

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Горбунов Алексей Александрович

Должность: Заместитель начальника университета по учебной работе

Дата подписания: 09.07.2025 11:42:54

Уникальный программный ключ:

286e49ee1471d400cc1f45539d51ed7bbf0e9cc7

ФГБОУ ВО «Санкт-Петербургский университет ГПС МЧС России»

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ТЕХНОЛОГИИ И СИСТЕМЫ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ**

**Специалитет по специальности
27.05.01 Специальные организационно-технические системы**

**Специализация
«Информационно-аналитическая деятельность в специальных
организационно-технических системах»**

1. Цель и задачи дисциплины

Цели освоения дисциплины:

- формирование у обучающихся теоретических и практических знаний в области технологии и систем больших данных;
- выработка приемов и навыков решения конкретных задач управления МЧС на базе технологии и систем больших данных.

Перечень компетенций, формируемых в процессе изучения дисциплины

Компетенции	Содержание
ПК-3	Способен применять технологии информационно-аналитической деятельности в специальных организационно-технических системах для решения задач поддержки принятия решений в области предупреждения и ликвидации ЧС.

Задачи дисциплины:

- изучение основных положений технологии и систем больших данных;
- изучение задач МЧС, решаемых с помощью технологий и систем больших данных;
- формирование практических навыков решения интеллектуальных задач МЧС методами и технологиями больших данных.

2. Перечень планируемых результатов обучения дисциплины, соотнесенных с планируемыми результатами освоения образовательной программы

Индикаторы достижения компетенции	Планируемые результаты обучения по дисциплине
ПК-3.1. Использует принципы построения информационных систем, методов обработки больших данных, особенности построения интеллектуальных информационных систем и защиты информации.	Знает: состояние, перспективы и организацию применения технологий больших данных в МЧС; основные термины и определения теории больших данных, принципы построения систем больших данных, принципы машинного обучения. Умеет использовать методы: сбора и подготовки больших данных, методы регрессионного анализа, прогнозирования событий временными рядами, методы Data Mining (ассоциативного и последовательного анализа, классификации, кластерного анализа и др.), методы постобработки и интерпретации результатов анализа больших данных.
ПК-3.2. Умеет обосновывать применение методов и инструментов обработки информации, в том числе обработки больших данных, для решения задач в области предупреждения и ликвидации ЧС.	Знает способы применения методов и инструментов сбора и подготовки больших данных, регрессионного анализа и прогнозирования событий временными рядами, методы Data Mining (ассоциативного и последовательного анализа, классификации, кластерного анализа и др.). Умеет обосновывать методы машинного обучения для разработки компьютерных

	моделей решения задач в области прогноза, предупреждения и ликвидации ЧС,
ПК-3.3. Демонстрирует навыки сбора и анализа больших данных, построения интеллектуальных информационных систем, обеспечивающих решение задач в области предупреждения и ликвидации ЧС.	Владеет навыками применения АП «Логин» для построения и решения задач машинного обучения и применения систем больших данных в области предупреждения и ликвидации ЧС.

3. Место дисциплины в структуре основной профессиональной образовательной программы

Дисциплина относится к части, формируемой участниками образовательных отношений основной профессиональной образовательной программы специалитета по специальности **27.05.01 – Специальные организационно-технические системы**, специализация – **Информационно-аналитическая деятельность в специальных организационно-технических системах**.

4. Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 5 зачетных единиц, 180 часов.

4.1 Распределение трудоемкости учебной дисциплины по видам работ по семестрам для очной формы обучения

Вид учебной работы	Трудоемкость		
	з.е.	час.	час. по семестрам
			А
Общая трудоемкость дисциплины по учебному плану	5	180	180
Контактная работа		72	72
Лекции		32	32
Практические занятия		40	40
Лабораторные работы			
Консультации перед экзаменом		2	2
Самостоятельная работа		70	70
Курсовая работа			
Зачет			
Зачет с оценкой			
Экзамен		36	36

4.2. Тематический план, структурированный по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий для очной формы обучения

