

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего
образования
«ИРКУТСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ»

Структурное подразделение «Институт информационных технологий и анализа данных»

УТВЕРЖДЕНА:

на заседании Совета института ИТиАД им. Е.И.Попова

Протокол №8 от 24 февраля 2025 г.

Рабочая программа дисциплины

«ОРГАНИЗАЦИЯ ЦЕНТРОВ ОБРАБОТКИ ДАННЫХ»

Направление: 09.03.01 Информатика и вычислительная техника

Вычислительные машины, комплексы, системы и сети

Квалификация: Бакалавр

Форма обучения: очная

Документ подписан простой
электронной подписью
Составитель программы:
Айзенберг Наталья
Ильинична
Дата подписания: 20.06.2025

Документ подписан простой
электронной подписью
Утвердил: Говорков Алексей
Сергеевич
Дата подписания: 20.06.2025

Документ подписан простой
электронной подписью
Согласовал: Аношко Алексей
Федорович
Дата подписания: 20.06.2025

Год набора – 2025

Иркутск, 2025 г.

1 Перечень планируемых результатов обучения по дисциплине, соотнесённых с планируемыми результатами освоения образовательной программы

1.1 Дисциплина «Организация центров обработки данных» обеспечивает формирование следующих компетенций с учётом индикаторов их достижения

Код, наименование компетенции	Код индикатора компетенции
ПКС-2 Способность реализовать проекты по созданию и модернизации ИТ-инфраструктуры предприятия	ПКС-2.2
ПКС-8 Способность создавать и поддерживать базы данных	ПКС-8.1

1.2 В результате освоения дисциплины у обучающихся должны быть сформированы

Код индикатора	Содержание индикатора	Результат обучения
ПКС-2.2	Разработка технического задания для проектирования центра обработки данных	Знать методики освоения и применения программных средств для решения практических задач, связанных с обработкой измерительной информации. Уметь применять программные средства для решения практических задач обработки данных; Владеть навыками освоения и использования новых программных средств для решения практических задач, связанных с разработкой алгоритмов обработки данных по результатам измерений.
ПКС-8.1	Способен решать задачи по мониторингу работы БД	Знать способы реализации алгоритмов обработки данных, способы моделирования выборочных данных, погрешностей измерений в них; Уметь реализовывать алгоритмы обработки данных на основе их формализованного представления на математическом языке, тестировать их с применением смоделированных измерительных данных, формулировать выводы по результатам; Владеть навыками построения вычислительных алгоритмов, генерации случайных чисел, подчиняющихся заданным законам распределения, моделирования с их помощью процесса получения выборочных

		оценок тех или иных величин.
--	--	------------------------------

2 Место дисциплины в структуре ООП

Изучение дисциплины «Организация центров обработки данных» базируется на результатах освоения следующих дисциплин/практик: «Математика», «Дискретная математика», «Математическая логика и теория алгоритмов», «Теория вероятностей, математическая статистика и случайные процессы»

Дисциплина является предшествующей для дисциплин/практик: «Основы проектной деятельности», «Производственная практика: технологическая (проектно-технологическая) практика»

3 Объем дисциплины

Объем дисциплины составляет – 5 ЗЕТ

Вид учебной работы	Трудоемкость в академических часах (Один академический час соответствует 45 минутам астрономического часа)	
	Всего	Семестр № 5
Общая трудоемкость дисциплины	180	180
Аудиторные занятия, в том числе:	64	64
лекции	32	32
лабораторные работы	32	32
практические/семинарские занятия	0	0
Самостоятельная работа (в т.ч. курсовое проектирование)	116	116
Трудоемкость промежуточной аттестации	0	0
Вид промежуточной аттестации (итогового контроля по дисциплине)	Зачет	Зачет

