

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Уфимский государственный нефтяной технический университет»

УТВЕРЖДАЮ

Директор IT-институт

З.Х. Павлова

(подпись)

29.06.2023

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

(48549)Статистические и вероятностные методы

Направление подготовки (специальность): **09.03.01 Информатика и вычислительная техника**

Направленность: **профиль«Технологии искусственного интеллекта в нефтегазовой отрасли»**

Уровень высшего образования: **бакалавриат**

Форма обучения: **очная;**

Кафедра, обеспечивающая преподавание дисциплины: **Вычислительная техника и инженерная кибернетика (ВТИК);**

Трудоемкость дисциплины: **3 з.е. (108час)**

Уфа 2023

Рабочую программу дисциплины разработал(и):

_____ старший преподаватель Шварева Е.Н.

Рецензент

_____ доцент, кандидат ф.-м.н. Зарипов Д. М.

Рабочая программа дисциплины рассмотрена и одобрена на заседании кафедры Вычислительная техника и инженерная кибернетика (ВТИК);, обеспечивающей преподавание дисциплины 31.08.2022, протокол №1.

Заведующий кафедрой

Вычислительная техника и инженерная кибернетика (ВТИК)_____..

СОГЛАСОВАНО

Заведующий кафедрой ВТИК_____..

Год приема 2023 г.

Рабочая программа зарегистрирована 19.09.2022 № 1 в ОПЛА и внесена в электронную базу данных

1. Место дисциплины в структуре ОПОП ВО

Дисциплины, предшествующие изучению данной дисциплины (исходя из формирования этапов по компетенциям): Основы нефтегазового дела; Основы переработки нефти и нефтепродуктов; Основы разработки нефтяных и газовых месторождений; Теория языков программирования; Транспортирование нефти, газа и продуктов их переработки

Дисциплины, для которых освоение данной дисциплины необходимо как предшествующее (исходя из формирования этапов по компетенциям): Базы данных; Интернет вещей; Криптографические алгоритмы; Методы трансляции; Объектно-ориентированное программирование; Операционные системы; Основы нефтегазохимии; Основы нефтепереработки; Преддипломная практика; Технологии бурения и разработки нефтегазовых месторождений; Технологическая (проектно-технологическая) практика; Трубопроводный транспорт углеводородов

Блок: Блок 1. Дисциплины (модули);

Обязательная или часть, формируемая участниками образовательных отношений (в том числе элективные дисциплины): Часть, формируемая участниками образовательных отношений;

Форма обучения: очная

Семестр, в котором преподается дисциплина	Трудоемкость дисциплины				Вид промежу- точной аттестации
	Зачет- ные едини- цы	Часы			
		Общая	В том числе		
			контакт- ная	СРО	
4	3	108	52	56	диф.зачет;
ИТОГО:	3	108	52	56	

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

№ пп.	Формируемые компетенции	Шифр/индекс компетенции
1	Способен классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта	ПК-1и-22 Г.-3
2	Способен разрабатывать и применять методы машинного обучения для решения задач	ПК-4и-22Г-1
3	Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	ПК-7и-22Г-1

В результате освоения дисциплины обучающийся должен:

Шифр компетенции	Индикаторы достижения компетенций	Шифр результата обучения	Результат обучения
ПК-1и-22 Г.	ПК 1.1 Классифицирует и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей	З(ПК-1и-22 Г.)	Знать: задачи систем искусственного интеллекта
		У(ПК-1и-22 Г.)	Уметь: умеет классифицировать задачи искусственного интеллекта
		В(ПК-1и-22 Г.)	Владеть: владеет идентификацией задач искусственного интеллекта
ПК-4и-22Г	ПК-4.1. Проводит анализ требований и определяет необходимые классы задач машинного обучения ПК-4.2. Определяет метрики оценки результатов моделирования и критерии качества построенных моделей ПК-4.3. Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения	З(ПК-4и-22Г)	Знать: метрики оценки результатов моделирования, критерии качества построенной модели
		У(ПК-4и-22Г)	Уметь: оценивать, выбирать, разрабатывать методы машинного обучения
		В(ПК-4и-22Г)	Владеть: анализом требований задач машинного обучения
ПК-7и-22Г	ПК-7.2. Выполняет подготовку и разметку структурированных и	З(ПК-7и-22Г)	Знать: структурированность

Шифр компетенции	Индикаторы достижения компетенций	Шифр результата обучения	Результат обучения
	неструктурированных данных для машинного обучения		данных для машинного обучения
		У(ПК-7и-22Г)	Уметь: размечать данные для машинного обучения
		В(ПК-7и-22Г)	Владеть: подготовкой данных для машинного обучения

3. Структура дисциплины

3.1. Виды учебной работы и трудоемкость (всего и по семестрам, в часах)

Форма обучения: очная

Вид учебной работы	Всего и по семестрам, часы	1	2	3	4	5	6	7	8	9	10	11	12
Контактная работа, всего в том числе:	52				52								
лекции (всего)	14				14								
-в т.ч. лекции on-line курс	0												
практические занятия (ПЗ)	36				36								
-в т.ч. практические занятия on-line курс	0												
лабораторные работы (ЛР)	0												
контролируемая самостоятельная работа (защита курсового проекта, курсовой работы и др. работ (при наличии))	0												
-в т.ч. лабораторные работы on-line курс	0												
иная контактная работа (сдача зачета, экзамена, консультации)	2				2								
проектная деятельность (ПД)	0												
Самостоятельная работа обучающихся (СРО), всего в том числе: (указать конкретный вид СРО)	56				56								
выполнение и подготовка к защите курсового проекта или курсовой работы	0												
выполнение и подготовка к защите РГР работы, реферата, патентных исследований, аналитических исследований и т.п	0												
изучение учебного материала, вынесенного на	14				14								

