Федеральное государственное образовательное бюджетное учреждение
высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)

Кафедра информационных технологий
Факультет информационных технологий и анализа больших данных

Документ подписан усиленной неквалифицированной электронной подписью
Организация: Финансовый университет при Правительстве РФ
Утвержено: Проректор по учебной и методической работе Е.А. Каменева
Сертификат: Ik5QWW6derL7UTdkZgP3IbjRnqDYZjHu
Дата: 08.10.2025 г.

Н. В. Гринева

Методология поиска источников данных и подготовки данных для анализа

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки
01.04.02 - Прикладная математика и информатика

Направленность программы магистратуры
Инженер машинного обучения

Рекомендовано
Факультет информационных технологий и анализа больших данных
(протокол №2 от 18.11.2025)

Одобрено
Кафедра информационных технологий
(протокол №11 от 12.11.2025)

Москва 2025

Содержание

Наименование разделов РПД		**Стр.**
1.	Наименование дисциплины	5
2.	Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине	5
3.	Место дисциплины в структуре образовательной программы	7
4.	Объем дисциплины(модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	8
5.	Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий	8
5.1.	Содержание дисциплины	8
5.2.	Учебно-тематический план	10
5.3.	Содержание семинаров, практических занятий	10
6.	Перечень учебно-методического обеспечения для	12

	самостоятельной работы обучающихся по дисциплине	
6.1.	Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	12
6.2.	Перечень вопросов, заданий, тем для подготовки к текущему контролю	14
7.	Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	16
8.	Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	23
9.	Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	25
10.	Методические указания для обучающихся по освоению дисциплины	26
11.	Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных	28

	справочных систем	
12.	Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	29

1. Наименование дисциплины

«Методология поиска источников данных и подготовки данных для анализа».

2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Код компетенции	**Наименование компетенции**	**Индикатор достижения компетенции**	**Результаты обучения**
ПКН-1	Способность самостоятельно приобретать и применять знания в области прикладной математики и информатики, а также поддерживать коллективную научную коммуникацию, организовывать научные мероприятия	Самостоятельно приобретает и применяет знания в области прикладной математики и информатики	**Знать:** методы описательного анализа данных **Уметь:** проводить описательный анализ, оценивать качество информации
		Демонстрирует самостоятельность при приобретении и применении новых знаний в области прикладной математики и информатики	**Знать:** методы поиска выбросов, в том числе в данных неизвестной структуры **Уметь:** находить выбросы в данных и обосновывать выбор метода
		Ведет эффективную научную коммуникацию в рамках командных мероприятий	**Знать:** методы оценки распределения данных **Уметь:** излагать основные особенности распределения

			данных в рамках командной работы
		Участвует в организации и проведении научных мероприятий	**Знать:** методы кластеризации данных **Уметь:** проводить кластеризацию данных в ходе научных исследований
ПКН-3	Способность проводить самостоятельные научные исследования в профессиональной области	Демонстрирует знания в области проведения самостоятельных научных исследований в профессиональной области	**Знать:** основные методы предобработки данных **Уметь:** оценивать характер исследовательских данных, обозначать основные направления предобработки
		Участвует в научных исследованиях в профессиональной области	**Знать:** методы восстановления целостности исследовательских данных в условиях их недостаточности **Уметь:** восстанавливать пропуски в данных, предназначенных для исследования
ПК-1	Способность вести разработку и управлять разработкой	Способность использовать инструментальные средства управления	**Знать:** инструментальные средства в разработке и управлении

	продуктов и решений на основе машинного обучения и анализа больших данных	разработкой прикладных решений на основе машинного обучения	прикладных задач на основе методов машинного обучения **Уметь:** проводить исследования, применять методы оптимизации и прогнозировани я с помощью методов машинного обучения
		Способность настраивать и поддерживать инфраструктуру операций разработки интеллектуальн ого программного обеспечения	**Знать:** инфраструктуру операций разработки интеллектуальн ого программного обучения в области поиска и обработки данных **Уметь:** применять на практике библиотеки и комплексы программ при разработке интеллектуальн ого программного обеспечения

3. Место дисциплины в структуре образовательной программы

Дисциплина «Методология поиска источников данных и подготовки данных для анализа» относится к «Модулю направленности программы магистратуры»

