

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«Луганский государственный университет
имени Владимира Даля»
(ФГБОУ ВО «ЛГУ им. В. Даля»)

Краснодонский факультет инженерии и менеджмента (филиал)
Кафедра информационных технологий и транспорта



УТВЕРЖДАЮ:

Директор
Панайотов К.К.

«21» апреля 2023 года

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

По дисциплине Интеллектуальный анализ данных
(название дисциплины по учебному плану)

По направлению подготовки 38.03.05 Бизнес-информатика
(код, название без кавычек)

Профиль подготовки Информационная бизнес-аналитика

Краснодон 2023

Лист согласования РПУД

Рабочая программа учебной дисциплины «Интеллектуальный анализ данных» по направлению подготовки 38.03.05 – Бизнес-информатика, профиль «Информационная бизнес-аналитика» - 36 с.

Рабочая программа учебной дисциплины «Интеллектуальный анализ данных» разработана в соответствии с Федеральным государственным образовательным стандартом высшего образования по направлению подготовки 38.03.05 Бизнес-информатика (утвержденный приказом Министерства образования и науки Российской Федерации от 29 июня 2020 года № 838, с изменениями и дополнениями от 26 ноября 2020 г)

СОСТАВИТЕЛЬ (СОСТАВИТЕЛИ):

к.э.н., доц. Попова Н.Н.

(ученая степень, ученое звание, должность фамилия, инициалы)

Рабочая программа дисциплины утверждена на заседании кафедры информационных технологий и транспорта «15» марта 2023 г., протокол № 7.

Заведующий кафедрой



Бихдрикер А.С.

Рекомендована на заседании учебно-методической комиссии факультета «20» марта 2023 г., протокол № 8.

Председатель учебно-методической
комиссии факультета



Замота О.Н.

Структура и содержание дисциплины

1. Цели и задачи дисциплины, ее место в учебном процессе

Целью дисциплины «Интеллектуальный анализ данных» является формирование у студентов комплекса теоретических знаний и методологических основ в области интеллектуального анализа данных, а также практических навыков, необходимых для внедрения и практического использования интеллектуальных алгоритмов анализа и обработки данных.

Задачи: изучение существующих технологий подготовки данных к анализу; изучение основных методов поиска в данных внутренних закономерностей, взаимосвязей, тенденций; овладение практическими умениями и навыками реализации технологий аналитической обработки данных, формирования и проверки гипотез о их природе и структуре, варьирования применяемыми моделями; формирование умений и навыков применения универсальных программных пакетов и аналитических платформ для анализа данных.

2. Место дисциплины в структуре ОПОП ВО

Дисциплина «Интеллектуальный анализ данных» относится к обязательной части цикла дисциплин.

Содержание дисциплины является логическим продолжением содержания дисциплин «Введение в специальность», «Бизнес-информатика», «Дискретный анализ в экономике», «Логические основы систем управления в экономике», «Математика», «Исследование операций».

Содержание дисциплины служит основой для освоения дисциплин «Математический инструментарий и модели оценки бизнеса», «Корпоративные информационные системы», «Проектирование социально-экономических систем», «Разработка и внедрение информационно-экономических систем», преддипломной практики и выпускной квалификационной работы.

3. Требования к результатам освоения содержания дисциплины

Код и наименование компетенции	Индикаторы достижений компетенции (по реализуемой дисциплине)	Перечень планируемых результатов
ОПК-6. Способен использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений	ОПК-6.2. Способен применять методы и алгоритмы интеллектуального анализа данных для поиска, выработки и применения новых решений в профессиональной деятельности	Знать: Методы и алгоритмы интеллектуального анализа данных
		Уметь: применять методы и алгоритмы интеллектуального анализа данных для поиска, выработки и применения новых решений в профессиональной деятельности
		Владеть: инструментальными средствами реализации методов и алгоритмов интеллектуального анализа данных

4. Структура и содержание дисциплины

4.1. Объем учебной дисциплины и виды учебной работы

Вид учебной работы	Объем часов (зач. ед.)		
	Очная форма	Очно-заочная форма	Заочная форма
Общая учебная нагрузка (всего)	108 (3 зач. ед)	108 (3 зач. ед)	-
Обязательная аудиторная учебная нагрузка (всего) в том числе:	68	24	-
Лекции	34	12	-
Семинарские занятия	-	-	-
Практические занятия	-	-	-
Лабораторные работы	34	12	-
Курсовая работа (курсовой проект)	-	-	-
Другие формы и методы организации образовательного процесса (<i>расчетно-графические работы, групповые дискуссии, ролевые игры, тренинг, компьютерные симуляции, интерактивные лекции, семинары, анализ деловых ситуаций и т.п.</i>)			-
Самостоятельная работа студента (всего)	40	84	-
Форма аттестации	зачет	зачет	-