самостоятельную проработку													
подготовка к лабораторным и/или практическим занятиям	35				35								
подготовка к сдаче зачета, экзамена	7				7								
иные виды работ обучающегося (при наличии)	0												
освоение on-line курса	0												
самостоятельная проектная деятельность (СПД)	0												
ИТОГО ПО ДИСЦИПЛИНЕ	108				108								

4. Содержание дисциплины

4.1. Темы (разделы) дисциплины и виды занятий (в часах)

Форма обучения: очная

(раздела)Номер темы	Название темы (раздела)	Семестр	Трудоемкость, часы					Шифр результата обучения
			Л	ПЗ	ЛР	СРО	Всего	
1	Основы статистики	4	8	18		30	56	З(ПК-7и-22Г) У(ПК-7и-22Г) В(ПК-7и-22Г)
2	Методы машинного обучения	4	6	18		26	50	З(ПК-1и-22 Г.) З(ПК-4и-22Г) У(ПК-1и-22 Г.) У(ПК-4и-22Г) В(ПК-1и-22 Г.) В(ПК-4и-22Г)
	ИТОГО:		14	36		56	106	

4.2. Содержание лекционного курса

№ пп.	Номер раздела	Название темы	Трудоемкость, часы		
			очная	очно-заочная	заочная
1	1-Основы статистики	Основные понятия статистики, законы распределения, Основы статистического анализа, описательная статистика, визуализация, центральная предельная теорема, идея статистического вывода, проверка гипотезы о законе распределения, Законы распределения, нормальный закон распределения, критерий Пирсона, t-распределение, t-критерий Стьюдента, Распределение Хи-квадрат Пирсона, критерий Фишера	2		
2	1-Основы статистики	Дисперсионный и корреляционно-регрессионный анализ Дисперсионный анализ однофакторный и многофакторный, коэффициент корреляции Пирсона и Спирмана, регрессионный анализ с одной или несколькими переменными	4		
3	1-Основы статистики	Предподготовка данных Нормализация, восстановление пропущенных значений, поиск выбросов и аномалий	2		
4	2-Методы машинного обучения	Анализ номинальных данных, классификация, деревья, вероятностные алгоритмы Логистическая	4		

		регрессия, кластерный анализ, метод главных компонент, метод ближайших соседей Деревья, дерево решений, случайный лес,			
5	2-Методы машинного обучения	Ансамбли, гибридизация Ансамбли алгоритмов, эвристические алгоритмы, гибридизация алгоритмов и метаоптимизация	2		
	-	ИТОГО:	14		

4.4. Перечень практических занятий

Номер раздела	№ ПЗ	Тема практического занятия	Трудоемкость, часы		
			очная	очно-заочная	заочная
1-Основы статистики	1	Обработка статистических данных и проверка о законе распределения Меры центральной тенденции, меры изменчивости, визуализация, проверка закона распределения, критерий Пирсона, t-критерий Стьюдента, расчет уровня значимости, Законы распределения, нормальный закон распределения, критерий Пирсона, t-распределение, t-критерий Стьюдента, РАспределение Хи-квадрат Пирсона, критерий Фишера	2		

1-Основы статистики	2	Дисперсионный анализ Однофакторный дисперсионный анализ, многофакторный анализ	8		
1-Основы статистики	3	Корреляционно-регрессионный анализ коэффициент ковариации, корреляции Пирсона, коэффициент детерминации, коэффициент корреляции Спирмана, регрессионный анализ с одной переменной, регрессионный анализ с несколькими переменными, метод наименьших квадратов	8		
2-Методы машинного обучения	4	Анализ номинальных данных, классификация Анализ номинальных данных, классификация, логистическая регрессия, кластерный анализ, метод главных компонент,	8		
2-Методы машинного обучения	5	Деревья Деревья, дерево решений, случайный лес,	6		
2-Методы машинного обучения	6	Ансамбли алгоритмов, эвристические алгоритмы, гибридизация алгоритмов и метаоптимизация Ансамбли алгоритмов, эвристические алгоритмы, гибридизация алгоритмов и метаоптимизация	4		
-		ИТОГО:	36		

4.5. Виды СРО

Номер раздела	Вид СРО	Трудоемкость, часы		
		очная	очно-заочная	заочная
1-Основы статистики	подготовка к сдаче зачета, экзамена	4		
1-Основы статистики	подготовка к лабораторным и/или практическим занятиям	18		
1-Основы статистики	изучение учебного материала, вынесенного на самостоятельную проработку	8		
2-Методы машинного обучения	подготовка к сдаче зачета, экзамена	3		
2-Методы машинного обучения	подготовка к лабораторным и/или практическим занятиям	17		
2-Методы машинного обучения	изучение учебного материала, вынесенного на самостоятельную проработку	6		
-	ИТОГО:	56		

Темы для самостоятельной работы обучающихся

Раздел 1. Основы статистики

Устойчивые (Робастные) методы. Мультиномиальная регрессия. Порядковая регрессия. Регрессия Пуассона. Регрессия выживаемости

Раздел 2. Методы машинного обучения

bootstrap, ансамбли алгоритмов, градиентный бустинг

5. Формы текущего контроля успеваемости и проведения промежуточной аттестации

Перечень оценочных средств текущего контроля и промежуточной аттестации по дисциплине приведен Фонде оценочных средств (приложение Б).