4. Объем дисциплины(модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Очная форма обучения

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Модуль 3 (в часах)
Общая трудоемкость дисциплины	**3/108**	**108**
Контактная работа - Аудиторные занятия	**24**	**24**
Лекции	8	8
Семинары, практические занятия	16	16
Самостоятельная работа	84	84
Вид текущего контроля	Контрольная работа;	Контрольная работа
Вид промежуточной аттестации	Зачет;	Зачет

5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1. Содержание дисциплины

Тема 1. Поиск и сбор данных для анализа

Основные источники данных: выгрузка данных из учетных систем, поиск данных в открытых источниках, приобретение данных из закрытых источников, проведение собственных исследований, экспертные оценки. Основы веб-майнинга. Характер собираемой информации: данные упорядоченные, неупорядоченные и транзакционные.

Тема 2. Описательный анализ данных

Оценки параметров генеральной совокупности. Критерии оценки репрезентативности и объема выборки. Описательные статистики количественных данных. Описательные статистики категориальных данных. Критерии оценки качества информации. Методы визуализации в оценке качества данных.

Тема 3. Очистка и предобработка данных

Основные виды аномалий данных: дубликаты и противоречия, пропуски, выбросы. Поиск дубликатов и противоречий. Методы работы с дубликатами и противоречиями.

Характер пропусков в данных: неслучайные (MNAR), случайные (MAR), абсолютно случайные (MCAR). Критерий Литтла. Визуализация в анализе пропусков. Методы импутации данных: импутация единым значением, импутация случайными величинами, параметрические модели импутации, импутация на основе методов машинного обучения. Множественная импутация.

Методы поиска выбросов. Квантильные методы. Параметрические методы, восстановление смесей. Непараметрические методы. Кластерные методы. Алгоритм локального уровня выброса (LOF). Изолирующие леса. Метод Леланда Уилкинсона. Критерии отнесения нетипичных значений к выбросам: на основе априорных предположений, на основе известных примеров. Сглаживание экстремальных значение. Винсоризация.

Нормализация данных, ее назначение. Методы кодирования категориальных величин.

Тема 4. Оценка распределения многомерных данных

Визуальное представление многомерных данных. Методы кластеризации в оценке распределения многомерных данных. Итеративные методы кластеризации: Expectation - Maximisation , k - средних, их модификации. Иерархические методы кластеризации.

Плотностные методы кластеризации. Нечеткая кластеризация. Самоорганизующиеся карты Кохонена как инструмент представления структуры многомерных данных.

5.2. Учебно-тематический план

Очная форма обучения

№ п/п	Наименование тем(разделов) дисциплины	Трудоемкость в часах				
		Всего	Контактная работа - Аудиторная работа			Самостоятельная работа
			Общая, в т.ч.:	Лекции	Семинары, практические занятия	
1	Поиск и сбор данных для анализа	14	4	2	2	10
2	Описательный анализ данных	26	6	2	4	20
3	Очистка и предобработка данных	30	6	2	4	24
4	Оценка распределения многомерных данных	38	8	2	6	30
	Итого	108	24	8	16	84

5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	**Перечень вопросов для обсуждения на семинарах, практических занятиях**	**Формы проведения занятий**
Поиск и сбор данных	Выгрузка данных из	Участие в решении

для анализа	учетных систем, поиск данных в открытых источниках. Основы веб-майнинга. Характер собираемой информации: данные упорядоченные, неупорядоченные и транзакционные.	задач на практических занятиях. Обсуждения по результатам самостоятельной работы.
Описательный анализ данных	Критерии оценки репрезентативности и объема выборки. Описательные статистики количественных данных. Описательные статистики категориальных данных. Методы визуализации в оценке качества данных.	Участие в решении задач на практических занятиях. Обсуждения по результатам самостоятельной работы.
Очистка и предобработка данных	Основные виды аномалий данных. Поиск дубликатов и противоречий. Методы работы с дубликатами и противоречиями. Характер пропусков в данных. Критерий Литтла. Визуализация в анализе пропусков. Методы импутации данных. Множественная импутация. Методы поиска выбросов. Квантильные методы. Параметрические методы, восстановление смесей. Кластерные методы. Алгоритм локального уровня выброса (LOF). Изолирующие леса. Критерии отнесения нетипичных значений к выбросам: на основе	Участие в решении задач на практических занятиях. Обсуждения по результатам самостоятельной работы.

