МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ГЕОДЕЗИИ И КАРТОГРАФИИ
(МИИГАиК)

Факультет геоинформатики и информационной безопасности
Тюменский государственный университет

**ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ**

Сертификат: 5BFEE1D286E9C95A0C0D2C611E434BD5
Владелец: Степанченко Алексей Леонидович
Действителен: с 23.09.2026 по 17.12.2027

«УТВЕРЖДАЮ»
Проректор
А. Л. Степанченко
24 марта 2025 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ) «Б1.О.07 МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА»

Уровень высшего образования: магистратура

Направление подготовки: 09.04.03 Прикладная информатика

Направленность (профиль) подготовки: Сквозные цифровые технологии в управлении недвижимостью

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Год набора (приема на обучение): 2025

Срок получения образования: 1 года 10 месяца(-ев)

Объем: в зачетных единицах: 9 з.е.
в академических часах: 324 ак.ч.

Москва, 2025

Разработчики:
Профессор академического департамента, доктор физико-математических наук Шевляков А. Н.

Рабочая программа дисциплины (модуля) составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.04.03 Прикладная информатика, утвержденного приказом Минобрнауки от 19.09.2017 № 916

Согласование и утверждение

№	Подразделение	Ответственное лицо	ФИО	Виза	Дата, протокол (при наличии)
1	факультет Геоинформатики и информационной безопасности	Председатель ученого совета факультета	Матерухин А. В.	Рассмотрено	14.10.2024, № 10-24

Содержание

1. Цель и задачи освоения дисциплины (модуля)

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

3. Место дисциплины в структуре ОП

4. Объем дисциплины и виды учебной работы

5. Содержание дисциплины
   5.1. Разделы, темы дисциплины и виды занятий
   5.2. Содержание разделов, тем дисциплины и формы текущего контроля

6. Оценочные материалы текущего контроля

7. Оценочные материалы промежуточной аттестации

8. Порядок проведения промежуточной аттестации

9. Материально-техническое и учебно-методическое обеспечение дисциплины
   9.1. Перечень основной и дополнительной учебной литературы
   9.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся
   9.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине
   9.4. Оборудование и технические средства обучения

10. Методические указания по освоению дисциплины (модуля)

1. Цель и задачи освоения дисциплины (модуля)

Цель освоения дисциплины - приобретение обучающимися знаний в области математического аппарата, используемого при проектировании и разработке систем искусственного интеллекта, а также методологии применения математических моделей и алгоритмов машинного обучения для анализа данных и извлечения знаний.

Задачи изучения дисциплины:

- изучить фундаментальные математические концепции, лежащие в основе современных методов искусственного интеллекта;
- сформировать умения формализовать прикладные задачи в терминах математических моделей машинного обучения и выбирать адекватные методы их решения;
- развить навыки применения изученного математического аппарата для построения и оценки алгоритмов искусственного интеллекта с использованием современных инструментальных средств и библиотек.

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

Компетенции, индикаторы и результаты обучения

ОПК-1 Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте

ОПК-1.1 Умеет самостоятельно приобретать, развивать и применять естественнонаучные и профессиональные знания для решения нестандартных задач

Знать:

ОПК-1.1/Зн1 Методологию моделирования физических процессов и явлений при разработке программного обеспечения

Уметь:

ОПК-1.1/Ум1 Интерпретировать результаты вычислений и симуляций с точки зрения их соответствия фундаментальным естественнонаучным законам

Владеть:

ОПК-1.1/Нв1 Навыками поиска междисциплинарных связей между информатикой и естественными науками для генерации инновационных решений

ОПК-1.2 Умеет самостоятельно приобретать, развивать и применять математические знания для решения нестандартных задач

Знать:

ОПК-1.2/Зн1 Методы формализации нестандартной прикладной задачи на математическом языке

Уметь:

ОПК-1.2/Ум1 Адаптировать классические математические методы под нетривиальные условия задачи

Владеть:

ОПК-1.2/Нв1 Навыками работы с математическими пакетами и библиотеками

ОПК-1.3 Умеет самостоятельно приобретать, развивать и применять социально-экономические знания для решения нестандартных задач

Знать:

ОПК-1.3/Зн1 Методологию оценки социальных и экономических рисков внедрения технологических решений

Уметь:

ОПК-1.3/Ум1 Применять методы социально-экономической диагностики для выявления скрытых потребностей пользователей

Владеть:

ОПК-1.3/Нв1 Навыками проведения технико-экономического анализа предлагаемых решений с использованием не только технических, но и социально-экономических критериев

ОПК-1.4 Самостоятельно находит, анализирует и адаптирует новые источники знаний (научные, нормативные, профессиональные) для решения задач в незнакомой предметной области.