4 Структура и содержание дисциплины

4.1 Сводные данные по содержанию дисциплины

Семестр № 5

№ п/п	Наименование раздела и темы дисциплины	Виды контактной работы						СРС		Форма текущего контроля
		Лекции		ЛР		ПЗ(СЕМ)				
		№	Кол. Час.	№	Кол. Час.	№	Кол. Час.	№	Кол. Час.	
1	2	3	4	5	6	7	8	9	10	11
1	Необходимые сведения из теории вероятностей	1, 2	4	1	4			1	12	Оценка знаний по соответств ующей теме, Устный опрос
2	Линейная регрессия	3, 4, 5	6	2	8			3	16	Оценка знаний по соответств ующей

										теме
3	Спецификация регрессии	6, 7, 8	6	3	6			6	12	Оценка знаний по соответствующей теме, Отчет по лабораторной работе, Устный опрос
4	Важные функциональные формы регрессии	9, 10	4	4	6			5	12	Отчет по лабораторной работе
5	Модели бинарного выбора	11, 12	4	5	4			2	16	Отчет по лабораторной работе
6	Кластерный анализ	13, 14	4	6	4			4	16	Оценка знаний по соответствующей теме
7	Инструменты анализа данных	15, 16	4					6	24	Отчет по лабораторной работе
8	Подготовка к зачету							4	8	Оценка знаний по соответствующей теме
	Промежуточная аттестация									Зачет
	Всего		32		32				116	

4.2 Краткое содержание разделов и тем занятий

Семестр № 5

№	Тема	Краткое содержание
1	Необходимые сведения из теории вероятностей	Теория вероятностей (совместно с комбинаторикой) выступает математическим фундаментом: математической статистики - науки об общих закономерностях для повторяющихся событий; эконометрики - науки о зависимых явлениях (в основном социальных); анализа данных - науки о восстановлении явлений по внешним проявлениям; машинного обучения и искусственного интеллекта - науки о воспроизводстве восстановленных явлений и их адаптации на основании данных.
2	Линейная регрессия	Понятие и основные принципы регрессионной модели зависимости одной переменной от другой или нескольких других переменных с линейной функцией зависимости. Модель линейной регрессии является часто используемой и наиболее

		изученной в эконометрике.
3	Спецификация регрессии	Рассмотрение возможных ошибок регрессии: эндогенность, коррелированность остатков, гетероскедастичность, пропущенные переменные, лишние объясняющие переменные, неверная функциональная форма.
4	Важные функциональные формы регрессии	Популярные виды функциональных зависимостей при оценке эконометрической модели
5	Модели бинарного выбора	Модель зависимости бинарной переменной (принимающей всего два значения — 0 и 1) от совокупности факторов.
6	Кластерный анализ	Методами кластерного анализа решается задача восстановления структуры данных: выявления параметров, по которым данные могут быть сгруппированы при условии отсутствия явных группировочных признаков
7	Инструменты анализа данных	В лекции рассматривается программное обеспечение для анализа данных
8	Подготовка к зачету	Подведение итогов курса. Обсуждение знаний и умений, полученных в курсе.

4.3 Перечень лабораторных работ

Семестр № 5

№	Наименование лабораторной работы	Кол-во академических часов
1	Расчёт моментов распределения	4
2	Прогноз ряда с трендом и сезонностью	8
3	Построение правильно специфицированной регрессии	6
4	Построение регрессии с факторными переменными	6
5	Построение логит-модели	4
6	Построение кластерной модели	4

4.4 Перечень практических занятий

Практических занятий не предусмотрено

4.5 Самостоятельная работа

Семестр № 5

№	Вид СРС	Кол-во академических часов
1	Анализ научных публикаций	12
2	Выполнение компьютерных экспериментов и компьютерных лабораторных работ в дистанционном режиме	16
3	Оформление отчетов по лабораторным и	16

	практическим работам	
4	Подготовка к зачёту	24
5	Подготовка к контрольным работам	12
6	Подготовка к практическим занятиям (лабораторным работам)	36

В ходе проведения занятий по дисциплине используются следующие интерактивные методы обучения: Разбор кейсов

5 Перечень учебно-методического обеспечения дисциплины

5.1 Методические указания для обучающихся по освоению дисциплины

5.1.1 Методические указания для обучающихся по лабораторным работам:

Работы 1, 2 выполняются в MS Excel, остальные в R

Работа 1. Расчёт моментов распределения

Постановка задачи

Необходимо сгенерировать данные для дискретной аппроксимации нормального распределения и вычислить по этим данным его центральные моменты до 4-го. Построить ещё одно распределение с таким же математическим ожиданием и дисперсией, но с другим эксцессом. Результаты сравнить графически и аналитически.