№ п/п	Наименование тем	Всего часов	Количество часов по видам занятий				Контроль	Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	Консультация		
А семестр								
1	Тема №1. Задачи управления МЧС России, большие данные и искусственный интеллект	16	4	4				8
2	Тема №2. Основы работы в АП Loginom Community	16	4	4				8
3	Тема №3. Введение в анализ больших данных	16	4	4/2*				8
4	Тема № 4. Предварительная обработка данных	16	2	6/2*				8
5	Тема № 5. Трансформация и визуализация данных	18	4	6				8
6	Тема № 6. Описательные статистики	16	4	4				8
7	Тема №7. Статистические исследования зависимостей	16	4	4				8
8	Тема №8. Искусственные нейронные сети.	16	4	4				8
9	Тема №9. Ассоциативный анализ чрезвычайных ситуаций	12	2	4				6
Консультация		2				2		
Экзамен		36					36	
Итого		180	32	40/4*		2	36	70

4.3 Содержание дисциплины для обучающихся очной формы обучения

Тема 1. Задачи управления МЧС России, большие данные и

искусственный интеллект.

Лекция 1.1. Вводная. Цели, задачи, объем, структура и содержание дисциплины «Технология и системы обработки больших данных». Методические рекомендации по изучению дисциплины «Технологии и системы обработки больших данных». Основные термины и определения.

Самостоятельная работа. Уяснить актуальность изучения дисциплины для специалистов по направлению 27.05.01 – специальные организационно-технические системы, специализация – информационно-аналитическая деятельность Роль дисциплины в системе повседневного и оперативного управления силами и средствами МЧС.

Рекомендуемая литература

Основная литература: [1];

Лекция 1.2. Перечень решений и задач по внедрению больших данных в деятельность МЧС России

Самостоятельная работа. Перспективные направления внедрения больших данных в МЧС России

Рекомендуемая литература

Основная литература: [1];

Практическое занятие 1.3. Атлас опасностей и рисков МЧС России, как пример применения теории анализа больших данных.

Самостоятельная работа. Ознакомление с правилами применения программного приложения Атласа опасностей и рисков.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Практическое занятие 1.4. Мобильное приложение МЧС России «Термические точки» как пример реализации интеллектуального анализа больших данных.

Самостоятельная работа. Ознакомление с правилами применения программного приложения «Термические точки».

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Тема 2. Основы работы в АП Loginom Community.

Лекция 2.1. Аналитические платформы. Назначение и возможности. Задачи аналитических платформ.

Самостоятельная работа. Ознакомление с методами анализа больших данных с помощью аналитических платформ

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Лекция 2.2. Назначение, задачи и возможности АП Loginom.
Самостоятельная работа. Ознакомление с правилами применения программного приложения «АП» Loginom.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Практическое занятие 2.3. Основные компоненты АП Loginom.
Запуск программы. Изучение рабочего окна программы. Действия с проектом.

Самостоятельная работа. Ознакомление с правилами применения программного приложения «АП» Loginom.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Практическое занятие 2.4. Интерфейс. Основные блоки интерфейса: главное меню, адресная строка, рабочее пространство, панель.

Самостоятельная работа. Ознакомление с правилами применения программного приложения «Термические точки».

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Тема 3. Введение в анализ больших данных.

Лекция 3.1. Хранилища данных. Идеи, лежащие в основе хранилища данных. Свойства хранилища данных. Типовая структура хранилища данных и его отличие от СУБД. Основные компоненты хранилища данных. OLAP-кубы. Схемы звезды и снежинки.

Самостоятельная работа. Ознакомление с правилами применения программного приложения «Термические точки».

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Лекция 3.2. Информационный и интеллектуальный анализ больших данных. Два подхода к анализу больших данных. Разведочный и интеллектуальный анализ больших данных. Пример больших данных. Система базовых механизмов анализа больших данных: механизмы подготовки, очистки и трансформации данных, механизмы Data mining, механизмы интерпретации данных.

Самостоятельная работа. Ознакомление с правилами применения программного приложения «Термические точки».

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Практическое занятие 3.3. Этапы интеллектуального анализа данных. Подготовка данных: создание набора данных. Очистка данных. Dataset.

Практическое занятие в форме практической подготовки. Пример создания набора данных в среде АП Loginom.

Самостоятельная работа. Ознакомление с правилами применения программного приложения «Термические точки».

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Практическое занятие 3.4. Понятие Dataset. Технология создания Dataset. Пример создания Dataset. Технология выбора Dataset из наборов данных.

Практическое занятие в форме практической подготовки. Методы создания Dataset. Пример создания Dataset.

Самостоятельная работа. Ознакомление с правилами применения программного приложения «Термические точки».