Знать:

ОПК-1.4/Зн1 Методы поиска актуальных источников знаний

Уметь:

ОПК-1.4/Ум1 Анализировать новые источники знаний для решения задач в незнакомой предметной области

Владеть:

ОПК-1.4/Нв1 Навыками анализа научной литературы

ОПК-1.5 Использует междисциплинарные методы и инструменты для разработки решений нестандартных или комплексных проблем

Знать:

ОПК-1.5/Зн1 Методы и инструменты для разработки программных продуктов

Уметь:

ОПК-1.5/Ум1 Использовать междисциплинарные методы и инструменты для разработки решений нестандартных или комплексных проблем

Владеть:

ОПК-1.5/Нв1 Навыками разработки программного обеспечения

ОПК-1.6 Создает и обосновывает оригинальные решения в условиях неопределенности или новой профессиональной среды, интегрируя ранее усвоенные знания.

Знать:

ОПК-1.6/Зн1 Методы анализа проектной ситуации

Уметь:

ОПК-1.6/Ум1 Обосновывать оригинальные решения в условиях неопределенности или новой профессиональной среды

Владеть:

ОПК-1.6/Нв1 Навыками анализа текущий условий разработки проектного решения

ОПК-7 Способен использовать методы научных исследований и математического моделирования в области проектирования и управления информационными системами

ОПК-7.1 Умеет использовать методы научных исследований и математического моделирования для проектирования и управления информационными системами и решения задач по их управлению

Знать:

ОПК-7.1/Зн1 Математические методы анализа и синтеза информационных систем, используемые в задачах искусственного интеллекта

ОПК-7.1/Зн2 Принципы построения математических моделей, описывающих структуру и поведение информационных систем и процессов управления ими

Уметь:

ОПК-7.1/Ум1 Применять методы математического моделирования для формализации требований к информационным системам, содержащим компоненты ИИ

ОПК-7.1/Ум2 Использовать математический аппарат для оценки эффективности проектных решений и управления параметрами информационных систем

Владеть:

ОПК-7.1/Нв1 Навыками применения методов научных исследований (анализ, сравнение, классификация) для выбора оптимальной математической модели при проектировании систем ИИ

ОПК-7.1/Нв2 Инструментарием математического моделирования для решения прикладных задач по управлению информационными системами

ОПК-7.2 Применяет математические модели в области проектирования и управления информационными системами

Знать:

ОПК-7.2/Зн1 Математические модели в области проектирования и управления информационными системами

Уметь:

ОПК-7.2/Ум1 Применять математические модели в области проектирования и управления информационными системами

Владеть:

ОПК-7.2/Нв1 Навыками построения математических моделей в области проектирования и управления информационными системами

ОПК-7.3 Использует методы научных исследований для обоснования и оценки эффективности проектных решений

Знать:

ОПК-7.3/Зн1 Методы научных исследований для обоснования и оценки эффективности проектных решений

Уметь:

ОПК-7.3/Ум1 Обосновывать проектные решения

Владеть:

ОПК-7.3/Нв1 Навыками оценки эффективности проектных решений

3. Место дисциплины в структуре ОП

Дисциплина (модуль) Б1.О.07 «Математические основы искусственного интеллекта» относится к обязательной части образовательной программы и изучается в семестре(ах): 2.

В процессе изучения дисциплины студент готовится к решению типов задач профессиональной деятельности, предусмотренных ФГОС ВО и образовательной программой.

4. Объем дисциплины (модуля) и виды учебной работы

Период	доемкость сы)	доемкость ЗЕТ)	ые занятия сы)	кие занятия сы)	ьная работа сы)	ая аттестация сы)

обучения	Общая тру (ча	Общая тру (ЗЕ	Лекционни (ча	Практическ (ча	Самостоятел (ча	Промежуточн (ча
2 семестр	324	9	16	32	258	Экзамен (18)
Всего	324	9	16	32	258	18

5. Содержание дисциплины (модуля)

5.1. Разделы, темы дисциплины и виды занятий

(часы промежуточной аттестации не указываются)