Указания к решению

1. Поточечное построение плотности осуществляется с помощью функции «НОРМРАСП». Эти данные выводятся на график.
2. По плотности строится дискретная функция. По дискретной функции вычисляются моменты.
3. На этом же листе делается копия распределения и на копии вручную изменяются значения плотности (а значит, и функции) так, чтобы новое распределение имело требуемые в задании характеристики.

Работа 2. Прогноз ряда с трендом и сезонностью

Постановка задачи

Необходимо по индивидуальному набору данных (Ряды_данных.xls, взять ряд с номером, равным номеру в списке группы) построить прогноз продаж на 12 месяцев с помесечной детализацией.

Указания к решению

- 1.

Работа 3. Построение правильно специфицированной регрессии

Постановка задачи

Необходимо по набору данных diamonds (встроен в библиотеку ggplot2 пакета R) построить 3 модели цены бриллиантов с разным набором объясняющих переменных, провести их сравнительный анализ и выбрать наиболее правильно специфицированную.

Указания к решению

К заданию приложен файл «lab_03_diamonds_start.R» стартующий работу, включая описание данных diamonds

1. Провести визуальный анализ
 - a. Определить подходящую функциональную форму: линейная, полулогарифмическая, логарифмическая
 - b. Какие переменные коррелируют с ценой, какие нет

- с. Как на цену влияет огранка, цвет и чистота камня
- 2. По результатам визуального анализа
- а. Построить полную модель по всем переменным
- б. Отбросить лишние регрессоры
- с. Проверить модель на квадратичность
- 3. Три лучшие модели сравнить между собой

Результатом работы является скрипт на R с комментариями

Работа 4. Построение регрессии с факторными переменными

Постановка задачи

Необходимо по набору данных flats_moscow.txt (прилагается к заданию) построить 3 модели с разным набором объясняющих переменных, провести их сравнительный анализ и выбрать наилучшую.

Работа 5. Построение логит-модели

Постановка задачи

Необходимо по набору данных titanic.csv (прилагается к заданию) построить модель с наилучшей чувствительностью и специфичностью.

Указания к решению

- 1. Использовать скрипт lab_05_titanic.r как стартовый для работы. Довести работу до выбора чувствительности и специфичности

Результатом работы является скрипт на R с комментариями

Работа 6. Построение кластерной модели

Постановка задачи

Необходимо по индивидуальному набору данных корпоративной отчетности компаний clusters.zip сделать кластерное разбиение, используя данные за 2007-2014 гг. как один общий набор данных. Требуется определить и обосновать количество кластеров.

Указания к решению

- 1. Архив с набор данных содержит стартовый скрипт clusters.r
- 2. Каждый файл данных содержит информацию об одном показателе. Например, Profit.csv – это величина прибыли компаний. Данные по разным годам содержатся в отдельных столбцах.
- 3. Стартовый скрипт обрабатывает данные за один год. В задании требуется объединить данные за несколько лет.
- 4. При решении задачи использовать как визуальный анализ (представление данных в проекциях и параллельных координатах) и формальный (значение силуэта кластеров)

Результатом работы является скрипт на R с комментариями

5.1.2 Методические указания для обучающихся по самостоятельной работе:

- 1. Подготовка к лабораторным работам. Студенты должны самостоятельно проработать ряд разделов тем теоретического курса (исходя из задания лабораторной работы и своей

теоретической подготовки), а также освоить необходимые приёмы работы с выбранной средой разработки. Поскольку важнейшей задачей профессиональной деятельности ИТ-специалиста является самостоятельный поиск информации о функциях, методах, системных вызовах, используемых инструментах визуализации и т.п., студентам предлагается самостоятельно освоить приёмы работы со средствами разработки, а затем они должны самостоятельно применить необходимые для решения их задач вспомогательные инструменты (например, средства отображения графиков).