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Тема № 4. Предварительная обработка данных

Лекция 4.1. Этапы предварительной обработки данных. Методика предварительной обработки данных. Этап разведывательного анализа. Этап оценки и повышения качества данных: выявление и устранение пропусков, выбросов и устранение противоречий в данных. Устранение ошибок в данных.

Самостоятельная работа. Задачи очистки и предобработки данных.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Практическое занятие 4.2. Методы очистки данных. Алгоритмы анализа строк. Частотный анализ. Контрольные числа.

Самостоятельная работа. Рассмотреть на примерах методы очистки и стандартизации данных.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2]

Практическое занятие 4.3. Методы стандартизации данных.

Самостоятельная работа. Рассмотреть на примерах методы

стандартизации данных.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Практическое занятие 4.4. Пример предварительной обработки данных о термических точках в АП Loginom. Перечень данных о термических точках. Очистка данных. Использование данных в различных форматах.

Самостоятельная работа. Рассмотреть на примерах методы очистки и стандартизации данных.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Глава 5. Трансформация и визуализация данных

Лекция 5.1. Трансформация данных в АП Loginom. Сортировка, фильтрация строк и замена. Группировка и разгруппировка. Квантование. Таблицы и диаграммы. OLAP-куб. Отчеты.

Самостоятельная работа. Задачи трансформации данных.

Рекомендуемая литература

Основная литература: [2];

Дополнительная литература: [2].

Лекция 5.2. Визуализация данных в АП Логином: Таблицы, Диаграммы, OLAP кубы, Статистические данные.

Самостоятельная работа. Базовые навыки трансформации и визуализации данных в АП Логином.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Практическое занятие 5.3. Работа с визуализаторами Таблица и Диаграмма в АП Loginom.

Самостоятельная работа. Базовые навыки применения визуализаторов Таблица и Дииаграмма в АП Loginom.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Практическое занятие 5.4. Работа с визуализаторами OLAP куб и Статистика в АП Loginom.

Самостоятельная работа. Базовые навыки применения визуализаторов OLAP куб и Статистика в АП Loginom.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1].

Практическое занятие 5.5. Пример применения описательных статистик в АП Loginom.

Самостоятельная работа. Базовые навыки применения описательных статистик в АП Логином.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Тема 6. Описательные статистики.

Лекция 6.1. Выборочные статистики, правило 3-х σ и доверительный интервал. Понятие описательных статистик. Выборка. Выборочные характеристики. Правило 3-х σ . Доверительный интервал.

Самостоятельная работа. Изучить свойства выборочных оценок: несмещенность, эффективность и состоятельность.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Лекция 6.2. Выборочные статистики. Гистограмма, полигон и диаграмма размаха. Таблицы распределения данных по группам. Распределение частот. Интервал группирования. Гистограмма группировка по частоте. Полигон. Диаграмма размаха. Интерквартильный диапазон.

Самостоятельная работа. Изучить свойства гистограммы, полигона и диаграммы размаха.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Практическое занятие 6.3. Пример применения описательных статистик в АП Loginom. Данные по пожарам в России за 2019 – 2022 г. Импорт данных в АП Loginom. Настройка визуализаторов. Анализ результатов.

Самостоятельная работа. Изучить свойства гистограммы, полигона и диаграммы размаха.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Практическое занятие 6.4. Пример применения описательных статистик в АП Loginom. Данные по наводнениям в России за 2019 – 2022 г. Импорт данных в АП Loginom. Настройка визуализаторов. Анализ результатов.

Самостоятельная работа. Изучить свойства гистограммы, полигона и

диаграммы размаха.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Тема 7. Статистические исследования зависимостей.

Лекция 7.1. Корреляционный анализ. Зависимость между случайными величинами. Коэффициент корреляции Пирсона.

Самостоятельная работа. Изучить свойства выборочных оценок: несмещенность, эффективность и состоятельность.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Практическое занятие 7.2. Корреляционный анализ в АП Loginom. Импорт файла в АП Loginom. Разработка проекта. Настройка узлов проекта. Решение задачи корреляционного анализа.

Лекция 7.3. Парный регрессионный анализ. Модель парной регрессии. Метод наименьших квадратов. Коэффициент детерминации.

Самостоятельная работа. Изучить методы проверки гипотез в отношении коэффициентов регрессии.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Практическое занятие 7.4. Построение модели парной линейной регрессии в АП Loginom.