Наименование раздела, темы	Всего	Лекционные занятия	Практические занятия	Самостоятельная работа	Планируемые результаты обучения, соотнесенные с результатами освоения программы
Раздел 1. Основные понятия теории вероятностей	**38**	**2**	**4**	**32**	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3
Тема 1.1. Вероятностное пространство и аксиоматика как основа моделирования неопределенности в задачах ИИ	38	2	4	32	
Раздел 2. Дискретные случайные величины	**38**	**2**	**4**	**32**	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3
Тема 2.1. Распределения дискретных случайных величин в задачах классификации и обработки текстов	38	2	4	32	
Раздел 3. Непрерывные случайные величины	**38**	**2**	**4**	**32**	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3
Тема 3.1. Непрерывные распределения и их роль в генерации признаков и обучении нейронных сетей	38	2	4	32	
Раздел 4. Система нескольких случайных величин	**38**	**2**	**4**	**32**	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5

Тема 4.1. Многомерные случайные величины и понятие ковариации в анализе признаков	38	2	4	32	ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3
Раздел 5. Основные понятия статистики	**38**	**2**	**4**	**32**	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3
Тема 5.1. Генеральная совокупность и выборка как основа data-инжиниринга	38	2	4	32	
Раздел 6. Оценка параметров генеральной совокупности	**38**	**2**	**4**	**32**	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3
Тема 6.1. Методы оценивания параметров и доверительные интервалы для контроля качества моделей	38	2	4	32	
Раздел 7. Проверка статистических гипотез	**38**	**2**	**4**	**32**	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3
Тема 7.1. Статистические критерии и принятие решений в условиях неопределенности при A/B-тестировании	38	2	4	32	
Раздел 8. Анализ статистических связей	**40**	**2**	**4**	**34**	ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3
Тема 8.1. Корреляционный и регрессионный анализ как основа прогнозирования в информационных системах	40	2	4	34	
Итого	**306**	**16**	**32**	**258**	

5.2. Содержание разделов, тем дисциплин и формы текущего контроля

Раздел 1. Основные понятия теории вероятностей

Тема 1.1. Вероятностное пространство и аксиоматика как основа моделирования неопределенности в задачах ИИ

Понятие пространства элементарных исходов, события и вероятности как меры. Аксиомы Колмогорова. Классическое, статистическое и геометрическое определения вероятности. Правила сложения и умножения вероятностей. Условная вероятность и понятие независимости событий.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 2. Дискретные случайные величины

Тема 2.1. Распределения дискретных случайных величин в задачах классификации и обработки текстов

Определение дискретной случайной величины (ДСВ). Ряд распределения. Функция вероятности. Числовые характеристики: математическое ожидание, дисперсия, среднеквадратическое отклонение. Биномиальное распределение (моделирование количества успехов) и распределение Бернулли. Распределение Пуассона (моделирование редких событий).

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 3. Непрерывные случайные величины

Тема 3.1. Непрерывные распределения и их роль в генерации признаков и обучении нейронных сетей

Определение непрерывной случайной величины (НСВ). Плотность распределения вероятностей и интегральная функция распределения. Числовые характеристики НСВ. Нормальное (Гауссовское) распределение и его фундаментальная роль (центральная предельная теорема). Равномерное распределение.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 4. Система нескольких случайных величин

Тема 4.1. Многомерные случайные величины и понятие ковариации в анализе признаков

Совместное распределение двух случайных величин. Ковариация и коэффициент корреляции Пирсона как мера линейной зависимости. Условные распределения. Регрессия — зависимость одной величины от другой.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 5. Основные понятия статистики

Тема 5.1. Генеральная совокупность и выборка как основа data-инжиниринга

Предмет математической статистики. Генеральная совокупность и выборка. Репрезентативность. Вариационные ряды. Точечные оценки (среднее арифметическое, выборочная дисперсия, мода, медиана). Понятие выборочного метода.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 6. Оценка параметров генеральной совокупности

Тема 6.1. Методы оценивания параметров и доверительные интервалы для контроля качества моделей

Понятие точечной и интервальной оценки. Свойства оценок (несмещенность, состоятельность, эффективность). Доверительные интервалы для среднего и дисперсии. Метод максимального правдоподобия (ММП) как фундаментальный подход к подбору параметров моделей.

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 7. Проверка статистических гипотез

Тема 7.1. Статистические критерии и принятие решений в условиях неопределенности при A/B-тестировании

Нулевая и альтернативная гипотеза. Ошибки первого и второго рода. Уровень значимости и мощность критерия. Параметрические и непараметрические критерии (t-критерий Стьюдента, критерий Хи-квадрат).

Текущий контроль

Вид (форма) контроля
Тестирование

Раздел 8. Анализ статистических связей

Тема 8.1. Корреляционный и регрессионный анализ как основа прогнозирования в информационных системах

Функциональная, статистическая и корреляционная связи. Коэффициенты корреляции (Пирсона для линейной связи, Спирмена для монотонной). Построение линейной регрессионной модели (парная регрессия). Оценка качества регрессионной модели.