	априорных предположений, на основе известных примеров. Сглаживание экстремальных значение. Винсоризация. Нормализация данных, ее назначение. Методы кодирования дискретных значений.	
Оценка распределения многомерных данных	Визуальное представление многомерных данных. Методы кластеризации в оценке распределения многомерных данных. Итеративные методы кластеризации: Expectation-Maximisation, k-средних, их модификации. Нечеткая кластеризация. Самоорганизующиеся карты Кохонена как инструмент представления структуры многомерных данных.	Участие в решении задач на практических занятиях. Обсуждения по результатам самостоятельной работы.

6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Поиск и сбор данных	Основные источники	Работа с учебной

для анализа	данных. Поиск данных в открытых источниках. Основы веб-майнинга.	литературой. Работа с источниками и поиск информаций в Интернете.
Описательный анализ данных	Оценки параметров генеральной совокупности. Критерии оценки репрезентативности и объема выборки. Критерии оценки качества информации.	Работа с учебной литературой. Решение типовых задач. Работа с источниками и поиск информаций в Интернете.
Очистка и предобработка данных	Методы работы с дубликатами и противоречиями. Характер пропусков в данных: неслучайные (MNAR), случайные (MAR), абсолютно случайные (MCAR). Критерий Литтла. Методы импутации данных: параметрические модели импутация, импутация на основе методов машинного обучения. Множественная импутация. Методы поиска выбросов. Кластерные методы. Алгоритм локального уровня выброса (LOF). Изолирующие леса. Метод Леланда Уилкинсона. Сглаживание экстремальных значение. Винсоризация. Методы кодирования категориальных величин.	Работа с учебной литературой. Решение типовых задач. Работа с источниками и поиск информаций в Интернете.
Оценка распределения многомерных данных	Иерархические методы кластеризации. Самоорганизующиеся карты Кохонена как инструмент представления	Работа с учебной литературой. Решение типовых задач. Работа с источниками и поиск информаций в Интернете.

	структуры многомерных данных.	

6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю

Примерный вариант контрольной работы

1. В файлах 1 и 2 содержатся данные по сотрудникам некоторой организации и их заработной платы, начисленной за январь месяц. Объедините эти файлы в единую таблицу. Обратите внимание, что поля могут называться по-разному и иметь разный формат полей (имена объединенной таблицы можно брать как из первой, так и из второй таблицы). Внутри каждого файла могут содержаться дубликаты; кроме того, дубликаты могут возникнуть при объединении файлов. Дубликаты необходимо удалить.

Этап, содержащий результат, назовите «Ответ 1».

2. На основе таблицы «Ответ 1» создайте единый список из 10% мужчин и 10% женщин, заработная плата которых за январь была наибольшей.

Этап, содержащий результат, назовите «Ответ 2».

3. В файле 3 содержатся данные по заработной плате работников того же предприятия за следующий месяц - февраль. Добавьте к таблице «Ответ 1» столбцы с заработной платой за февраль и должностями. При этом какие-то работники могут отсутствовать в январе (это значит, что они еще не были наняты на работу) или в феврале (это значит, что они уволились). Нужно оставить в списке всех, кто работал в феврале (тех, кого нет в феврале, нужно исключить, а тех, кого нет в январе, оставить).

Этап, содержащий результат, назовите «Ответ 3».

4. Проверьте случаи, когда должности работников в январе и феврале не совпадают. В некоторых случаях это может быть связано с сокращениями, лишними пробелами, заглавными буквами и т.п. Такие случаи исправьте.

5. Посчитайте среднюю заработную плату следующим образом: если работник не работал в январе или же изменил должность в

феврале, средняя зарплата будет равна февральской. Если работник работал оба месяца на одной должности, средняя зарплата высчитывается как среднее арифметическое из зарплат за январь и февраль.

6. Отберите в таблицу 20% работников предприятия с наименьшей средней заработной платой.