2. Самостоятельное изучение отдельных разделов дисциплины. Набор тем для самостоятельного изучения определяется темами лабораторных работ. Кроме того, для самостоятельного изучения предлагаются некоторые разделы темы “Основы регрессионного анализа”.

3. Подготовка отчётов по лабораторным работам. Студенты оформляют письменные отчёты по лабораторным работам, содержащие краткое описание выполненных шагов и полученных результатов, а также полученные экспериментальные данные, готовятся к их защите. В ходе подготовки самостоятельно прорабатываются темы, необходимые для выполнения и защиты работы, а также недостаточно хорошо усвоенный студентом теоретический материал.

6 Фонд оценочных средств для контроля текущей успеваемости и проведения промежуточной аттестации по дисциплине

6.1 Оценочные средства для проведения текущего контроля

6.1.1 семестр 5 | Оценка знаний по соответствующей теме

Описание процедуры.

Вопросы по теме № 1:

1. Что такое случайная величина?
2. При описании каких процессов применяется нормальный закон распределения?
3. Как в языках программирования генерируется случайная величина с равномерным законом распределения?
4. Как осуществляется генерация случайных величин, имеющих нормальное распределение?

Вопросы по теме № 2:

1. Что такое выброс?
2. Как влияют выбросы на погрешность измерений?
3. Как имитируется наличие выбросов в выборке?
4. Что такое измерение?

Вопросы по теме № 3:

1. Что такое математическое ожидание?
2. Какие требования предъявляются к оценкам?
3. Что такое параметр масштаба?
4. Как влияет наличие “выбросов” в выборке на выборочное среднее?

Вопросы по теме № 4:

1. Что такое робастная оценка?
2. Как находится альфа-усечённая оценка?
3. Какие виды помехоустойчивых оценок вы знаете?
4. Как можно проиллюстрировать графически вычисление альфа-усечённой оценки?

Вопросы по теме № 5:

1. Что такое доверительный интервал?
2. Что такое доверительная вероятность?
3. Для чего применяется распределение Стьюдента?
4. Для чего применяется t распределение?

Вопросы по теме № 6:

1. Что такое гипотеза?
2. Что такое закон распределения случайной величины?
3. Что понимается под объёмом выборки?
4. Для чего требуется проверять выборку на нормальность?

Критерии оценивания.

Отлично: Студент уверенно отвечает на вопрос и дополнительные вопросы, демонстрирует чёткое комплексное понимание связи рассматриваемых методов и алгоритмов с прочими аспектами обработки данных и математической статистики, их роль и место в практической деятельности

Хорошо: Студент относительно уверенно отвечает на оба вопроса, однако демонстрирует существенные пробелы в знаниях при ответах на дополнительные вопросы, не имеет комплексного представления о роли и месте тех или иных понятий или методов в практической деятельности специалистов по обработке данных

Удовлетворительно: Студент в целом отвечает не менее чем на 1 из 2 поставленных вопросов, однако демонстрирует лишь формальные заученные ответы, на заданные дополнительные вопросы не отвечает или отвечает крайне неуверенно. Также оценка “удовлетворительно” применяется при несвоевременной сдаче лабораторных работ даже если студент демонстрирует теоретическую подготовку на уровне “хорошо”.

Неудовлетворительно: Студент не способен ответить на оба вопроса, не способен ответить ни на один из дополнительных вопросов, демонстрирует полное незнание материала и т.п.

6.1.2 семестр 5 | Отчет по лабораторной работе

Описание процедуры.

Работа 1. Расчёт моментов распределения

Необходимо сгенерировать данные для дискретной аппроксимации нормального распределения и вычислить по этим данным его центральные моменты до 4-го. Построить ещё одно распределение с таким же математическим ожиданием и дисперсией, но с другим эксцессом. Результаты сравнить графически и аналитически.

Работа 2. Прогноз ряда с трендом и сезонностью

Необходимо по индивидуальному набору данных (Ряды_данных.xls, взять ряд с номером, равным номеру в списке группы) построить прогноз продаж на 12 месяцев с помесечной детализацией.