Текущий контроль

Вид (форма) контроля
Тестирование

6. Порядок проведения промежуточной аттестации

Порядок проведения промежуточной аттестации регламентируется локальными нормативными актами Университета:
- Положением о формах, периодичности, порядке текущего контроля успеваемости и промежуточной аттестации обучающихся по основным образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии», п.6 «Организация и проведение промежуточной аттестации».
- Положением о порядке осуществления образовательного процесса, проведения текущего контроля успеваемости и промежуточной аттестации обучающихся с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный университет геодезии и картографии», п.8 «Проведение промежуточной аттестации с применением ЭО и ДОТ».

Промежуточная аттестация: Экзамен, 2 семестр.
Экзамен оценивается по пятибалльной шкале:
- зачтено с оценкой «отлично» выставляется, если обучающийся имеет глубокие, исчерпывающие знания в объеме освоенной программы; знание основной (обязательной) литературы; дает правильные и уверенные ответы, свидетельствующие о наличии твердых знаний и навыков в использовании технических средств; полное, четкое, грамотное и логически стройное изложение материала; демонстрирует свободное применение теоретических знаний при анализе практических вопросов;
- зачтено с оценкой «хорошо» выставляется обучающемуся в случае соблюдения тех же требований, но наличии в ответе обучающегося по некоторым перечисленным показателям недостатков принципиального характера, что вызывает замечания или поправки преподавателя;
- зачтено с оценкой «удовлетворительно» выставляется обучающемуся в случае соблюдения тех же требований, но в случае наличия в ответе ошибок, необходимости существенной помощи в виде поправок и наводящих вопросов преподавателя;
- не зачтено («не удовлетворительно») – обучающийся не смог ответить на 2/3 вопросов; не справился с практическим заданием или выполнено менее 50 % практических заданий.
Вот 30 тестовых вопросов по курсу «Математические основы искусственного интеллекта» с вариантами ответов (цифры 1-4) и правильными ответами.

7. Оценочные материалы текущего контроля

Раздел 1. Основные понятия теории вероятностей
Контролируемые ИДК: ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3

Тема 1.1. Вероятностное пространство и аксиоматика как основа моделирования неопределенности в задачах ИИ

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Что понимается под вероятностью события в рамках аксиоматического определения Колмогорова?

1. Отношение числа благоприятных исходов к общему числу исходов
2. Числовая функция, определенная на множестве событий и удовлетворяющая аксиомам неотрицательности, нормированности и аддитивности
3. Предел относительной частоты появления события при бесконечном числе испытаний
4. Степень уверенности в наступлении события

2. Если события A и B независимы, то вероятность их одновременного наступления P(A∩B) равна:

1. P(A) + P(B)
2. P(A) * P(B|A)
3. P(A) * P(B)
4. P(A) + P(B) — P(A∪B)

3. Формула полной вероятности используется, когда:

1. Нужно найти вероятность события при условии, что другая гипотеза уже произошла
2. Событие может наступить только вместе с одной из несовместных гипотез, образующих полную группу
3. События являются независимыми
4. Известны только частоты появления событий

4. Теорема Байеса позволяет вычислить:

1. Вероятность гипотезы после наступления события (апостериорную вероятность)
2. Вероятность суммы событий
3. Дисперсию случайной величины
4. Математическое ожидание произведения величин

Раздел 2. Дискретные случайные величины

Контролируемые ИДК: ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3

Тема 2.1. Распределения дискретных случайных величин в задачах классификации и обработки текстов

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Какой закон распределения описывает количество «успехов» в последовательности n независимых испытаний с двумя исходами?

1. Распределение Пуассона
2. Равномерное распределение
3. Биномиальное распределение
4. Геометрическое распределение

2. Математическое ожидание дискретной случайной величины — это:

1. Наиболее вероятное значение величины
2. Среднее арифметическое всех возможных значений, взвешенное по их вероятностям
3. Корень квадратный из дисперсии
4. Разность между максимальным и минимальным значением

3. Для случайной величины, распределенной по Бернулли (принимает значения 1 с вероятностью p и 0 с вероятностью q=1-p), дисперсия равна:

1. p
2. p*q
3. p^2
4. 1-p

4. Распределение Пуассона обычно используется для моделирования:

1. Количества редких событий за фиксированный промежуток времени
2. Роста температуры в зависимости от времени
3. Распределения роста людей
4. Зависимости между двумя величинами

Раздел 3. Непрерывные случайные величины

Контролируемые ИДК: ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3

Тема 3.1. Непрерывные распределения и их роль в генерации признаков и обучении нейронных сетей

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Основное свойство плотности распределения вероятностей f(x):

1. f(x) не может быть больше 1
2. Интеграл от f(x) по всей числовой прямой равен 1
3. f(x) всегда является четной функцией
4. f(x) равна вероятности P(X=x)

2. Какой параметр нормального распределения $N(\mu, \sigma^2)$ отвечает за «ширину» колокола и разброс значений?