Самостоятельная работа. Обратная модель парной линейной регрессии. Фиктивная линейная связь и очистка переменных.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [2].

Тема №8. Искусственные нейронные сети.

Лекция 8.1. Нейрон и нейронная сеть. Биологический и искусственный нейрон. Состав нейрона синапсы, дендриты, аксон. Сумматор, функция возбуждения нейрона. Типы архитектур искусственного нейрона. Обучение с учителем, без учителя и с подкреплением.

Самостоятельная работа. Изучить подходы к решению задач классификации.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1, 2].

Лекция 8.2. Сеть прямого распространения сигналов. Внутренний слой сети. Количество слоев. Алгоритм обучения перцептрона.

Самостоятельная работа. Изучить подходы к решению задач выявления классификации.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1, 2].

Практическое занятие 8.3. Решение задачи классификации с помощью. сети прямого распространения сигнала в АП Loginom. Перемещение Excel файла в область построения сценария. Создание и настройка сценария. Запускаем сценарий на обучение, затем загружаем файл, который необходимо классифицировать и классифицируем его.

Самостоятельная работа. Изучить подходы к решению задач выявления ассоциаций.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1, 2].

Практическое занятие 8.4. Решение задачи прогнозирования с помощью. сети прямого распространения сигнала в АП Loginom. Перемещение Excel файла в область построения сценария. Перемещение Excel файла в область построения сценария. Создание и настройка сценария. Запускаем сценарий на обучение, затем загружаем файл, который необходимо прогнозировать и выполняем прогноз.

Самостоятельная работа. Изучить подходы к решению задач выявления ассоциаций.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1, 2].

Тема 9. Ассоциативный анализ чрезвычайных ситуаций.

Лекция 9.1. Технология выявления ассоциаций методом ассоциативного анализа. Задачи ассоциативного анализа. Ассоциации между двумя или несколькими событиями. Ассоциативные шаблоны.

Самостоятельная работа. Изучить подходы к решению задач выявления ассоциаций.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1, 2].

Практическое занятие 9.2. Решение задачи ассоциативного анализа событий в МЧС с помощью алгоритма Apriori.

Самостоятельная работа. Изучить основные показатели значимости ассоциативных правил.

Рекомендуемая литература

Основная литература: [1];

Дополнительная литература: [1, 2].

Практическое занятие 9.3. Решение задачи ассоциативного анализа событий в МЧС в среде АП Loginom.

Самостоятельная работа. Изучить технологию выявления ассоциаций в среде аналитической платформы Логином.

Рекомендуемая литература

Основная литература: [2];

Дополнительная литература: [1, 2].

5. Методические рекомендации по организации изучения дисциплины

При реализации программы дисциплины используются лекционные и практические занятия.

Общими целями занятий являются:

обобщение, систематизация, углубление, закрепление теоретических знаний по конкретным темам дисциплины;

формирование умений применять полученные знания на практике, реализация единства интеллектуальной и практической деятельности;

выработка при решении поставленных задач профессионально значимых качеств: самостоятельности, ответственности, точности, творческой инициативы.

Целями лекции являются:

дать систематизированные научные знания по дисциплине, акцентируя внимание на наиболее сложных вопросах;

стимулировать активную познавательную деятельность обучающихся, способствовать формированию их творческого мышления.

В ходе практического занятия обеспечивается процесс активного взаимодействия обучающихся с преподавателем; приобретаются практические навыки и умения. Цель практического занятия: углубление и закрепление знаний, полученных на лекции, формирование навыков использования знаний для решения практических задач; выполнение тестовых заданий по проверке полученных знаний и умений.

Самостоятельная работа обучающихся направлена на углубление и закрепление знаний, полученных на лекциях и практических занятиях, выработку навыков самостоятельного активного приобретения новых, дополнительных знаний, подготовку к предстоящим занятиям.

6. Оценочные материалы по дисциплине

Текущий контроль успеваемости обеспечивает оценивание хода освоения дисциплины, проводится в соответствии с содержанием дисциплины по видам занятий в форме опроса/решения задач под руководством преподавателя и самостоятельно/в форме тестирования.

Промежуточная аттестация обеспечивает оценивание промежуточных и окончательных результатов обучения по дисциплине, проводится в форме экзамена.