7. Отберите в другую таблицу по 20% работников каждого отдела с наименьшей средней заработной платой.

8. Создайте таблицу, в которой останутся только работники, которые попали И в таблицу, полученную на этапе 6 И в таблицу, полученную на этапе 7.

Этап, содержащий результат, назовите «Ответ 4».

9. В файле 4 содержатся данные по премиям (в процентах), начисленным некоторым работникам в феврале. Добавьте их к таблице «Ответ 3». Если почему-то комбинация «табельный номер» и «фамилия, имя, отчество» в файле 4 не будут совпадать с теми, что уже есть в таблице «Ответ 3», такие записи нужно считать ошибочными, и добавлять их не нужно.

10. Высчитайте размер премии (в рублях) для каждого из работников, учитывая, что она исчисляется от средней заработной платы.

11. Посчитайте общий размер начислений («всего начислено») в феврале (заработная плата плюс премия) для каждого работника.

12. Выведите таблицу - ведомость начислений в форме: «название отдела», «ФИО», «табельный номер», «заработная плата за февраль», «премия», «всего начислено». Данные должны быть отсортированы по отделам, затем по ФИО.

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях Кафедры информационных технологий Факультета информационных технологий и анализа больших данных.

7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

Перечень компетенций с указанием индикаторов их достижения в процессе освоения образовательной программы содержится в разделе 2. *Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине.*

Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

ПКН-1 Способность самостоятельно приобретать и применять знания в области прикладной математики и информатики, а также поддерживать коллективную научную коммуникацию, организовывать научные мероприятия

1) Самостоятельно приобретает и применяет знания в области прикладной математики и информатики

Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции

Знать: методы описательного анализа данных

Уметь: проводить описательный анализ, оценивать качество информации

Типовые контрольные задания

Провести описательный (дескриптивный) анализ предложенного массива данных, определить основные проблемы в данных. Дать комплексную оценку качества данных, предложить пути его улучшения

2) Демонстрирует самостоятельность при приобретении и применении новых знаний в области прикладной математики и информатики

Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции

Знать: методы поиска выбросов, в том числе в данных неизвестной структуры

Уметь: находить выбросы в данных и обосновывать выбор метода

Типовые контрольные задания

В предложенном многомерном массиве данных найти и оценить аномальные значения. Определить, какие из них относятся к выбросам, исходя из априорного предположения о вероятности появления выбросов в представленных данных, равной 0,005.

3) Ведет эффективную научную коммуникацию в рамках командных мероприятий

Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции

Знать: методы оценки распределения данных

Уметь: излагать основные особенности распределения данных в рамках командной работы

Типовые контрольные задания

Оценить распределение данных в представленном массиве. Написать краткое заключение для других участников научной команды, в котором отразить распределение и описательные статистики каждой из величин, наличие аномалий в распределении.

4) Участвует в организации и проведении научных мероприятий

Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции

Знать: методы кластеризации данных

Уметь: проводить кластеризацию данных в ходе научных исследований

Типовые контрольные задания

Таблица содержит данные о сферах научных интересов 200 участников научной конференции. Проведите кластеризацию данных и на основе выделенных кластеров сделайте обоснованное

предположение о том, какие научные секции целесообразно провести в ходе конференции.

ПКН-3 Способность проводить самостоятельные научные исследования в профессиональной области

1) Демонстрирует знания в области проведения самостоятельных научных исследований в профессиональной области

Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции

Знать: основные методы предобработки данных

Уметь: оценивать характер исследовательских данных, обозначать основные направления предобработки

Типовые контрольные задания

Оценить массив предложенных данных, выявить имеющиеся проблемы в данных, составить поэтапный план предобработки информации для ее дальнейшего использования в исследованиях.

2) Участвует в научных исследованиях в профессиональной области

Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции

Знать: методы восстановления целостности исследовательских данных в условиях их недостаточности

Уметь: восстанавливать пропуски в данных, предназначенных для исследования

Типовые контрольные задания

Оценить пропуски в данных в предложенном массиве, определить характер и объемы пропусков. На основе этой информации выбрать и обосновать методы и провести импутацию данных.