1. Математическое ожидание (μ)
2. Медиана
3. Дисперсия (σ^2) или среднеквадратичное отклонение (σ)
4. Коэффициент асимметрии

3. Центральная предельная теорема утверждает, что:

1. Все данные в природе распределены нормально
2. Сумма большого числа независимых случайных величин распределена приблизительно нормально
3. Дисперсия суммы равна сумме дисперсий
4. Нормальное распределение симметрично относительно нуля

4. В машинном обучении равномерное распределение часто используется для:

1. Инициализации весов нейронной сети
2. Нормализации признаков
3. Регуляризации моделей
4. Оценки качества классификации

Раздел 4. Система нескольких случайных величин

Контролируемые ИДК: ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3

Тема 4.1. Многомерные случайные величины и понятие ковариации в анализе признаков

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Ковариация Cov(X, Y) показывает:

1. Направление линейной зависимости между X и Y
2. Тесноту любой (в том числе нелинейной) связи
3. Дисперсию произведения X и Y
4. Отношение средних значений X и Y

2. Если коэффициент корреляции Пирсона r = 0, то это означает:

1. Отсутствие зависимости между переменными
2. Отсутствие линейной зависимости между переменными
3. Сильную обратную линейную зависимость
4. Переменные функционально зависимы

3. В задачах отбора признаков (feature selection) мы стремимся:

1. Оставить признаки с высокой корреляцией друг с другом
2. Оставить признаки, слабо коррелирующие с целевой переменной
3. Оставить признаки, сильно коррелирующие с целевой переменной и слабо — между собой
4. Удалить все признаки с нулевой корреляцией

4. Условное математическое ожидание M[X|Y] в контексте ИИ — это основа для:

1. Кластеризации данных

2. Задач регрессии (функция регрессии)
3. Понижения размерности
4. Детекции аномалий

Раздел 5. Основные понятия статистики

Контролируемые ИДК: ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3

Тема 5.1. Генеральная совокупность и выборка как основа data-инжиниринга

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Что такое репрезентативная выборка?

1. Выборка, содержащая все элементы генеральной совокупности
2. Выборка, максимально точно отражающая свойства и структуру генеральной совокупности
3. Выборка, состоящая только из средних значений
4. Случайная выборка большого объема

2. Какая характеристика относится к мерам положения (центральной тенденции) выборки?

1. Дисперсия
2. Медиана
3. Размах вариации
4. Коэффициент вариации

3. Разведочный анализ данных (EDA) в проектах машинного обучения начинается с:

1. Построения нейронной сети
2. Сбора дополнительных данных
3. Описательной статистики и визуализации распределений
4. Разделения выборки на train и test

Раздел 6. Оценка параметров генеральной совокупности

Контролируемые ИДК: ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3

Тема 6.1. Методы оценивания параметров и доверительные интервалы для контроля качества моделей

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Каким свойством обладает «хорошая» статистическая оценка?

1. Смещенность
2. Несостоятельность
3. Эффективность (минимальная дисперсия среди всех оценок)
4. Зависимость от объема выборки

2. Метод максимального правдоподобия (ММП) подбирает параметры модели таким образом, чтобы:

1. Минимизировать сумму квадратов ошибок
2. Максимизировать вероятность появления наблюдаемых данных
3. Минимизировать дисперсию оценки
4. Оценить доверительный интервал

3. Доверительный интервал для среднего показывает:

1. Диапазон, в который с заданной вероятностью попадает истинное значение параметра
2. Диапазон, в который попадают 95% всех значений выборки
3. Максимальное и минимальное значение признака
4. Точное значение параметра генеральной совокупности

4. В машинном обучении функция потерь (loss function) часто строится на основе:

1. Теоремы Байеса
2. Принципа максимума энтропии
3. Метода максимального правдоподобия
4. Правила сложения вероятностей

Раздел 7. Проверка статистических гипотез

Контролируемые ИДК: ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3

Тема 7.1. Статистические критерии и принятие решений в условиях неопределенности при A/B-тестировании

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Что такое «ошибка первого рода» (α) в проверке гипотез?