6.1. Примерные оценочные материалы:

6.1.1. Текущего контроля

Типовые вопросы для опроса:

1. Роль дисциплины в процессе подготовки специалиста
2. Основные термины и определения больших данных
3. Примеры больших данных в МЧС.
4. Разведочный анализ.
5. Анализ данных с помощью моделей обнаружения знаний в данных.
6. Базовые механизмы информационного анализа данных.
7. Базовые механизмы интеллектуального анализа данных.
8. Очистка данных.
9. Предобработка данных.
10. Заполнение пропусков и редактирование выбросов.
11. Технология трансформации данных.
12. Технология визуализации данных
13. Математическое ожидание и его выборочное среднее.
14. Дисперсия, среднее квадратическое отклонение и их выборочные оценки.
15. Ковариация, коэффициент корреляции и их выборочные оценки.
16. Модель парной регрессии.
17. Метод наименьших квадратов.
18. Коэффициент детерминации.
19. Постановка задачи проверки гипотез в отношении коэффициентов регрессии.
20. Постановка задачи проверки гипотезы в отношении коэффициента детерминации.
21. Парная нелинейная регрессия.
22. Регрессии нелинейные по объясняющим переменным и их линеаризация.
23. Регрессии нелинейные по параметрам и их линеаризация.
24. Модель множественной регрессии.
25. Прогнозирование событий по статистическим данным.

26. Временные ряды.
27. Разработка модели прогноза временного ряда.
28. Понятие и задачи Ассоциативного анализа.
29. Алгоритм Apriori ассоциативного анализа.
30. Последовательности двух и более событий.
31. Деревья решений.
32. Задачи, решаемые с помощью деревьев решений.
33. Алгоритмы деревьев решения.
34. Классификация методом деревьев решений.
35. Понятие кластерного анализа.
36. Иерархические и итеративные методы кластерного анализа.
37. Гипотеза компактности в кластерном анализе.
38. Пример кластерного анализа.
39. Постобработка результатов анализа.
40. Интерпретация результатов анализа.
41. Вычислительная сложность анализа больших данных.
42. Экспериментальное определение времени решения задачи анализа больших данных.

Типовые задачи:

1. Технология очистки данных в АП Логином.
2. Технология предобработки данных в АП Логином.
3. Технология заполнения пропусков и редактирования выбросов в АП Логином.
4. Технология трансформации данных в АП Логином.
5. Технология визуализации данных в АП Логином.
6. Решение задачи прогнозирования в АП Логином.
7. Решение задачи ассоциативного анализа в АП Логином.
8. Решение задачи выявления последовательностей в АП Логином.
9. Решение задачи классификации в АП Логином.
10. Решение задачи кластерного анализа в АП Логином.

Типовые задания для тестирования:

№ п.п Вопрос

Варианты ответа

1	Большие данные это	- данные, которые характеризуются низкой достоверностью а подчас и противоречивостью. - данные, которые характеризуются большим объемом и большой скоростью поступления. - данные, которые характеризуются большим объемом и большой скоростью поступления, неупорядоченностью данных и многообразием источников данных, низкой достоверностью а подчас и противоречивостью.
2	Задачи анализа больших данных решаются	- методами интеллектуальных систем и алгоритмов, а также методами математической статистики. - методами интеллектуальных систем и алгоритмов. -методами и средствами математической статистики.
3	Анализ больших данных в современном понимании, это технология решения задач анализа явлений	- для которых еще нет адекватных математических моделей, а есть только наборы экспериментальных данных. -для которых нет адекватных математических моделей. -методами имитационного моделирования.
4	Разведочный анализ данных использует	-классические методы визуального отображения данных. -многомерные кубы, диаграммы, гистограммы, карты, деревья и т.п. -самостоятельный анализ данных, извлеченных из источников.
5	Модели обнаружения знаний в данных	- осуществляют предварительную обработку данных от различных источников и применяют к ним интеллектуальные или статистические методы анализа. - осуществляют предварительную обработку данных от различных источников и применяют к ним интеллектуальные методы анализа. - осуществляют предварительную обработку данных от различных источников и применяют к ним статистические методы анализа.
6	В подавляющем большинстве случаев анализ больших данных сводится	- к применению одних и тех же базовых механизмов, распределенных по этапам. - к применению различных базовых механизмов, своих для каждого этапа анализа. - к применению классических методов статистического анализа..
7	Экспертные системы находят применение в тех областях, в которых	-знания представляют собой коллективный опыт. -знания представляют собой личный опыт специалистов высокого уровня. -знания являются неточными и противоречивыми. -верно все вышесказанное.