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Учебно-методическое обеспечение

Сведения об обеспеченности дисциплины основной, дополнительной и учебно-методической литературой приведены в формах № 1-УЛ и № 2-УЛ (приложение А).

6.2. Перечень современных профессиональных баз данных и информационных справочных систем, рекомендуемых для освоения дисциплины

Названия современных профессиональных баз данных и информационных справочных систем, рекомендуемых для освоения дисциплины	Ссылки на официальные сайты
cyberleninka.ru	Электронная научная библиотека
elibrary.ru	Российская база научных публикаций
http://biblioclub.ru/	Университетская библиотека онлайн
http://docs.python.org/3/	Информационная база Python
http://ru.wikiversity.org/wiki	Программирование и научные вычисления на языке Python В свободном доступе.
https://docs.python.org/3/	Официальный сайт Python
https://e.lanbook.com/book/	Электронно-библиотечная система Лань
https://github.com/jrjohansson/scientific-python-lectures	Lectures on scientific computing with Python. В свободном доступе.
https://www.gpntb.ru/	Государственная публичная научно-техническая библиотека России
https://www.python.org	Python. The official Python web site. В свободном доступе.
https://www.r-project.org/	Официальный сайт “Проект R для статистических вычислений”
https://www.rsl.ru/	Российская государственная библиотека
https://znanium.com/	Электронно-библиотечная система
http://www.raai.org/resurs/resurs.shtml	российская ассоциация искусственного интеллекта
http://znanium.com/	ЭБС Znanium.com
IPRbooks	Электронная библиотечная система
rukont.ru	Электронная библиотечная система
scopus.com	международная база научных публикаций
Государственная публичная научно-техническая библиотека России	https://www.gpntb.ru/
Научная электронная библиотека	https://elibrary.ru/defaultx.asp
Российская государственная библиотека	https://www.rsl.ru/
Система дистанционного обучения УГНТУ, учебный курс преподавателя	http://do.rusoil.net
Университетская библиотека онлайн	http://biblioclub.ru/
Учебный курс преподавателя в СДО УГНТУ	http://do.rusoil.net
ЭБС Znanium.com	http://znanium.com/
ЭБС Лань	https://e.lanbook.com/
ЭБС Национальная электронная библиотека	https://rusneb.ru/
ЭБС «Университетская библиотека онлайн»	https://biblioclub.ru/
Электронная библиотека УГНТУ	http://www.bibl.rusoil.net

7. Материально-техническое обеспечение дисциплины

7.1. Перечень специальных аудиторий, кабинетов, лабораторий и пр., используемых при реализации дисциплины с перечнем основного оборудования

№ пп.	Номер помещения	Оснащенность помещения (перечень основного оборудования)	Наименование помещения
1	1-420в	Компьютер Intel Core 2 Duo E8200(1);Компьютер WIN i3-550(2);Компьютер персональный i3-4170/21,5" PHILIPS 226V4LAB(1);МФУ hp Laser Jet Pro M1536dnf<CE538A>A4(1);Монитор 19" Acer(1);Монитор ASUS VA24DQ Black 23,8", шт(3);Сервисное устройство для очистки Katun 3 м(1);Системный блок CRONA:i7 11700/DDR4 2*16GB/SSD 250Gb Samsung/HDD 1Tb/600W(3);Шкаф(ы) для хранения	Помещения для хранения и профилактического обслуживания учебного оборудования
2	1-432	Компьютер Nettop Pegatron Walle L6 PV D-SUB(15);Компьютер тип K2 i3-3220/21,5" LG 22EA63TP(1);Проектор мультимедийный Sony VPL-CX150(1);Системный блок UNIVERSAL D1(18);Стол, стулья	Учебная аудитория для проведения групповых и индивидуальных консультаций
3	1-435	Компьютер Pegatron Nettop MiniPC Wall-e L6(12);Монитор Samsung S-LC24F390FHIXCI(9);Монитор Samsung S-LC24FG73FQIXCI(5);Проектор Optoma EH334(1);Системный блок UNIVERSAL D1(14);Стол, стулья	Учебная аудитория для проведения занятий семинарского типа – укомплектована специализированной (учебной) мебелью, техническими средствами обучения.
4	1-441	Компьютер Pegatron Nettop MiniPC Wall-e L6 Pinetrail Atom D510(13);Стол, стулья	Учебная аудитория для текущего контроля и промежуточной аттестации – укомплектована специализированной (учебной) мебелью, техническими средствами обучения.
5	1-444	Компьютер Nettop Pegatron Walle L6 PV D-SUB(1);Настенный экран Master Picture 244x244 MW(1);Проектор Acer ProjectorP1203(1);мультимедиа-проектор;Учебно-наглядные пособия по дисциплине,набор демонстрационного оборудования; Столы, стулья;	Учебная аудитория для проведения занятий лекционного типа – укомплектована специализированной (учебной) мебелью, набором демонстрационного оборудования и учебно-наглядными пособиями, обеспечивающими тематические иллюстрации, соответствующие рабочим учебным программам дисциплин (модулей).
6	3-201	Защитная RFID Система LSG405HF(1);Компьютер i3-2120(1);Компьютер i3-3220 K1 BenQ 21,5"(3);Компьютер i3-3240 21.5" Acer(2);Компьютер ПК НИКС\i3-4170\21.5"(1);Компьютер персональный-неттоп Celeron J1900/4Gb(1);Контрольно-кассовая машина Пионер 114Ф с ФН(1);МФУ hp Laser Jet Pro M1132<CE847A>A4(1);МФУ hp LaserJet Pro M1132<CE847A>(A4 13	Помещение для самостоятельной работы – укомплектовано специализированной (учебной) мебелью, оснащено компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечено доступом в электронную информационно-образовательную среду организации.

