования по типовым вопросам. Типовые вопросы для промежуточной аттестации в форме зачета для студентов очной 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. Структурированные и неструктурированные данные. Категориальные и непрерывные переменные. Методы отбора признаков (переменных). Библиотека pandas: объекты Series и DataFrame.
3. Обработка данных: поиск пропущенных значений, основные методы обработки пропущенных значений, обработка пропущенных значений с помощью pandas, поиск и удаление дублирующихся значений в pandas.
4. Обработка данных: описательные статистики, поиск аномалий (включая гистограммы, ящичковые диаграммы, ядерные оценки плотности), анализ выбросов и шумов. Нормализация и стандартизация данных.
5. Визуализация данных: виды графиков и диаграмм. Основные инструменты визуализации данных в Python.
6. Математика в Data Science: основные понятия, градиентный спуск.
7. Матричные преобразования, векторы, функции регрессии.

8. Кластерный анализ: иерархический кластерный анализ, построение дендрограмм, методы k-средних. Кластерный анализ в Python.
9. Машинное обучение: основные понятия, задачи, которые можно решить с помощью машинного обучения. Алгоритмы машинного обучения с учителем и без учителя (перечислить). Инструменты Python, используемые в машинном обучении.
10. Машинное обучение: этапы моделирования. Отбор признаков на основе модели. Оценка качества построенных моделей.
11. Машинное обучение: ансамбли алгоритмов. Способы реализации ансамблей на языке python.
12. Задача классификации: постановка задачи, пример моделей, понятие переобучения, оценка качества классификации, тонкая настройка модели.
13. Задача регрессии: постановка задачи, пример моделей, понятие переобучения, оценка качества классификации, тонкая настройка модели.
14. «Жизненный цикл» проекта по аналитике больших данных.
15. Корреляция. Регрессионный анализ.
16. Примеры использования корреляции в области больших данных.
17. Задачи в области больших данных, решаемые методом регрессионного анализа.
18. Постановка задачи классификации.
19. Постановка задачи кластеризации.
20. Задача построения ассоциативных правил.

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