8	Информационный анализ данных осуществляется с помощью следующих механизмов	- CRM механизмов (механизмов ориентированных на клиентов) и ERP (механизмов планирования ресурсов предприятия). - OLAP - оперативной системы аналитической обработки данных. - верно все вышесказанное.
9	Системы интеллектуального анализа данных в основном используют информацию	- из баз данных. - из банков данных. - из хранилищ данных.
10	Хранилища данных это	-базы данных специальной структуры. -набор файлов специальной структуры. -набор банков данных специальной структуры.
11	В основе концепции хранилищ данных лежат идеи	- интеграции данных в едином хранилище. - тематическое и временное структурирование, согласование и агрегирование данных. - разделение данных, используемых операционной системой и наборов данных для решения задач анализа. - верно все вышесказанное.
12	OLAP куб создается из соединения таблиц	- с применением схемы звезды. - с применением схемы снежинки. - верно все вышесказанное.
13	OLAP куб представляет собой	-проекцию исходного многомерного куба данных на куб меньшей размерности. -проекцию исходного многомерного куба данных на куб большей размерности. -проекцию исходного многомерного куба данных на куб такой же размерности.
14	Сложные таблицы с большим количеством полей, следует сводить	-к нескольким более простым таблицам. -к таблице, данные в которой агрегированы. -верно все вышесказанное.
15	Интеллектуальный анализ данных это	- процесс обнаружения знаний в большом количестве сырых(необработанных) данных. - процесс обнаружения знаний в большом количестве сырых(необработанных) данных. - верно все вышесказанное
16	Интеллектуальный анализ данных состоит из следующих этапов	-подготовки данных, очистки данных и трансформации данных. -подготовки данных, очистки данных, трансформации данных, механизмов Data Mining и интерпретации данных. - Data Mining и интерпретации данных.
17	Этап подготовки данных заключается	-в создании набора данных, в том числе в консолидации данных из различных источников. - в определение выборки, которая будет в дальнейшем анализироваться. - верно все вышесказанное.
18	Очистка данных заключается	- в устранении пропусков, шумов и аномальных значений выборки.

			- в устранении пропусков и шумов значений выборки. - в устранении пропусков и аномальных значений выборки.
19	Трансформация данных заключается		- в подготовке данных для использования в числовом виде. - в подготовке данных для использования в категориальном виде. - в подготовке данных в виде воспринимаемых используемыми приложениями.
20	На этапе Data Mining решаются задачи		- построения математических моделей в интересах анализа ситуаций для которых таких моделей еще не существует. - определения параметров моделей общий вид которых известен. - правильно все вышесказанное.
21	К задачам Data Mining относятся задачи		- выявления ассоциаций и выявления последовательностей. - выявления ассоциаций, выявления последовательностей, классификации, кластеризации и прогнозирования. - классификации, кластеризации и прогнозирования.
22	Ассоциации указывают, что		- вместе с событием X с той или иной степенью вероятности происходит событие Y. - события X и Y взаимно независимы. - верно все вышесказанное.
23	Ассоциативный анализ предназначен для		- разработки ассоциативных правил. - для анализа ситуации с использованием ассоциативных правил. - верно все вышесказанное.
24	Последовательные шаблоны это		- ассоциативные правила, в которых важен порядок следования событий. - ассоциативные правила, в которых неважен порядок следования событий. - верно все вышесказанное.
25	В качестве критерия остановки процесса классификации может быть выбран критерий остановки		- по минимальному допустимому количеству объектов в выборке. - по точности на контрольной выборке. - верно все вышесказанное.
26	Классификация позволяет		- отнести объект к определенному классу, в зависимости от характеризующих его признаков. - определить вероятность принадлежности объекта к определенному классу, в зависимости от характеризующих его признаков. - верно все вышесказанное.
27	Классификация объектов может выполняться по величине		- энтропийного признака неоднородности. - индекса Джини. - верно все вышесказанное.
28	Дерево решений состоит из вершин		- двух типов. - трех типов. - четырех типов.
29	Для классификации объектов могут быть использованы		- ассоциативные правила. - деревья решений. - методы кластерного анализа