1. Принять нулевую гипотезу, когда она ложна
2. Отвергнуть нулевую гипотезу, когда она верна
3. Отвергнуть альтернативную гипотезу
4. Получить p-value = 0.05

2. Что означает p-value = 0.03 при проверке гипотезы (при уровне значимости α = 0.05)?

1. Нулевую гипотезу следует принять
2. Нулевую гипотезу следует отвергнуть
3. Данные собраны неверно
4. Необходимо увеличить выборку

3. Для сравнения средних двух выборок и проверки, значимо ли они различаются, обычно используется:

1. Критерий Стьюдента (t-критерий)
2. Критерий Пирсона (χ^2)
3. Критерий Фишера
4. Критерий Кохрена

4. A/B-тестирование в IT-компаниях — это практическое применение:

1. Теории оптимизации
2. Теории массового обслуживания
3. Проверки статистических гипотез
4. Регрессионного анализа

Раздел 8. Анализ статистических связей

Контролируемые ИДК: ОПК-1.1 ОПК-1.2 ОПК-1.3 ОПК-1.4 ОПК-1.5 ОПК-1.6 ОПК-7.1 ОПК-7.2 ОПК-7.3

Тема 8.1. Корреляционный и регрессионный анализ как основа прогнозирования в информационных системах

Форма контроля/оценочное средство: Тестирование

Вопросы/Задания:

1. Коэффициент детерминации R^2 показывает:

1. Тесноту связи между признаками
2. Направление связи
3. Какую долю дисперсии зависимой переменной объясняет построенная модель
4. Статистическую значимость модели

2. Если коэффициент корреляции Спирмена близок к +1, это означает:

1. Сильную прямую линейную связь

2. Сильную прямую монотонную (не обязательно линейную) связь
3. Отсутствие связи
4. Сильную обратную линейную связь

3. В уравнении линейной регрессии y = kx + b, коэффициент k (угловой коэффициент) показывает:
1. На сколько единиц изменится y при изменении x на единицу
2. Долю дисперсии y
3. Ошибку модели
4. Точку пересечения с осью ординат

8. Оценочные материалы промежуточной аттестации

2 семестр, Экзамен
Контролируемые ИДК: ОПК-1.1 ОПК-7.1 ОПК-1.2 ОПК-7.2 ОПК-1.3 ОПК-7.3 ОПК-1.4 ОПК-1.5 ОПК-1.6
Вопросы/Задания:

1. Аксиоматическое определение вероятности по Колмогорову. Пространство элементарных исходов, сигма-алгебра событий, вероятность как мера. Почему аксиоматический подход важен для построения вероятностных моделей в машинном обучении?

2. Условная вероятность и независимость событий. Вывод формулы полной вероятности и теоремы Байеса. Приведите пример применения теоремы Байеса в задачах классификации (наивный байесовский классификатор).

3. Формула Байеса: априорная и апостериорная вероятности. Объясните роль априорного распределения в байесовском подходе к машинному обучению. В чем заключается идея байесовской регуляризации?

4. Понятие дискретной случайной величины (ДСВ). Ряд распределения, функция вероятности. Числовые характеристики: математическое ожидание, дисперсия, моменты. Как интерпретируется математическое ожидание в контексте функции потерь при обучении модели?

5. Биномиальное распределение: условия возникновения, формула Бернулли, числовые характеристики. Приведите пример задачи классификации, где возникает биномиальная случайная величина.

6. Распределение Пуассона как предел биномиального распределения. Свойства и применение для моделирования редких событий. Где в анализе данных могут встречаться данные, распределенные по Пуассону?

7. Понятие непрерывной случайной величины (НСВ). Функция плотности распределения вероятностей и интегральная функция распределения, их свойства и взаимосвязь. Чем принципиально отличаются способы задания ДСВ и НСВ?

8. Нормальное (Гауссовское) распределение: параметры, форма кривой, функция плотности. Правило трех сигм. Объясните, почему нормальное распределение занимает центральное место в теории вероятностей и машинном обучении.

9. Центральная предельная теорема: формулировка и практическое значение. Как ЦПТ обосновывает использование нормального распределения при анализе выборочных средних и в методе Монте-Карло?

10. Равномерное распределение: плотность, числовые характеристики. Приведите примеры использования равномерного распределения в задачах машинного обучения (генерация синтетических данных, инициализация параметров).

11. Совместное распределение двух случайных величин. Ковариация и коэффициент корреляции Пирсона, их смысл, свойства и ограничения. Почему корреляция не означает причинно-следственную связь?