Работа 3. Построение правильно специфицированной регрессии

Необходимо по набору данных diamonds (встроен в библиотеку ggplot2 пакета R) построить 3 модели цены бриллиантов с разным набором объясняющих переменных,

провести их сравнительный анализ и выбрать наиболее правильно специфицированную.

Работа 4. Построение регрессии с факторными переменными

Необходимо по набору данных flats_moscow.txt (прилагается к заданию) построить 3 модели с разным набором объясняющих переменных, провести их сравнительный анализ и выбрать наилучшую.

Работа 6. Построение кластерной модели

Необходимо по индивидуальному набору данных корпоративной отчетности компаний clusters.zip сделать кластерное разбиение, используя данные за 2007-2014 гг. как один общий набор данных. Требуется определить и обосновать количество кластеров.

Критерии оценивания.

При защите проверяется: работоспособность разработанной программы, соответствие результатов работы предъявляемым требованиям, правильность и творческий подход к выполнению заданий, знание теоретического материала, необходимого для выполнения работ, качество и логичность написанного отчёта.

По лабораторной работе демонстрируется работоспособная программа на языке программирования высокого уровня (язык программирования выбирается на усмотрение студента, по согласованию с преподавателем, рекомендуется использовать язык, освоенный в ходе изучения дисциплины «Программирование»), которая проверяется преподавателем и защищается обучающимся. Для успешной сдачи лабораторной работы студенту необходимо предоставить отчёт, содержащий необходимые результаты и выводы, а также защитить представленную работу. В ходе защиты студенту необходимо дать краткое изложение основных этапов выполнения лабораторной работы, устно ответить на теоретические вопросы по теме лабораторной работы, а также продемонстрировать умение ориентироваться в программном коде и знание соответствующего математического аппарата.

6.1.3 семестр 5 | Устный опрос

Описание процедуры.

Основные вопросы устного опроса:

1. Измерения и наблюдения.
2. Основная задача обработки результатов измерений.
3. Статистическая обработка измерений.
4. Критерии качества оценок.
5. Методы оценивания параметров.
6. Метод максимального правдоподобия.
7. Логарифмическая функция правдоподобия.
8. Метод наименьших квадратов.
9. Достоинства и недостатки метода наименьших квадратов.
10. Проверка гипотез для математического ожидания.
11. Проверка гипотез для дисперсии.
12. Проверка гипотез о законах распределения случайной величины.
13. Основные задачи регрессионного анализа.
14. Линейная и нелинейная регрессия.
15. Определение коэффициентов линейной модели.
16. Центрирование данных.
17. Точность оценки регрессии.

18. Исследование уравнения регрессии.
19. Доверительные интервалы для коэффициентов регрессии.
20. Доверительные интервалы для прогноза.
21. Критерии значимости регрессии.
22. Множественная линейная регрессия.
23. Коэффициент множественной корреляции.
24. Проверка общей линейной гипотезы в регрессионных моделях.
25. Взвешенный метод наименьших квадратов.
26. Исследование остатков.
27. Корреляция между остатками.

Критерии оценивания.

Отлично: Студент уверенно отвечает на вопрос и дополнительные вопросы, демонстрирует чёткое комплексное понимание связи рассматриваемых методов и алгоритмов с прочими аспектами обработки данных и математической статистики, их роль и место в практической деятельности

Хорошо: Студент относительно уверенно отвечает на оба вопроса, однако демонстрирует существенные пробелы в знаниях при ответах на дополнительные вопросы, не имеет комплексного представления о роли и месте тех или иных понятий или методов в практической деятельности специалистов по обработке данных

Удовлетворительно: Студент в целом отвечает не менее чем на 1 из 2 поставленных вопросов, однако демонстрирует лишь формальные заученные ответы, на заданные дополнительные вопросы не отвечает или отвечает крайне неуверенно. Также оценка “удовлетворительно” применяется при несвоевременной сдаче лабораторных работ даже если студент демонстрирует теоретическую подготовку на уровне “хорошо”.