30	Обучение классификации есть	-обучение с учителем. - обучение без учителя. - верно все вышесказанное.
31	Кластеризация объектов позволяет	-объединить объекты в кластеры, когда сами кластеры не определены. - отнести объект к одному из известных кластеров. - определить вероятность отнесения объекта к одному из известных кластеров.
32	Для расчета расстояния между кластерами на диаграмме рассеяния могут быть использованы	- евклидово расстояние. - неевклидово расстояние. - квадратичное расстояние.
33	Объект относится к кластеру, если	- расстояние от объекта до центра кластера больше, чем радиус кластера. - расстояние от объекта до центра кластера меньше, чем радиус кластера.
34	Кластерный анализ опирается на	- гипотезу компактности. - гипотезу сравнимости шкал (в правильном выборе масштаба и единиц измерения признаков). - верны обе гипотезы.
35	Обучение кластеризации есть	-обучение с учителем. - обучение без учителя. - верно все вышесказанное.
36	Для кластеризации объектов могут быть использованы	-ассоциативные правила. -деревья решений. - методы классификации
37	Прогнозирование это	-особая форма предсказания, которая на основе статистики событий, имевших место в прошлом, прогнозирует ход и исход аналогичных событий в будущем. -особая форма предсказания, которая на основе экспертных оценок специалистов прогнозирует ход и исход событий в будущем. -особая форма предсказания, которая по приметам, имеющих место в настоящем, прогнозирует возникновение определенных событий в будущем.
38	Постобработка и интерпретация результатов анализа данных предназначена	-для тестирования полученных знаний. - для практического применения полученных знаний. .- верно все вышесказанное.
39	Быстродействие систем анализа данных зависит от	- вычислительной сложности алгоритмов. -способа программной реализации алгоритмов. - производительности и объема памяти используемых компьютеров. - от всех перечисленных факторов.
40	Для оценки быстродействия систем анализа данных может быть использован аппарат	- кластерного анализа. - ассоциативного анализа. - регрессионного анализа.

6.1.2. Промежуточной аттестации

Примерный перечень вопросов, выносимых на экзамен (семестр А)

Теоретические вопросы

1. Основные термины и определения теории анализа больших данных.
2. Анализ состояния, перспектив и организации развития технологии и систем обработки больших данных в МЧС.
3. Два подхода к анализу больших данных.
4. Базовые механизмы информационного и интеллектуального анализа больших данных.
5. Очистка и предобработка данных. Заполнение пропусков. Редактирование выбросов.
6. Очистка и предобработка данных. Заполнение пропусков. Редактирование выбросов.
7. Трансформация и визуализация данных.
8. Случайные величины и их выборочные оценки.
9. Парный регрессионный анализ.
10. Построение модели парной линейной регрессии в Excel.
11. Обратная модель парной линейной регрессии.
12. Фиктивная линейная связь и очистка переменных.
13. Построение модели парной нелинейной регрессии в Excel.
14. Регрессии нелинейные по объясняющим переменным.
15. Регрессии нелинейные по параметрам.
16. Модель множественной регрессии.
17. Проверка гипотез в отношении параметров регрессии
18. Проверка гипотезы в отношении коэффициента детерминации.
19. Технология прогнозирования событий по статистическим данным, имевших место в прошлом.
20. Решение задачи построения временных рядов.
21. Разработка математической модели прогноза временного ряда.
22. Решение задачи прогнозирования в среде аналитической платформы Логинум.
23. Ассоциативный анализ.
24. Алгоритмы ассоциативного анализа.
25. Решение задачи ассоциативного анализа с помощью алгоритма Apriori.
26. Основные показатели значимости ассоциативных правил.
27. Методика решения задачи ассоциативного анализа в АП Логинум.
28. Технология выявления последовательностей.
29. Методика решения задачи ассоциативного анализа в АП Логинум.
30. Деревья решений.

31. Методы классификации событий по известным значениям с помощью деревьев решений.
32. Алгоритмы классификации событий.
33. Индексы неоднородностей Джини, ошибочной классификации и энтропийный индекс.
34. Алгоритмы прогнозирования событий с помощью дерева решений.
35. Решение задачи классификации событий методом деревьев решений.
36. Методика решения задачи классификации событий в АП Логином методом деревьев решений.
37. Кластерный анализ.
38. Иерархические и итеративные методы кластерного анализа.
39. Алгоритмы кластерного анализа.
40. Методы выбора количества кластеров.
41. Пример использования кластерного анализа.
42. Методика решения задачи кластеризации событий в среде аналитической платформы Логином.
43. Постобработка и интерпретация результатов анализа.
44. Методика решения задачи постобработки и интерпретации результатов в АП Логином.
45. Методы оценки быстродействия систем анализа больших данных.
46. Методы определения вычислительной сложности системы анализа больших данных.