4.2. Содержание разделов дисциплины

Тема 1. ВВЕДЕНИЕ В АНАЛИЗ ДАННЫХ

Введение. Основные понятия. Пирамида знаний. Постановка задачи анализа данных. История развития технологий анализа данных. Методика анализа данных. Методы сбора и подготовки исходного набора данных. Моделирование неопределённости. Вероятностный подход. Условные Байесовские сети.

Тема 2. ВВЕДЕНИЕ В ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ

Предмет и задачи интеллектуального анализа данных. Примеры практических приложений. KDD (Knowledge discovery in databases). Data Mining. Задачи классификации, кластеризации и регрессии. Оценка качества алгоритмов Data Mining. Метод скользящего контроля (кросс-валидация).

Тема 3. ВИЗУАЛЬНЫЙ АНАЛИЗ ДАННЫХ

Визуальный анализ данных. Применение различных средств визуализации.

Тема 4. АФФИНИТИВНЫЙ АНАЛИЗ. ПОИСК ПОСЛЕДОВАТЕЛЬНЫХ ШАБЛОНОВ

Введение в аффинитивный анализ (affinity analysis). Алгоритм a priori. Иерархические ассоциативные правила.

Тема 5. КЛАСТЕРНЫЙ АНАЛИЗ

Введение в кластеризацию. Классификация методов кластеризации. Алгоритм кластеризации k-means. Сети Кохонена (KCN – Kohonen network). Карты Кохонена (SOM – self organizing map). Проблемы алгоритмов кластеризации.

Тема 6. КЛАССИФИКАЦИЯ И РЕГРЕССИЯ. СТАТИСТИЧЕСКИЕ МЕТОДЫ

Введение в классификацию и регрессию. Простая линейная регрессия. Оценка соответствия простой линейной регрессии реальным данным. Простая регрессионная

модель. Множественная линейная регрессия. Модель множественной линейной регрессии. Регрессия с категориальными входными переменными. Методы отбора переменных в регрессионные модели. Ограничения применимости регрессионных моделей. Основы логистической регрессии. Интерпретация модели логистической регрессии.

Тема 7. КЛАССИФИКАЦИЯ И РЕГРЕССИЯ. МАШИННОЕ ОБУЧЕНИЕ

Введение в деревья решений. Алгоритмы построения деревьев решений. Алгоритмы ID3 и C4.5. Алгоритм CART. Упрощение деревьев решений. Введение в нейронные сети. Искусственный нейрон. Принципы построения нейронных сетей. Алгоритмы обучения нейронных сетей. Алгоритм обратного распространения ошибки.

Тема 8. НЕЙРОСЕТЕВЫЕ АЛГОРИТМЫ КЛАССИФИКАЦИИ

Понятия искусственных нейронных сетей. Перцептрон Розенблатта, правило обучения. Многослойные сети прямого распространения. Алгоритм обратного распространения ошибки. Машины опорных векторов. Вероятностные нейронные сети. Дилемма смещения и дисперсии.

Тема 9. АНАЛИЗ И ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ

Временной ряд и его компоненты. Модели прогнозирования. Прогнозирование в торговле и логистике.

Тема 10. АНСАМБЛИ МОДЕЛЕЙ

Введение в ансамбли моделей. Бэггинг. Бустинг. Альтернативные методы построения ансамблей.

Тема 11. ИСПОЛЬЗОВАНИЕ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ В ПЕРСПЕКТИВНЫХ НАПРАВЛЕНИЯХ РАЗВИТИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Обзор наиболее успешных продуктов, библиотек, toolkit, framework в выбранном направлении из области анализа данных. Вопросы: история развития, степень развития, назначение, функции, конкретные примеры использования, самостоятельное применение, лицензии, стандарты, применяемые средства анализа данных, предоставляемые API, совместимые технологии и т.д.