ПК-1 Способность вести разработку и управлять разработкой продуктов и решений на основе машинного обучения и анализа больших данных

1) Способность использовать инструментальные средства управления разработкой прикладных решений на основе машинного обучения

Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции

Знать: инструментальные средства в разработке и управлении прикладных задач на основе методов машинного обучения

Уметь: проводить исследования, применять методы оптимизации и прогнозирования с помощью методов машинного обучения

Типовые контрольные задания

Скачать данные с предложенного сайта, выбрать и настроить необходимые библиотеки для обработки большого массива данных, построить прогнозную модель. Метод выбрать самостоятельно и обосновать.

2) Способность настраивать и поддерживать инфраструктуру операций разработки интеллектуального программного обеспечения

Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции

Знать: инфраструктуру операций разработки интеллектуального программного обучения в области поиска и обработки данных

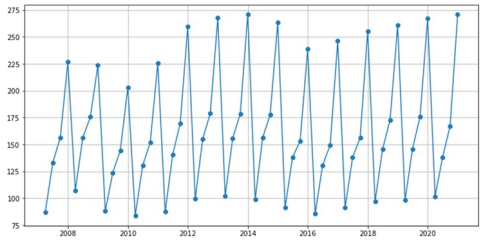
Уметь: применять на практике библиотеки и комплексы программ при разработке интеллектуального программного обеспечения

Типовые контрольные задания

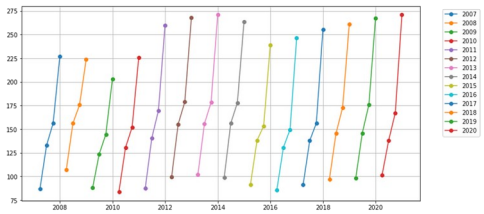
Построить самоорганизующиеся карты Кохонена как инструмент представления структуры многомерных данных и провести анализ. Предложить стратегии по управлению рисками.

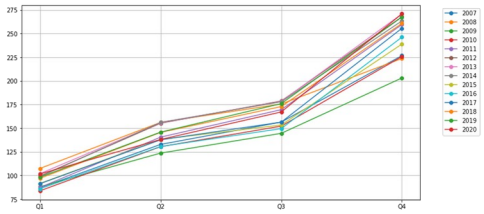
Примеры практико-ориентированных заданий

1. Для данных Индекса реальных инвестиций в основной капитал: построить график реальных данных, определить, есть ли сезонность, есть ли выбросы. Провести сглаживание ряда, построить прогноз.



Наблюдаем выраженную сезонность. Продемонстрируем её разбив по годам:



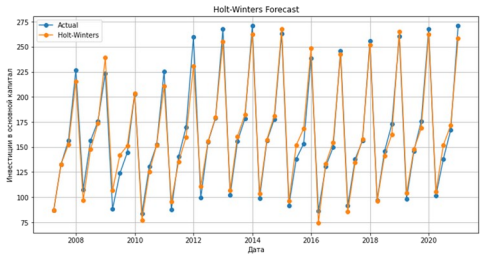


Сезонность в инвестициях может быть связана с определенными факторами, такими как налоговые сроки и сезонные колебания спроса на товары и услуги. Например, если налоговый год заканчивается в конце календарного года, то многие компании могут решать увеличить инвестиции в конце года, чтобы использовать налоговые льготы. Это может привести к повышению инвестиций в конце года и сезонной составляющей в индексе реальных инвестиций в капитал.

Для временных рядов с сезонностью типичным выбором кривой роста является экспоненциальное сглаживание Хольта-Винтерса.

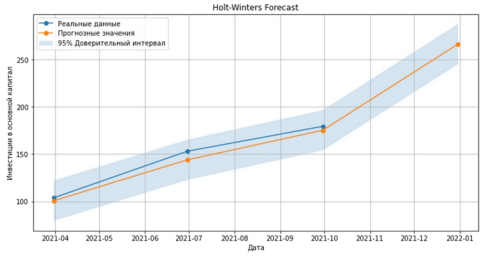
Она позволяет учитывать как тренд, так и сезонность в данных и предсказывать значения на будущее.

Аддитивная модель будет лучше подходить, потому что в данных не наблюдается увеличения амплитуды с течением времени. Тренд и сезонность изменяются относительно постоянно.