Практические вопросы

1. Технология очистки данных в АП Логином.
2. Технология предобработки данных в АП Логином.
3. Технология заполнения пропусков и редактирования выбросов в АП Логином.
4. Технология трансформации данных в АП Логином.
5. Технология визуализации данных в АП Логином.
6. Решение задачи прогнозирования в АП Логином.
7. Решение задачи ассоциативного анализа в АП Логином.
8. Решение задачи выявления последовательностей в АП Логином.
9. Решение задачи классификации в АП Логином.
10. Решение задачи кластерного анализа в АП Логином.
11. Решение задачи прогнозирования в АП Логином.

6.2. Шкала оценивания результатов промежуточной аттестации и критерии выставления оценок.

Форма контроля	Показатели оценивания	Критерии оценивания	Оценка
----------------	-----------------------	---------------------	--------

экзамен	правильность и полнота ответа	дан правильный, полный ответ на поставленный вопрос, показана совокупность осознанных знаний по дисциплине, доказательно раскрыты основные положения вопросов; могут быть допущены недочеты, исправленные самостоятельно в процессе ответа.	отлично
		дан правильный, недостаточно полный ответ на поставленный вопрос, показано умение выделить существенные и несущественные признаки, причинно-следственные связи; могут быть допущены недочеты, исправленные с помощью преподавателя.	хорошо
		дан недостаточно правильный и полный ответ; логика и последовательность изложения имеют нарушения; в ответе отсутствуют выводы.	удовлетворительно
		ответ представляет собой разрозненные знания с существенными ошибками по вопросу; присутствуют фрагментарность, нелогичность изложения; дополнительные и уточняющие вопросы не приводят к коррекции ответа на вопрос.	неудовлетворительно

7. Ресурсное обеспечение дисциплины

7.1. Лицензионное и свободно распространяемое программное обеспечение

Перечень лицензионного и свободно распространяемого программного обеспечения:

- Astra Linux Common Edition релиз Орел [ПО-25В-603] - Операционная система общего назначения "Astra Linux Common Edition" [Коммерческая (Full Package Product). Номер в Едином реестре российских программ для электронных вычислительных машин и баз данных - 4433]; Лицензия на право пользования № 217800111-ore-2.12-client-6196.

7.2. Профессиональные базы данных и информационные справочные системы

Системный анализ - https://systems-analysis.ru/systems_analysis.html

Системный анализ информационных систем <https://lektsii.org/7-94478.html>
Системный анализ информационно управляющих систем -
tudmed.ru/view/dondik-em-sistemnyy-analiz-informacionno-upravlyayuschih-sistem
Информационная система «Единое окно доступа к образовательным ресурсам» [Электронный ресурс]. – Режим доступа: <http://window.edu.ru/>, доступ только после самостоятельной регистрации
Научная электронная библиотека «eLIBRARY.RU» [Электронный ресурс]. – Режим доступа: <https://www.elibrary.ru/>, доступ только после самостоятельной регистрации

7.3. Литература

Основная литература: литература:

1. Параскевов, А. В. Большие данные: учебник / А. В. Параскевов, А. Э. Сергеев. — Москва, Вологда : Инфра-Инженерия, 2024. — 148 с. — ISBN 978-5-9729-2120-1. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/143597.html>

Дополнительная литература:

1. Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/81324.html>

2. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. — Москва : МИСИ-МГСУ, ЭБС АСВ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/101802.html>

7.4. Материально-техническое обеспечение

Для проведения и обеспечения лекционных занятий используются помещения, которые представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой бакалавриата, оснащенные оборудованием и техническими средствами обучения: автоматизированное рабочее место преподавателя, маркерная доска, мультимедийный проектор, документ-камера, посадочные места обучающихся.

Помещения для практических занятий и самостоятельной работы обучающихся оснащены компьютерной техникой из расчета 1 компьютер на одного обучающегося, с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде

университета.

Автор: Заслуженный работник высшей школы России, доктор технических наук, профессор Куватов Валерий Ильич.