4.3. Лекции

№ п/п	Название темы	Объем часов		
		Очная форма	Очно- заочная форма	Заочная форма
1	Введение в анализ данных	3	1	-
2	Введение в интеллектуальный анализ данных.	3	1	-
3	Визуальный анализ данных	3	1	-
4	Аффинитивный анализ. Поиск последовательных шаблонов	3	1	-
5	Кластерный анализ	3	1	-
6	Классификация и регрессия. Статистические методы	3	1	-
7	Классификация и регрессия. Машинное обучение	3	1	-
8	Нейросетевые алгоритмы классификации	3	1	-
9	Анализ и прогнозирование временных рядов	3	1	-
10	Ансамбли моделей	3	1	-
11	Использование методов интеллектуального анализа данных в перспективных направлениях развития информационных технологии	4	2	-
Итого:		34	12	-

4.4. Практические (семинарские) занятия

Учебным планом не предусмотрены

4.4. Лабораторные работы

№ п/п	Название темы	Объем часов		
		Очная форма	Очно- заочная форма	Заочная форма
1	Введение в анализ данных	3	1	-
2	Введение в интеллектуальный анализ данных.	3	1	-
3	Визуальный анализ данных	3	1	-
4	Аффинитивный анализ. Поиск последовательных шаблонов	3	1	-
5	Кластерный анализ	3	1	-
6	Классификация и регрессия. Статистические методы	3	1	-
7	Классификация и регрессия. Машинное обучение	3	1	-
8	Нейросетевые алгоритмы классификации	3	1	-
9	Анализ и прогнозирование временных рядов	3	1	-
10	Ансамбли моделей	3	1	-
11	Использование методов интеллектуального анализа данных в перспективных направлениях развития информационных технологии	4	2	-
Итого:		34	12	-

4.6. Самостоятельная работа студентов

№ п/п	Название темы	Вид СРС	Объем часов		
			Очная форма	Очно- заочная форма	Заочная форма
1	Введение в анализ данных	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
2	Введение интеллектуальный анализ данных.	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
3	Визуальный анализ данных	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
4	Аффинитивный анализ. Поиск последовательных шаблонов	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
5	Кластерный анализ	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-

6	Классификация и регрессия. Статистические методы	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
7	Классификация и регрессия. Машинное обучение	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
8	Нейросетевые алгоритмы классификации	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
9	Анализ и прогнозирование временных рядов	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
10	Ансамбли моделей	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	3	7	-
11	Использование методов интеллектуального анализа данных в перспективных направлениях развития информационных технологии	Подготовка к лабораторным работам, к текущему и промежуточному контролю знаний и умений.	6	10	-
12	Зачет	Подготовка к зачету	4	4	-
Итого:			40	84	-

4.7. Курсовые работы/проекты.

Курсовая работа рабочим учебным планом не предусмотрена.

5. Образовательные технологии

Преподавание дисциплины ведется с применением следующих видов образовательных технологий:

Преподавание дисциплины ведется с применением следующих видов образовательных технологий:

Информационные технологии: использование электронных образовательных ресурсов (электронный конспект, размещенный во внутренней сети, или т.п.) при подготовке к лекциям и лабораторным работам.

Мультимедийные технологии: презентации к лекциям.

6. Учебно-методическое и информационное обеспечение дисциплины:

а) основная литература:

1. Чубукова И.А., Data Mining / Чубукова И.А. - М.: Национальный Открытый Университет "ИНТУИТ", 2016. (Основы информационных технологий) - ISBN 978-5-94774-819-2 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN9785947748192.html>.

2. Мельниченко А.С., Математическая статистика и анализ данных: учеб. пособие / А.С. Мельниченко - М.: МИСиС, 2018. - 45 с. - ISBN 978-5-906953-62-9 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN9785906953629.html>.

3. Кожаринов А.С., Моделирование и анализ информационных и бизнес-процессов в информационных системах: метод. указ. к выполнению курсовых работ / А.С. Кожаринов. - М.: МИСиС, 2017. - 27 с. - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: http://www.studentlibrary.ru/book/Misis_362.html.

б) дополнительная литература:

1. Кук Д., Машинное обучение с использованием библиотеки H2O / Кук Д. - М.: ДМК Пресс, 2018. - 250 с. - ISBN 978-5-97060-508-0 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN9785970605080.html>.

2. Джеймс Г., Введение в статистическое обучение с примерами на языке R / Джеймс Г., Уиттон Д., Хастингс Т., Тибширани Р. - М.: ДМК Пресс, 2017. - 456 с. - ISBN 978-5-97060-495-3 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN9785970604953.html>.

3. Рашка С., Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения / Рашка С. - М.: ДМК Пресс, 2017. - 418 с. - ISBN 978-5-97060-409-0 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN9785970604090.html>.