Неудовлетворительно: Студент не способен ответить на оба вопроса, не способен ответить ни на один из дополнительных вопросов, демонстрирует полное незнание материала и т.п.

6.2 Оценочные средства для проведения промежуточной аттестации

6.2.1 Критерии и средства (методы) оценивания индикаторов достижения компетенции в рамках промежуточной аттестации

Индикатор достижения компетенции	Критерии оценивания	Средства (методы) оценивания промежуточной аттестации
ПКС-2.2	Студенты грамотно используют различных программные средства для решения практических задач, связанных с реализацией и исследованием алгоритмов обработки данных	Устное собеседование по теоретическим вопросам, выполнение лабораторных работ
ПКС-8.1	Студент самостоятельно	Устное

	разрабатывает и тестирует с применением модельных данных алгоритмы решения лабораторных работ на основе их описания на математическом языке, способен в полной мере объяснить их логику и функционирование, понимает необходимые условия их работы, факторы влияющие на точность результата и т.п.	собеседование по теоретическим вопросам, выполнение лабораторных работ
--	--	--

6.2.2 Типовые оценочные средства промежуточной аттестации

6.2.2.1 Семестр 5, Типовые оценочные средства для проведения зачета по дисциплине

6.2.2.1.1 Описание процедуры

1. Студент в течении семестра сдает все устные опросы
2. Студент сдает 5 лабораторных работ (в течении семестра или в зачётную неделю)

6.2.2.1.2 Критерии оценивания

Зачтено	Не зачтено
1. Устные опросы сданы на оценку не менее "удовлетворительно" 2. Студент сдал 5 лабораторных работ	Не выполнено хотя бы один из критериев: 1. Все устные опросы сданы на оценку не менее "удовлетворительно" 2. Сдано 5 лабораторных работ

7 Основная учебная литература

1. Буренков С. И. Теория вероятностей и математическая статистика : конспект лекций / С.И. Буренков, И. М. Сидоров, А. А. Трухан, 2008. - 255.
2. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике : учебное пособие для вузов / В. Е. Гмурман, 2008. - 403.
3. Эконометрика : учеб. для вузов по специальности 061700 "Статистика" / И. И. Елисеева [и др.]; под ред. И. И. Елисеевой, 2007. - 574,[1].
4. Кремер Н. Ш. Эконометрика : учеб. для вузов / Н. Ш. Кремер, Б. А. Путко, 2007. - 310.

8 Дополнительная учебная литература и справочная

1. Тихомиров Н. П. Эконометрика : учеб. для вузов по специальности "Мат. методы в экономике" / Н. П. Тихомиров, Е. Ю. Дорохина, 2007. - 510.
2. Аистов А. В. Эконометрика шаг за шагом : учеб. для вузов по направлению подгот. "Экономика" / А. В. Аистов, А. Г. Максимов, 2006. - 177.
3. Гитис Леонид Хаскелевич. Статистическая классификация и кластерный анализ / Л. Х. Гитис, 2003. - 156.

4. Дюран Б. Кластерный анализ / Б. Дюран, П. Оделл, 1977. - 128.

9 Ресурсы сети Интернет

1. <http://library.istu.edu/>
2. <https://e.lanbook.com/>
3. <https://www.r-studio.com/ru/>

10 Профессиональные базы данных

1. <http://new.fips.ru/>
2. <http://www1.fips.ru/>

11 Перечень информационных технологий, лицензионных и свободно распространяемых специализированных программных средств, информационных справочных систем

1. Excel Link concurrent AcademicEdition

12 Материально-техническое обеспечение дисциплины

1. Экран Projecta с электроприводом Compact
2. Проектор Epson EB-W04LCD.WXGA 1280*800.3000:1.2800 ANSI Lumens
3. Доска магнитно-маркерная INDEX настенная ,размер 1x1.8 м
4. Компьютер Intel Core i3 3220/DDR3 4Gb/Hdd 1Tb/GF 1Gb/DVD+RW/Sound Net km/LCD23"/