7.2. Перечень лицензионного и свободно распространяемого программного обеспечения, используемых в учебном процессе при освоении дисциплины

№ пп.	Наименование ПО	Лицензионная чистота (реквизиты лицензии, свидетельства о гос. регистрации и т.п., срок действия)
1	Python	Дата выдачи лицензии 01.01.2006, Поставщик: Свободное программное обеспечение
2	Python	Дата выдачи лицензии 01.01.1991, Поставщик: Свободное программное обеспечение
3	Python v 3.7.1	Дата выдачи лицензии 01.01.2000, Поставщик: Свободное программное обеспечение
4	Python 3	Дата выдачи лицензии 01.01.2006, Поставщик: Свободное программное обеспечение
5	Pyton OnLine	Дата выдачи лицензии 01.01.2006, Поставщик: Свободное программное обеспечение
6	R открытая программная среда для статистических вычислений и графики	Дата выдачи лицензии 01.01.2006, Поставщик: Свободное программное обеспечение

8. Организация обучения лиц с ограниченными возможностями здоровья

Для лиц с ограниченными возможностями здоровья, обучающихся по данной образовательной программе, разрабатывается индивидуальная программа освоения дисциплины с учетом особенностей их психофизического развития, их индивидуальных возможностей и состояния здоровья.

Приложение А

Форма № УЛ-1

СВЕДЕНИЯ

об обеспеченности дисциплины основной и дополнительной учебной литературой

Наименование дисциплины: (48549)(48549)Статистические и вероятностные методыНаправление подготовки (специальность): 09.03.01 Информатика и вычислительная техникаНаправленность: профиль«Технологии искусственного интеллекта в нефтегазовой отрасли»Форма обучения: очная;Кафедра, обеспечивающая преподавание дисциплины: Вычислительная техника и инженерная кибернетика (ВТИК);

Тип	Назначение учебных изданий	Семестр			Библиографическое описание	Кол-во экз.	Адрес нахождения электронного	Кoeffициент
		очная	очно-заочная	заочная				
1	2	3	4	5	6	7	8	9
Основная литература	Для выполнения СРО;Для выполнения практических занятий;Для изучения теории;	4			Гмурман, В. Е. Теория вероятностей и математическая статистика : учебное пособие / В. Е. Гмурман. - 12-е изд. - М. : Юрайт, 2013. - 479 с. - URL: http://bibl.rusoil.net/base_docs/UGNTU/CD/Gmurman1.pdf . - Текст : электронный.	1	http://bibl.rusoil.net	1.00

Основная литература	Для выполнения СРО; Для выполнения практических занятий;	4			Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике : учебное пособие / В. Е. Гмурман. - 11-е изд., перераб. и доп. - М. : Юрайт, 2013. - 404 с. - URL: http://bibl.rusoil.net/base_docs/UGNTU/CD/Gmurman.pdf . - Текст : электронный.	1	http://bibl.rusoil.net	1.00
Дополнительная литература	Для выполнения СРО; Для выполнения практических занятий; Для изучения теории;	4			Гусаров, В. М. Статистика : учебное пособие : [16+] / В. М. Гусаров, Е. И. Кузнецова. – 2-е изд., перераб. и доп. – Москва : Юнити, 2012. – 480 с. – URL: https://biblioclub.ru/index.php?page=book&id=117190 (дата обращения: 22.04.2022). – Текст : электронный.	1	https://biblioclub.ru	1.00
Примечание – Графы 1-5,8 заполняются кафедрой, графы 7 и 9 - библиотекой								

Составил:

_____ старший преподаватель Шварева Е.Н.

Год приема 2023 г.

СВЕДЕНИЯ**об обеспеченности дисциплины учебно-методическими изданиями**

Наименование дисциплины: (48549)(48549)Статистические и вероятностные методы

Направление подготовки (специальность): 09.03.01 Информатика и вычислительная техника

Направленность профиль«Технологии искусственного интеллекта в нефтегазовой отрасли»

Форма обучения очная;

Кафедра, обеспечивающая преподавание дисциплины: Вычислительная техника и инженерная кибернетика (ВТИК);

Назначение учебных изданий	Семестр			Библиографическое описание	Кол-во экз.		Адрес нахождения электронного	Коэффициент
	очная	очно-заочная	заочная		Всего	в том числе		
1	2	3	4	5	6	7	8	9

Для выполнения СРО;Для выполнения практических занятий;	4			Учебно-методический комплекс дисциплины "Математика" : учебно-методический комплекс / Р. Н. Бахтин [и др.] ; УГНТУ, ИАУ, каф. Математики. - Уфа : Изд-во УГНТУ. - URL: http://bibl.rusoil.net/base_docs/UGNTU\Matematika/14UMK2010.PDF . - Текст : электронный. Раздел 14 : Математическая статистика : теоретические основы; метод. указания для студентов; материалы для самостоятельной работы студентов. - 2010.	1	0	http://bibl.rusoil.net	1.00
Для выполнения СРО;Для выполнения практических занятий;	4			Статистика на языке программирования R : учебно-методическое пособие для лабораторных, практических и самостоятельных работ / УГНТУ, каф. Математики ; сост.: С. Н. Коледин [и др.]. - Уфа : УГНТУ, 2018. - 1,24 Мб. - URL: http://bibl.rusoil.net/base_docs/UGNTU/Matematika/Koledin5.pdf . - Текст : электронный.	1	0	http://bibl.rusoil.net	1.00
Примечание – Графы 1-5,8 заполняются кафедрой, графы 6,7 и 9 - библиотекой								

Составил:

_____ старший преподаватель Шварева Е.Н.

—
Год приема 2023 г.