12. Условные распределения и условное математическое ожидание. Как понятие условного математического ожидания M[Y|X] связано с задачей регрессионного анализа? Что такое функция регрессии?

13. Многомерное нормальное распределение: понятие, свойства. Почему оно широко используется в статистических моделях (например, в гауссовских процессах и линейном дискриминантном анализе)?

14. Генеральная совокупность и выборка. Требования к репрезентативности выборки. Понятие случайной выборки. Опишите проблемы, возникающие при работе с нерепрезентативными данными в задачах машинного обучения.

15. Выборочные характеристики: выборочное среднее, выборочная дисперсия, мода, медиана, квантили. Сравните выборочное среднее и медиану с точки зрения устойчивости к выбросам (робастности).

16. Эмпирическая функция распределения и гистограмма. Как они используются на этапе разведочного анализа данных (EDA) для понимания структуры данных и выбора модели?

17. Точечные оценки параметров распределения. Свойства оценок: несмещенность, состоятельность, эффективность. Почему выборочная дисперсия с делителем n-1 является несмещенной оценкой дисперсии?

18. Метод моментов и метод максимального правдоподобия (ММП) как способы получения оценок. Сравните эти подходы. Приведите пример оценки параметра нормального распределения методом ММП.

19. Понятие доверительного интервала. Доверительный интервал для математического ожидания при известной и неизвестной дисперсии. Как интерпретировать доверительный интервал в контексте практического применения модели?

20. Интервальное оценивание. Уровень доверия и точность оценки. Как объем выборки влияет на ширину доверительного интервала?

21. Статистическая гипотеза: нулевая и альтернативная гипотезы. Ошибки первого и второго рода. Уровень значимости и мощность критерия. Объясните смысл этих понятий на примере A/B-тестирования.

22. P-value: определение, интерпретация, связь с уровнем значимости. Распространенные ошибки при интерпретации p-value. Что означает фраза «p-value < 0.05»?

23. Параметрические и непараметрические критерии. Критерий Стьюдента (t-критерий): назначение, условия применимости. Для решения каких задач в анализе данных он применяется?

24. Критерий χ^2 (хи-квадрат) Пирсона: назначение (проверка независимости признаков и согласия распределений). Приведите пример использования критерия χ^2 в задачах обработки категориальных данных.

25. Понятие статистической значимости модели в машинном обучении. Как с помощью проверки гипотез можно сравнить две модели и определить, значимо ли одна модель лучше другой?

26. Корреляционный анализ. Коэффициент корреляции Пирсона и коэффициент ранговой корреляции Спирмена: сравнение, области применения, чувствительность к выбросам. В каких случаях предпочтительнее использовать коэффициент Спирмена?

27. Регрессионный анализ. Парная линейная регрессия: модель, метод наименьших квадратов (МНК) для оценки параметров. Интерпретация коэффициентов регрессии.

28. Коэффициент детерминации R^2: определение, смысл, интерпретация. Чем R^2 отличается от коэффициента корреляции? Какие у него недостатки?

29. Предпосылки регрессионного анализа (условия Гаусса-Маркова). Что происходит с оценками параметров, если нарушается условие гомоскедастичности (постоянства дисперсии ошибок)?

30. Множественная регрессия. Проблема мультиколлинеарности и ее влияние на качество модели. Как мультиколлинеарность связана с корреляцией между признаками и какие методы применяют для ее устранения?

9. Материально-техническое и учебно-методическое обеспечение дисциплины

9.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. Хамидуллин Р. Я. Теория вероятностей и математическая статистика: учебное пособие / Р. Я. Хамидуллин. - Теория вероятностей и математическая статистика - Москва: Университет «Синергия», 2020. - 276 с. - 978-5-4257-0398-9. - Текст: электронный // IPR SMART: [сайт]. - URL: https://www.iprbookshop.ru/101341.html - Режим доступа: по подписке

2. Кацко И. А. Теория вероятностей и математическая статистика: учебник для вузов / И. А. Кацко, П. С. Бондаренко, Г. В. Горелова. - 3-е изд., испр. и доп. - Санкт-Петербург: Лань, 2023. - 436 с. - 978-5-507-45492-1. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/302663 - Режим доступа: по подписке

3. Методы искусственного интеллекта в обработке данных и изображений: монография / А. Ю. Дёмин, А. К. Стоянов, В. Б. Немировский, В. А. Дорофеев. - Томск: ТПУ, 2016. - 130 с. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/106257 - Режим доступа: по подписке