Как мы видим кривая роста Хольта-Винтерса с сезонностью в 4 отлично аппроксимирует реальные данные

Построим интервальный и точечный прогнозы. Более того у нас есть 3 квартала в 2021 году, реальные данные по которым у нас имеются, но в кривой роста они учтены не были. Посмотрим насколько хорошо был построен прогноз:



Прогноз отлично попадает в 95% доверительный интервал.

Примерные вопросы для подготовки к Зачету

1. Основные источники данных для анализа

2. Сущность, назначение и основные методы веб-майнинга.

3. Неупорядоченные, упорядоченные и транзакционные данные: особенности анализа.

4. Статистические оценки параметров генеральной совокупности.

5. Точечные статистические оценки

6. Интервальные статистические оценки.

7. Оценка репрезентативности и объема выборки.

8. Описательные статистики количественных данных.

9. Описательные статистики категориальных данных.

10. Методы визуализации в оценке качества данных.

11. Поиск дубликатов и противоречий.

12. Характер пропусков в данных. Неслучайные пропуски. Случайные пропуски. Совершенно случайные пропуски.

13. Визуализация пропусков в данных.

14. Зависимость методов восстановления пропусков от их величины и характера.

15. Импутация пропущенных данных одной величиной (средней, медианой).

16. Импутация пропущенных данных случайными величинами.

17. Импутация пропущенных данных параметрическими методами.

18. Импутация пропущенных данных с использованием методов машинного обучения.

19. Поиск выбросов в значениях величин квантильными методами. Винсоризация.

20. Параметрические методы поиска выбросов. Восстановление смесей.

21. Непараметрические методы поиска выбросов.

22. Методы машинного анализа в поиске выбросов: фактор локального выброса (LOF), «изолирующие леса», метод Леланда Уилкинсона.

23. Методы кодирования категориальных значений.

24. Визуальное представление многомерных данных.

25. Методы кластеризации в оценке распределения многомерных данных.

26. Итеративные алгоритмы кластеризации: Expectation-Maximisation, k-средних.

27. Иерархические методы кластеризации.

28. Плотностные методы кластеризации.

29. Нечеткая кластеризация.

30. Самоорганизующиеся карты Кохонена как инструмент представления структуры многомерных данных.

8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература:

1. Демидова, Л. А. Интеллектуальный анализ данных на языке Python [Электронный ресурс] : учебно-методическое пособие / Демидова Л. А. Москва : РТУ МИРЭА, 2021 92 с. Книга из коллекции РТУ МИРЭА - Информатика https://e.lanbook.com/book/218693. [БИК ID: RU-LAN-BOOK-218693]

2. Тесленко, И.Б. Анализ больших данных : Учебное пособие / И.Б. Тесленко, В.Е. Крылов, А.М. Губернаторов [и др.]. Электрон. дан. Москва : КноРус, 2025 295 с. Режим доступа: book.ru Internet access https://book.ru/book/955989 ISBN 978-5-406-14006-2. [БИК ID: RU\bookru\bibl\955989]

3. Паршинцева, Л.С. Многомерный анализ данных на Python : Учебник / Л.С. Паршинцева, А.А. Паршинцев Электрон. дан. Москва : КноРус, 2024 129 с. Режим доступа: book.ru Internet access

https://book.ru/book/951954 ISBN 978-5-406-12606-6. [БИК ID: RU\bookru\bibl\951954]

Дополнительная литература:

1. Алетдинова, А. А. Интеллектуальный анализ больших данных [Электронный ресурс] : учебное пособие / Алетдинова А. А.,Муртазина М. Ш. Новосибирск : НГТУ, 2023 66 с. Утверждено Редакционно-издательским советом университета в качестве учебного пособия Книга из коллекции НГТУ - Информатика https://e.lanbook.com/book/404567 ISBN 978-5-7782-4899-1. [БИК ID: RU-LAN-BOOK-404567]

2. Соловьев, В.И. Анализ данных в экономике: теория вероятностей, прикладная статистика, обработка и визуализация данных в Microsoft Excel : учебник для направления бакалавриата "Экономика и управление" / В.И. Соловьев ; Финуниверситет Москва : Кнорус, 2019 498 с. + Тираж 500 экз. (Бакалавриат) ISBN 978-5-406-06940-0. [БИК ID: RU\FA\books\66486]