Приложение Б

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Уфимский государственный нефтяной технический университет»

УТВЕРЖДАЮ

Директор IT-институт

_____ 3.Х. Павлова
29.06.2023

**Фонд оценочных средств по текущей успеваемости и промежуточной аттестации по дисциплине
(48549)Статистические и вероятностные методы**

Направление подготовки (специальность): 09.03.01 Информатика и вычислительная техника

Направленность: профиль «Технологии искусственного интеллекта в нефтегазовой отрасли»

Уровень высшего образования: бакалавриат

Форма обучения: очная

Кафедра, обеспечивающая преподавание дисциплины: Вычислительная техника и инженерная кибернетика (ВТИК);

Трудоемкость дисциплины: 3 з.е. (108час)

Уфа 2023

ФОС по текущей успеваемости и промежуточной аттестации по дисциплине разработал (и):

_____ старший преподаватель Шварева Е.Н.

—
Рецензент

_____ доцент, кандидат ф.-м.н. Зарипов Д. М.

—

ФОС по текущей успеваемости и промежуточной аттестации по дисциплине рассмотрен и одобрен на заседании кафедры Вычислительная техника и инженерная кибернетика (ВТИК);, обеспечивающей преподавание дисциплины 31.08.2022, протокол №1.

Заведующий кафедрой Вычислительная техника и инженерная кибернетика (ВТИК);_____..

СОГЛАСОВАНО

Заведующий кафедрой ВТИК _____..

Год приема 2023 г.

ФОС по текущей успеваемости и промежуточной аттестации по дисциплине
зарегистрирован 19.09.2022 № 1 в отделе МСОП и внесен в электронную базу данных

1. Паспорт фонда оценочных средств

№ п/п	Контролируемые разделы (темы) дисциплины	Шифр результата обучения	Результат обучения	Индикатор достижения компетенций	Показатели достижения результатов освоения компетенций	Вид оценочного средства
1	Основы статистики	В(ПК-7и-22Г)	структурированность данных для машинного обучения	ПК-7.2. Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения	поиск выбросов и аномалий	Письменный и устный опрос
		З(ПК-7и-22Г)		ПК-7.2. Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения	представление данных для машинного обучения	Письменный и устный опрос
		У(ПК-7и-22Г)		ПК-7.2. Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения	восстановление данных, рекомендательные системы	Письменный и устный опрос Тестирование
4	Методы машинного обучения	В(ПК-1и-22 Г.)	задачи систем искусственного интеллекта	ПК 1.1 Классифицирует и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей	идентификацией задач систем искусственного интеллекта	Письменный и устный опрос
		В(ПК-4и-22Г)	метрики оценки результатов моделирования, критерий качества	ПК-4.1. Проводит анализ требований и определяет необходимые	определением классов задач МО	Письменный и устный

			построенной модели	классы задач машинного обучения		опрос
				ПК-4.2. Определяет метрики оценки результатов моделирования и критерии качества построенных моделей	оценкой результатов моделирования	Письменный и устный опрос
				ПК-4.3. Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения	выбором методов МО	Письменный и устный опрос
		З(ПК-1и-22 Г.)	задачи систем искусственного интеллекта	ПК 1.1 Классифицирует и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей	классификацию задач искусственного интеллекта	Письменный и устный опрос
		З(ПК-4и-22Г)	метрики оценки результатов моделирования, критерий качества построенной модели	ПК-4.1. Проводит анализ требований и определяет необходимые классы задач машинного обучения	требования к задачам МО	Письменный и устный опрос
				ПК-4.2. Определяет метрики оценки результатов моделирования и критерии качества построенных моделей	критерии качества моделей МО	Письменный и устный опрос
				ПК-4.3. Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения	оценку методов МО	Письменный и устный опрос
		У(ПК-1и-22	задачи систем искус-	ПК 1.1 Классифицирует	выбирать средства ре-	Письмен-

		Г.)	ственного интеллекта	и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей	шения задач искусственного интеллекта	ный и устный опрос Тестирование
		У(ПК-4и-22Г)	метрики оценки результатов моделирования, критерий качества построенной модели	ПК-4.1. Проводит анализ требований и определяет необходимые классы задач машинного обучения	анализировать задачи МО	Письменный и устный опрос Тестирование
				ПК-4.2. Определяет метрики оценки результатов моделирования и критерии качества построенных моделей	применять методы МО	Письменный и устный опрос Тестирование
				ПК-4.3. Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения	разрабатывать методы МО	Письменный и устный опрос Тестирование

2. Перечень оценочных средств для текущего контроля и промежуточной аттестации по дисциплине

п/п	Вид оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в фонде	Шкала оценки
1	2	3	4	5
1	Письменный и устный опрос	Оценочное средство для текущего контроля успеваемости и	Совокупность вопросов, заданий, упражнений, тестов	оценка «отлично» выставляется обучающемуся, если обучающийся правильно ответил на теоретические вопросы; пока-