4. Пенькова Т. Г. Модели и методы искусственного интеллекта: учебное пособие / Т. Г. Пенькова, Ю. В. Вайнштейн. - Красноярск: СФУ, 2019. - 116 с. - 978-5-7638-4043-8. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/157579 - Режим доступа: по подписке

Дополнительная литература

1. Лихачев А. В. Введение в теорию вероятностей и математическую статистику: учебное пособие / А. В. Лихачев. - Новосибирск: НГТУ, 2019. - 102 с. - 978-5-7782-3903-6. - Текст: электронный // Издательство Лань: [сайт]. - URL: https://e.lanbook.com/book/152261 - Режим доступа: по подписке

2. Теория вероятностей и математическая статистика: учебное пособие (курс лекций) / Е. О. Тарасенко, И. В. Зайцева, П. К. Корнеев, А. В. Гладков. - Теория вероятностей и математическая статистика - Ставрополь: Северо-Кавказский федеральный университет, 2018. - 229 c. - 2227-8397. - Текст: электронный // IPR SMART: [сайт]. - URL: https://www.iprbookshop.ru/92605.html - Режим доступа: по подписке

3. Александрова О. В. Теория вероятностей и математическая статистика: практикум / О. В. Александрова, Т. В. Жмыхова. - Теория вероятностей и математическая статистика - Макеевка: Донбасская национальная академия строительства и архитектуры, ЭБС АСВ, 2019. - 108 c. - 2227-8397. - Текст: электронный // IPR SMART: [сайт]. - URL: https://www.iprbookshop.ru/92352.html - Режим доступа: по подписке

4. Теория вероятностей и математическая статистика: учебное пособие / И. Л. Макарова, С. Ж. Симаворян, А. Р. Симонян, Е. И. Улитина. - Теория вероятностей и математическая статистика - Сочи: Сочинский государственный университет, 2020. - 130 c. - 2227-8397. - Текст: электронный // IPR SMART: [сайт]. - URL: https://www.iprbookshop.ru/106592.html - Режим доступа: по подписке

5. Мажуль И. И. Введение в теорию вероятностей: учебное пособие / И. И. Мажуль. - Введение в теорию вероятностей - Новосибирск: Новосибирский государственный технический университет, 2017. - 204 c. - 978-5-7782-3384-3. - Текст: электронный // IPR SMART: [сайт]. - URL: https://www.iprbookshop.ru/91188.html - Режим доступа: по подписке

9.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся

Профессиональные базы данных

Не используются.

Ресурсы «Интернет»

1. http://search.ebscohost.com/ - Academic Search Ultimate EBSCO publishing
2. https://link.springer.com/ - eBook Collections Springer Nature
3. https://scholar.google.ru/ - Гугл Академия
4. https://e.lanbook.com/ - Электронная библиотечная система "Лань"
5. https://rusneb.ru/ - Национальная электронная библиотека (НЭБ)
6. https://www.iprbookshop.ru/ - IPR BOOKS
7. https://znanium.ru/ - Электронно-библиотечная система "Znanium"
8. https://www.elibrary.ru/ - НЭБ eLIBRARY

9.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине

Перечень программного обеспечения
(обновление производится по мере появления новых версий программы)

Не используется.

Перечень информационно-справочных систем
(обновление выполняется еженедельно)
Не используется.

9.4. Оборудование и технические средства обучения

Компьютерный класс
Маркерная доска - 0 шт.
Моноблок - 64 шт.
Проектор - 4 шт.

10. Методические указания по освоению дисциплины (модуля)

С целью успешного освоения данного предмета студенту необходимо выполнять домашние работы по указанным преподавателям темам и представленным на занятиях материалам. Для более успешного выполнения работы студенту предлагается использование учебно-методических материалов, соответствующих указанной теме.

Самостоятельная работа студентов описывается и регулируется методическими рекомендациями по дисциплине.

Описание возможностей изучения дисциплины лицами с ОВЗ и инвалидами

При необходимости обучения студентов-инвалидов и лиц с ограниченными возможностями здоровья аудиторные занятия могут быть заменены или дополнены изучением полнотекстовых лекций, презентаций, видео- и аудиоматериалов в электронной информационно-образовательной среде (ЭИОС) университета. Индивидуальные задания подбираются в адаптированных к ограничениям здоровья формах (письменно или устно, в форме презентаций). Выбор методов обучения зависит от их доступности для инвалидов и лиц с ограниченными возможностями здоровья.

Рабочая программа дисциплины для инвалидов и лиц с ограниченными возможностями здоровья разрабатывается индивидуально с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.