3. Ямашкин, С. А. Системный анализ пространственных данных [Электронный ресурс] / Ямашкин С. А.,Ямашкин А. А. Саранск : МГУ им. Н.П. Огарева, 2020 44 с. Книга из коллекции МГУ им. Н.П. Огарева - Информатика https://e.lanbook.com/book/204680 ISBN 978-5-7103-4050-9. [БИК ID: RU-LAN-BOOK-204680]

4. Воскобойников, Ю. Е. Статистический анализ экспериментальных данных в пакетах MathCAD и Excel [Электронный ресурс] : учебное пособие для вузов / Воскобойников Ю. Е. 2-е изд., стер. Санкт-Петербург : Лань, 2022 212 с. Книга из коллекции Лань - Информатика https://e.lanbook.com/book/256109 ISBN 978-5-507-45039-8. [БИК ID: RU-LAN-BOOK-256109]

5. Сандуляк, Д. А. Анализ и обработка экспериментальных данных [Электронный ресурс] : учебно-методическое пособие / Сандуляк Д. А.,Полисмакова М. Н. Москва : РТУ МИРЭА, 2023 64 с. Книга из коллекции РТУ МИРЭА - Инженерно-технические науки https://e.lanbook.com/book/382541 ISBN 978-5-7339-1902-7. [БИК ID: RU-LAN-BOOK-382541]

6. Алборов, Р. А. Статистический анализ данных с использованием современных информационных технологий [Электронный ресурс] / Алборов Р. А.,Остаев Г. Я.,Концевая С. М. Ижевск : УдГАУ, 2022 12 с.

Книга из коллекции УдГАУ - Информатика https://e.lanbook.com/book/422687. [БИК ID: RU-LAN-BOOK-422687]

7. Панов, М. А. Анализ данных с использованием языка программирования Python [Электронный ресурс] : учебное пособие / Панов М. А. Екатеринбург : УрГЭУ, 2024 329 с. Книга из коллекции УрГЭУ - Информатика https://e.lanbook.com/book/481577. [БИК ID: RU-LAN-BOOK-481577]

9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Электронная библиотека Финансового университета (ЭБ) http://elib.fa.ru/ (http://library.fa.ru/files/elibfa.pdf)

2. Электронно-библиотечная система BOOK.RU http://www.book.ru

3. Школа больших данных [Электронный ресурс]: сайт. – Режим доступа: https://www.bigdataschool.ru/

4. Электронно-библиотечная система "Университетская библиотека ОНЛАЙН" http://biblioclub.ru/

5. Образовательная платформа "ЮРАЙТ" https://urait.ru/

6. Научная электронная библиотека eLibrary.ru http://elibrary.ru

7. Информационно-образовательный портал Финуниверситета: https://org.fa.ru

8. Библиотека электронных ресурсов исторического факультета МГУ им. М.В. Ломоносова: http://www.hist.msu.ru/ER/index.html

9. •Деловая онлайн-библиотека Alpina Digital http://lib.alpinadigital.ru/

10. •Электронные продукты издательства Elsevier http://www.sciencedirect.com

11. 4.Официальный сайт Федеральной службы государственной статистики. URL: http://www.gks.ru (Росстат)

12. Финансовая справочная система «Финансовый директор» http://www.1fd.ru/

13. Ресурсы информационно-аналитического агентства по финансовым рынкам Cbonds.ru https://cbonds.ru/

14. www.cbr.ru - Официальный сайт Банка России

15. "Деловая онлайн библиотека" издательства "Альпина Паблишер" http://lib.alpinadigital.ru/en/library

10. Методические указания для обучающихся по освоению дисциплины

Самостоятельная работа студентов проходит аудиторно и внеаудиторно. Организации самостоятельной работы служит календарно-тематический план изучения дисциплины. В этом плане указана тематика лекций, практических занятий, вопросы и задания для самостоятельного изучения.