		промежуточной аттестации. Позволяет выявить и восполнить пробелы в знаниях; повторить, закрепить, систематизировать материал; оценить знания, умения, теоретические и практические навыки; определить уровень сформированных у студентов компетенций по дисциплине (модулю)	для выполнения контрольных работ, домашних заданий, РГР и иных учебных работ. Комплект билетов для текущей и промежуточной аттестации	зал отличные знания в рамках усвоенного учебного материала; ответил на все дополнительные вопросы; решены все задачи; оценка «<i>хорошо</i>» выставляется обучающемуся, если обучающийся ответил на теоретические вопросы с небольшими неточностями; показал хорошие знания в рамках усвоенного учебного материала; ответил на большинство дополнительных вопросов; оценка «<i>удовлетворительно</i>» выставляется обучающемуся, если обучающийся ответил на теоретические вопросы с существенными неточностями; показал удовлетворительные знания в рамках усвоенного учебного материала; при ответах на дополнительные вопросы было допущено несколько неправильных ответов; оценка «<i>неудовлетворительно</i>» выставляется обучающемуся, если при ответе на теоретические вопросы обучающийся продемонстрировал недостаточный уровень знаний; на дополнительные вопросы отвечает неправильно;
2	Тестирование	Система стандартизированных простых и комплексных заданий, позволяющая автоматизировать процедуру измерения уровня знаний, умений и владений обучающегося.	Фонд тестовых заданий.	оценка «<i>отлично</i>» выставляется обучающемуся, если обучающийся дал правильные ответы на 90...100 % вопросов; оценка «<i>хорошо</i>» выставляется обучающемуся, если обучающийся дал правильные ответы на 75...89 % вопросов; оценка «<i>удовлетворительно</i>» выставляется обучающемуся, если обучающийся дал правильные ответы на 60...74 % вопросов; оценка «<i>неудовлетворительно</i>» выставляется обучающемуся, если обучающийся дал правильные ответы на 0...59 % вопросов;

Приложение В

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Уфимский государственный нефтяной технический университет»

Письменный и устный опрос.

Перечень вопросов (задач, заданий, тем, комплекта тестовых заданий):

Вопросы:

Вопросы

Меры центральной тенденции

Меры изменчивости

Квантили распределения

Нормальное распределение

Центральная предельная теорема.

Доверительный интервал

Идея статистического вывода

α -уровень значимости

Распределение Стьюдента

F-критерий Фишера

Множественное попарное сравнение

Поправка Бонферрони

q-критерий Тьюки

Коэффициент ковариации

Коэффициент корреляции

Коэффициент детерминации

Гомоскедастичность

U-критерий Манна-Уитни

Критерий Краскела-Уоллиса

Эволюционные алгоритмы оптимизации

Популяционные алгоритмы оптимизации

Стандартизация (Z-преобразование)

Критерий Манна-Уитни.

Тест Шапиро-Уилк

Однофакторный дисперсионный анализ

Двухфакторный дисперсионный анализ

Корреляция и регрессия

Одномерный регрессионный анализ

Многофакторный регрессионный анализ

Анализ номинативных данных

Логистическая регрессия

Модель с одним номинативным предиктором

Модель с двумя номинативными предикторами

Смешанная регрессионная модель

Метаоптимизация

Тестирование.

Перечень вопросов (задач, заданий, тем, комплекта тестовых заданий):

1. Рассчитайте среднеквадратическое отклонение данных выборочных значений: 1 5 2 7 1 9 3 8 5 9
 - 3,16*
 - 1,14
 - 10,12
 - 1,53

2. Как соотносятся дисперсии двух выборок
 (1)? 1 3 5 6 6 7 9 11
 (2)? 5 7 9 10 10 11 13 15
 - $D2 = D1 * 4$
 - $D2 = D1 *$
 - $D2 = D1 + 4$

3. Как соотносятся стандартные (среднеквадратические) отклонения двух выборок:
 (1)? 2 4 5 8 9 10 14 16
 (2)? 6 12 15 24 27 30 42 48
 - $SD2 = SD1$
 - $SD2 = SD1 * 3 *$
 - $SD2 = SD1 / 3$
 - $SD2 = SD1 * 9$

4. Измерили рост 1000 человек, данное распределение оказалось нормальным со средним равным 175 и стандартным отклонением равным 8 ($M = 175, sd = 8$). Рост одного из испытуемых составил 186,2. Чему равняется z - значение, рассчитанное для этого испытуемого?
 - Недостаточно данных, чтоб ответить на этот вопрос.
 - 1,4*
 - - 1,4
 - 0

5. Если отдельное наблюдение меньше, чем выборочное среднее, то соответствующее z - значение будет:
 - Отрицательным*
 - Положительным
 - Положительным
 - Нулем

6. Считается, что значение IQ (уровень интеллекта) у людей имеет нормальное распределение со средним значением равным 100 и стандартным отклонением равным 15 ($M = 100, sd = 15$). Какой приблизительно процент людей обладает $IQ > 125$?
 - 15%
 - 10%
 - 5%*
 - 25%

7. Считается, что значение IQ (уровень интеллекта) у людей имеет нормальное распределение со средним значением равным 100 и стандартным отклонением равным 15 ($M = 100, sd = 15$). Какой приблизительно процент людей обладает IQ на промежутке от 70 до 112
 - 92%
 - 60%
 - 77%*
 - 85%

8. В симметричной выборке...
- разница между медианой и средним значением близка к нулю*
 - отклонение близко к нулю
 - среднее значение близко к нулю
9. Первая и третья квартиль значений признака Р равны 2, 4 соответственно. Какие из следующих значений будут считаться выбросами?
- 6,5
3
0
0,5
-1,5*
8*
7,5*
10. Среднее значение, отклонение и медиана десяти значений признака Р равны 10, 1.1 (одна целая одна десятая), и 9 соответственно. Какие из следующих значений будут выбросами? Не забудьте в процессе решения проверить симметричность выборки.
- 7
6,5*
14*
14
13
6*
11. среднее значение выборки некоторого признака Р равно 50. Для этой выборки был применен критерий Шовене поиска выбросов. Оказалось, что элементы выборки равные 42 и 57 были признаны выбросами. Какие из указанных ниже чисел критерий Шовене гарантированно определит как выбросы?
- 56
58*
44
41*
12. Если объем выборки достаточно велик (больше 100), то такая выборка является репрезентативной.
- Да
 - Зависит от способа формирования выборки*
 - Нет
13. В каких случаях вместо среднего значения лучше использовать моду или медиану в качестве центральной тенденции?
- Если распределение является симметричным и унимодальным
Если присутствуют заметные выбросы*
Если распределение асимметрично*
14. Укажите, в какой из выборок наибольшее стандартное отклонение:
- (1) 1 3 2 4 5 7 1 8
(2) 100 300 250 400 230 280 320 112
(3) 15 10 13 7 28 31 20 32
- 3
 - 2*
 - 1
15. Выберите верные утверждения (одно или несколько):

- Для каждого наблюдения в выборке можно рассчитать соответствующее z значение*
 - Z значение может быть рассчитано без знания стандартного отклонения по выборке
 - Если все выборочные значения положительные, то и после z преобразования все значения останутся положительными.
 - Z значение не может быть больше трех
1. Как соотносятся средние значения двух рядов чисел:
 - 1) 1 3 1 3 7 8 9 10 12 12 13 18 20 19
 - 2) 2 6 2 6 14 16 18 20 24 24 26 36 40 38
 - Средние значения равны
 - Среднее второго ряда чисел в два раза больше*
 - Среднее второго ряда чисел в два раза меньше
 2. Может ли показатель стандартного отклонения принимать отрицательные значения?
 - Может, если все значения в выборке равны друг другу.
 - Не может, стандартное отклонение всегда равно нулю.
 - Не может, стандартное отклонение всегда неотрицательное.*
 - Может, если все значений выборки отрицательные.
 3. Если отдельное наблюдение в выборке равняется 1000, при условии, что выборочное среднее равняется 10, то что можно сказать про такое наблюдение:
 - Такое наблюдение в принципе не может принадлежать выборке со средним значением равным 10, так как в 100 раз больше, чем выборочное среднее, а вероятность такого события стремится к нулю
 - Чтобы судить о том, насколько необычным является это наблюдение, необходимо знать, чему равняется стандартное отклонение.*
 - Можно рассматривать как необычное (выброс), т.к. оно очень далеко отклоняется от среднего значения
 4. некоторый признак распределен нормально, выборочное среднее равняется 100, а дисперсия равняется 25 ($M = 100$, $D = 25$). Тогда приблизительно 95% всех наблюдений находится в диапазоне:
 - от 90 до 110*
 - от 50 до 100
 - от 50 до 150
 5. Если увеличить размер выборки, то сильные отклонения выборочных средних от истинного среднего как будут возникать
 - Реже, распределение выборочных средних станет более узким*
 - Чаще, распределение выборочных средних станет более широким
 - Это не повлияет на характер распределения выборочных средних
 6. Как соотносятся стандартная ошибка среднего и выборочное стандартное отклонение исследуемого признака (при размере выборки $n > 1$)?
 - Стандартная ошибка всегда равняется стандартному отклонению
 - Стандартная ошибка всегда больше, чем стандартное отклонение
 - Стандартная ошибка всегда меньше, чем стандартное отклонение*
 7. Если рассчитали 95% доверительный интервал для среднего значения, то какой вывод можно сделать?

- Среднее значение в генеральной совокупности точно превышает нижнюю границу 95% доверительного интервала.
- Среднее значение в генеральной совокупности точно принадлежит рассчитанному доверительному интервалу.
- Если многократно повторять эксперимент, то 95 % выборочных средних значений будут принадлежать рассчитанному нами доверительному интервалу.
- Если многократно повторять эксперимент, для каждой выборки рассчитывать свой доверительный интервал, то в 95 % случаев истинное среднее будет находиться внутри доверительного интервала.*
- Мы можем быть на 95% уверены, что среднее значение в генеральной совокупности принадлежит рассчитанному доверительному интервалу.*

8. Если бы в нашем примере мы увеличили объем выборки в два раза (при условии, что показатель стандартного отклонения остался неизменным), то 95% доверительный интервал каким будет?

- стал более широким
- стал более узким*
- возможны оба варианта

9. В центре 95% доверительного интервала, рассчитанного по выборочным значениям, что находится

- Среднее значение генеральной совокупности
- Значение стандартной ошибки среднего
- Выборочное среднее значение*

10. Предположим, нулевой гипотезой исследования являлось предположение, что среднее в генеральной совокупности равняется 100. Получили $p = 0,12$ и не смогли отклонить нулевую гипотезу. Однако позже выяснилось, что среднее в генеральной совокупности равняется 114. Как можно описать результаты данного исследования?

- Вы совершили ошибку второго рода*
- Вы совершили ошибку первого рода
- Вы не совершали ни ошибку первого рода, ни ошибку второго рода.

11. В первом эксперименте для сравнения двух средних $X_1=17, X_2=16$ применялся t-критерий Стьюдента, и эти различия оказались значимы ($p = 0,001$). Во втором исследовании, также при помощи t - критерия, сравнивались два средних $X_1=17, X_2=36$, и эти различия не значимы при $p = 0,8$. В чем может быть причина таких результатов?

- Возможно, в первом эксперименте больше объем выборок и меньше изменчивость исследуемого признака*
- Ни размер выборки, ни изменчивость исследуемого признака не может быть причиной таких результатов
- Возможно, во втором эксперименте больше объем выборок и меньше изменчивость исследуемого признака

11. Исследовались различия между средними значениями числа рибосом двух различных типов бактерий. Что означает высказывание: "95% доверительный интервал для разности между средними значениями рибосом двух типов бактерий ?тип№1??тип№2 оказался равен от 100 до 300".