При подготовке к лекции целесообразно предварительно познакомиться с ее содержанием по рекомендованным пособиям и выделить наиболее трудные вопросы. Во время лекций следует конспектировать содержание лекции. После занятий следует провести работу с конспектом: отредактировать записи, оформить конспект. При оформлении целесообразно выделять специальным образом названия тем и формулировки вопросов, основные определения, формулировки теорем и примеры. Сделанные записи нужно сверить с учебниками и учебными пособиями и в случае расхождений проконсультироваться с преподавателем.

Методические указания по проведению практических занятий:

По структуре практические занятия следует разделить на учебные и контрольные.

●Учебные практические занятия структурно состоят из следующих компонент:

1. Формулировка темы занятия, определение ключевых вопросов

2. Рассмотрение теоретических вопросов, связанных с текущим практическим занятием;

3. Постановка задания.

4. Индивидуальное выполнение заданий на компьютере;

5. Обсуждение результатов. Разбор типичных ошибок, возникших в самостоятельной работе;

6. Корректировка заданий для самостоятельной работы студентов.

● Контрольные практические занятия структурно состоят из следующих компонент:

1. Постановка задания, уточнение параметров представления отчетных данных.

2. Проведение аудиторной контрольной работы

3. Корректировка заданий для самостоятельной работы студентов.

При подготовке к практическому занятию необходимо повторить или, если это требуется, изучить соответствующий теоретический материал. Во время занятия нужно точно записывать формулировки решаемых задач, вопросы, указания преподавателя к решению и разбираемые решения. После занятий необходимо просмотреть записанные решения и восстановить в решениях имеющиеся пробелы. В случае затруднений отметить соответствующие задания и обратиться за консультацией к преподавателю. Практические занятия проходят, как правило, в интерактивной форме и преподаватель учитывает активность студентов, направленную на решение предложенных задач, и в поиске ответов на вопросы. Не следует бояться дать неверный ответ или допустить иную ошибку: исправление и анализ ошибок в режиме общения с преподавателем и сокурсниками в ходе практического занятия способствуют освоению учебного материала и предупреждают появление ошибок в дальнейшем.

На практических занятиях используется проблемно-деятельностный подход для решения практических задач. Сущность проблемно-деятельностного обучения заключается в том, что в процессе учебных занятий создаются специальные условия, в которых обучающийся, опираясь на приобретенные знания, мысленно и практически действует в целях поиска и обоснования наиболее оптимальных вариантов ее решения. Создается

проблемная задача, студенты знакомятся с задачей, анализируют ее, выделяют лежащее в ее основе противоречие, создают и обосновывают модель своих возможных действий по разрешению проблемной ситуации, пробуют разрешить возникшую проблему на основе имеющихся у них знаний, выстраивают модель своих действий по ее решению.

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем

- Комплект лицензионного программного обеспечения:

1. Python 3.8
2. Язык R
3. Jupyter Notebook (Windows 10)
4. Kaspersky

- Современные профессиональные базы данных и информационные справочные системы:

1. Язык программирования Python 3. https://pythonworld.ru/
2. eLIBRARY.RU
3. ИС «Поисковая платформа»
4. База данных финансовой информации AmadeusBureauvanDijk [Официальныйсайт].URL:https://amadeus.bvdinfo.com/version2013617/home.serv?product=amadeusneo.
5. Информационно-справочная система «СПАРК»
6. Демографический ежегодник России / Росстат. URL: https://rosstat.gov.ru/folder/210/document/13207
7. Индикаторы инновационной деятельности / ВШЭ. URL: https://www.hse.ru/primarydata/ii

8. Регионы России. Социально-экономические показатели / Росстат. URL: https://rosstat.gov.ru/folder/210/document/13204

9. Российский статистический ежегодник / Росстат. URL: https://rosstat.gov.ru/folder/210/document/12994

10. БД “СПАРК-ИНТЕРФАКС» URL: https://vk.com/www.spark-interfax.ru

- Сертифицированные программные и аппаратные средства защиты информации

1. не предусмотрены

12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

1. **Компьютерный класс** для проведения учебных занятий, предусмотренных программой, в том числе групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенный оборудованием и техническими средствами обучения: мебель аудиторная (столы, стулья, доска аудиторная), персональные компьютеры, набор демонстрационного оборудования (проектор, экран)