- Число рибосом у первого типа бактерий в среднем на 100-300 рибосом превышает число рибосом у второго типа.*

- Число рибосом у второго типа бактерий в среднем на 100-300 рибосом превышает число рибосом у первого типа.
 - Среднее значение числа рибосом у первого типа бактерий равняется 100, а у второго 300.
 - Эти данные не позволяют отклонить нулевую гипотезу о равенстве средних значений ?
type1 = ?type2.
12. Если для проверки нормальности распределения на выборке в 100 наблюдений мы применили критерий Shapiro-Wilk test и получили p-уровень значимости, равный 0.001, то что можно сказать про распределение?
- Распределение значимо отклоняется от нормального*
 - Распределение является бимодальным
 - Распределение можно считать нормальным
13. Если при сравнении четырех групп между собой (по 5 наблюдений в каждой группе) получили $F = 3.5$, тогда какой вывод можно сделать о гипотезе?
- Принимаем нулевую гипотезу о равенстве средних, т.к. $p < 0,05$
 - Отклоняем нулевую гипотезу о равенстве средних, т.к. $p < 0,05^*$
 - Принимаем нулевую гипотезу о равенстве средних, т.к. $p > 0,05$
 - Отклоняем нулевую гипотезу о равенстве средних, т.к. $p > 0,05$
14. Отклонение нулевой гипотезы какой вывод позволяет нам сделать ?
- Только две группы значимо различаются
 - Значимых различий обнаружить не удалось
 - Все группы значимо различаются между собой
 - Как минимум две группы значимо различаются*
15. Взаимодействие факторов в дисперсионном что означает?
- Только один фактор оказывает значимое влияние на зависимую переменную
 - Нет возможности отклонить нулевую гипотезу о равенстве средних значений
 - Влияние одного фактора по-разному проявляется на зависимую переменную в зависимости от градаций другого фактора
16. Причиной отклонения нулевой гипотезы в многофакторном дисперсионном анализе (о равенстве средних во всех группах) что может стать ?
- Значимое взаимодействие факторов
 - Значимый эффект одного из фактора
 - Значимый эффект для всех факторов
 - Каждый из предложенных вариантов
17. Могут ли очень заметные выбросы негативно повлиять на результаты дисперсионного анализа и почему?
- Нет, т. к. дисперсионный анализ сравнивает дисперсии групп, а не средние значения.
 - Да, но только в случае многофакторного анализа.
 - Да, так как могут нежелательно повлиять на значения средних межгрупповых и внутригрупповых квадратов (особенно в случае небольшого количества наблюдений).*

18. Какие ограничения желательно проверять при использовании дисперсионного анализа:
- Нормальное распределение зависимой переменной в группах, гомогенность (однородность) дисперсий
 - F-значение не должно превышать 30
 - Приблизительное равенство межгруппового и внутригруппового среднего квадрата
 - Количество групп (не больше трех), и гомогенность дисперсий (однородность)
19. Рассчитайте стандартную ошибку среднего, если выборочное среднее равняется 10, дисперсия 4, при $N = 100$
- 0.2
20. Даны значения признака Р для шести объектов выборки $P = (-1, 0, 4, 2, 1, 2)$ Чему равно среднее значение признака Р (ответ округлить до двух знаков после запятой).
- 1,33

Аннотация к рабочей программе дисциплины (48549)Статистические и вероятностные методы

Направление подготовки (специальность): 09.03.01 Информатика и вычислительная техника

Направленность: профиль«Технологии искусственного интеллекта в нефтегазовой отрасли»

Уровень высшего образования: бакалавриат

Форма обучения: очная;

Кафедра, обеспечивающая преподавание дисциплины: Вычислительная техника и инженерная кибернетика (ВТИК);

Компетенции, формируемые в результате освоения дисциплины

ПК-1и-22 Г. Способен классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта:

-ПК 1.1 Классифицирует и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей

ПК-4и-22Г Способен разрабатывать и применять методы машинного обучения для решения задач:

-ПК-4.1. Проводит анализ требований и определяет необходимые классы задач машинного обучения

-ПК-4.2. Определяет метрики оценки результатов моделирования и критерии качества построенных моделей

-ПК-4.3. Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения

ПК-7и-22Г Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта:

-ПК-7.2. Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения

Результат обучения

Знать:

ПК-1и-22 Г.-3 задачи систем искусственного интеллекта

ПК-4и-22Г-1 метрики оценки результатов моделирования, критерий качества построенной модели

ПК-7и-22Г-1 структурированность данных для машинного обучения

Уметь:

ПК-1и-22 Г.-3 умеет классифицировать задачи искусственного интеллекта

ПК-4и-22Г-1 оценивать, выбирать, разрабатывать методы машинного обучения

ПК-7и-22Г-1 размечать данные для машинного обучения

Владеть:

ПК-1и-22 Г.-3 владеет идентификацией задач искусственного интеллекта

ПК-4и-22Г-1 анализом требований задач машинного обучения

ПК-7и-22Г-1 подготовкой данных для машинного обучения

Краткая характеристика дисциплины

Основы статистики; Методы машинного обучения;

Трудоёмкость (з.е. / часы)

3 з.е. (108час)

диф.зачет;

Вид промежуточной аттестации

Разработчик(и):

_____ старший преподаватель Шварева Е.Н.

СОГЛАСОВАНО

Заведующий кафедрой ВТИК _____..