4. Афонин П.Н., Статистический анализ с применением современных программных средств: учебное пособие / Афонин П.Н., Афонин Д.Н. - СПб.: ИЦ Интермедия, 2017. - 100 с. - ISBN 978-4383-0080-9 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN978438300809.html>.

5. Тазиева Р.Ф., Обработка экспериментальных данных. Ч.1: учебное пособие: в 2 ч. / Р.Ф. Тазиева, А.Н. Титов - Казань: Издательство КНИТУ, 2017. - 96 с. - ISBN 978-5-7882-2261-5 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN9785788222615.html>.

6. Александровская Ю.П., Многомерный статистический анализ в экономике: учебное пособие / Александровская Ю. П. - Казань: Издательство КНИТУ, 2017. - 96 с. - ISBN 978-5-7882-2191-5 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN9785788221915.html>.

7. Шелудько В.М., Язык программирования высокого уровня Python. Функции, структуры данных, дополнительные модули: учебное пособие / Шелудько В. М. - Ростов н/Д: Изд-во ЮФУ, 2017. - 107 с. - ISBN 978-5-9275-2648-2 - Текст: электронный // ЭБС "Консультант студента": [сайт]. - URL: <http://www.studentlibrary.ru/book/ISBN9785927526482.html>.

в) методические указания:

1. Методические указания к лабораторным работам по дисциплине «Интеллектуальный анализ данных» для студентов направления подготовки 38.03.05 – Бизнес-информатика [Электронный ресурс] / сост. А.В. Велигура. – Луганск: ЛНУ им. В. Даля, 2019. – 75 с.

2. Методические указания к самостоятельной работе по дисциплине «Интеллектуальный анализ данных» для студентов направления подготовки 38.03.05 – Бизнес-информатика [Электронный ресурс] / сост. А.В. Велигура. – Луганск: ЛНУ им. В. Даля, 2019. – 26 с.

3. Методические указания к практическим занятиям по дисциплине «Интеллектуальный анализ данных». Занятие 1. Проектирование и наполнение хранилища данных [Электронный ресурс] / сост.: А.В. Велигура, Н.Н. Попова. – Луганск: ЛНУ им. В. Даля, 2017. – 18 с.

г) Интернет-ресурсы:

1. Министерство образования и науки Российской Федерации – <http://минобрнауки.рф/>
2. Федеральная служба по надзору в сфере образования и науки – <http://obrnadzor.gov.ru/>
3. Министерство образования и науки Луганской Народной Республики – <https://minobr.su>
4. Министерство промышленности и торговли Луганской Народной Республики – <https://www.minpromlnr.su/main.php/>
5. Министерство экономического развития Луганской Народной Республики – <https://merlnr.su/>
6. Министерство финансов Луганской Народной Республики – <https://minfinlnr.su/>
7. Народный совет Луганской Народной Республики – <https://nslnr.su>
8. Государственный комитет статистики Луганской Народной Республики – <https://www.gkslnr.su/>
9. Портал Федеральных государственных образовательных стандартов высшего образования – <http://fgosvo.ru/>
10. Федеральный портал «Российское образование» – <http://www.edu.ru/>
11. Информационная система «Единое окно доступа к образовательным ресурсам» – <http://window.edu.ru/>
12. Федеральный центр информационно-образовательных ресурсов – <http://fcior.edu.ru/>
13. Статистические сборники ВШЭ - <https://www.hse.ru/org/hse/primarydata/>
14. OpenOffice.org: Теория и практика - <https://www.altlinux.org/Books:Openoffice>
15. Базовый курс по OpenOffice - <https://4creates.com/training/49-bazovyy-kurs-po-openoffice.html>

Электронные библиотечные системы и ресурсы

16. Электронно-библиотечная система «Консультант студента» – <http://www.studentlibrary.ru/>
17. Электронно-библиотечная система «StudMed.ru» – <https://www.studmed.ru/>

Информационный ресурс библиотеки образовательной организации

18. Научная библиотека имени А. Н. Коняева – <http://biblio.dahluniver.ru/>

7. Материально-техническое и программное обеспечение дисциплины

Для проведения лекционных занятий требуется аудитория, оборудованная мультимедийным проектором с экраном. Для проведения лабораторных и практических занятий необходим компьютерный класс и презентационная техника (проектор, экран, компьютер/ноутбук)

Программное обеспечение:

Функциональное назначение	Бесплатное программное обеспечение	Ссылки
Офисный пакет	Libre Office 6.3.1	https://www.libreoffice.org/

		https://ru.wikipedia.org/wiki/LibreOffice
Операционная система	UBUNTU 19.04	https://ubuntu.com/
Браузер	Firefox Mozilla	http://www.mozilla.org/ru/firefox/fx
Графический редактор	GIMP (GNU Image Manipulation Program)	http://www.gimp.org/ http://gimp.ru/viewpage.php?page_id=8 http://ru.wikipedia.org/wiki/GIMP
Редактор PDF	Adobe Acrobat Reader	https://get.adobe.com/ru/reader/
Аудиоплеер	VLC	http://www.videolan.org/vlc/

9. Оценочные средства по дисциплине

Паспорт оценочных средств по учебной дисциплине «Интеллектуальный анализ данных»

Перечень компетенций (элементов компетенций), формируемых в результате освоения учебной дисциплины (модуля) или практики

№ п/п	Код контролируемой компетенции	Формулировка контролируемой компетенции	Индикаторы достижений компетенции (по реализуемой дисциплине)	Контролируемые темы учебной дисциплины, практики	Этапы формирования (семестр изучения)
5	ОПК-6	Способен использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений	ОПК-6.2	Тема 1 Тема 2 Тема 3 Тема 4 Тема 5 Тема 6 Тема 7 Тема 8 Тема 9 Тема 10 Тема 11	5

Показатели и критерии оценивания компетенций, описание шкал оценивания

№ п/п	Код контролируемой компетенции	Индикаторы достижений компетенции (по реализуемой дисциплине)	Перечень планируемых результатов	Контролируемые разделы (темы) учебной дисциплины (модуля), практики	Наименование оценочного средства ²
3	ОПК-6	ОПК-6.2	Знать: Методы и алгоритмы интеллектуального анализа данных Уметь: применять методы и алгоритмы интеллектуального анализа данных для поиска, выработки и применения новых решений в профессиональной деятельности Владеть: инструментальными и средствами реализации методов и алгоритмов интеллектуального анализа данных	Тема 1 Тема 2 Тема 3 Тема 4 Тема 5 Тема 6 Тема 7 Тема 8	Собеседование (устный или письменный опрос), тесты

Оценочные средства по дисциплине «Интеллектуальный анализ данных»

Перечень вопросов (для проведения собеседования (устный или письменный опрос))

1. Область применения хранилищ данных.
2. Основные требования к данным, вводимым в хранилище данных.
3. Возможности создания хранилищ данных в реляционных БД.
4. Понятие витрин данных.
5. Понятие многомерного представления при описании структур данных.
6. Алгоритм добычи данных - метод деревьев решений.
7. Алгоритм добычи данных – кластеризация.
8. Технология аналитической обработки данных (OLAP).
9. Варианты реализации OLAP.
10. Распределенные корпоративные хранилища данных.
11. Зависимые и независимые киоски данных.
12. Характеристика виртуальных кубов.
13. Архитектура хранилища HOLAP.
14. Архитектура хранилища ROLAP.
15. Архитектура хранилища MOLAP.
16. Особенности ETL-процесса.
17. Методы и этапы процесса очистки данных.

18. Задачи Data Mining. Классификация задач DataMining.
19. Задача классификации и регрессии.
20. Задача поиска ассоциативных правил.
21. Задача кластеризации.
22. Модели DataMining.
23. Методы DataMining.
24. Возможности современных программ, используемых для создания хранилищ данных.
25. Характеристика информационной технологии принятия решений.
26. Роль хранилища данных в информационной технологии принятия решений.
27. 12 правил Кодда оценки средств для OLAP.
28. Построение кубов данных. SQL Server Analysis Services 2005. Инструменты разработки и администрирования, клиентские инструменты.
29. Построение кубов данных. Использование Business Intelligence Development Studio для создания проекта.
30. Построение кубов данных. Создание источника данных.
31. Построение кубов данных. Создание представления источника данных (DSV).
32. Построение кубов данных. Создание куба с помощью мастера.
33. Построение кубов данных. Развертывание и просмотр куба.
34. Построение реляционной витрины данных.
35. Заполнение реляционной витрины данных.
36. Настройка витрины данных.
37. Построение OLAP-срезов.
38. OLAP-средства программы MS Excel.
39. Создание OLAP-куба в среде MS Excel.
40. Средства анализа OLAP-куба.
41. Создание OLAP-куба с помощью Deductor.
42. Обработка OLAP-куба в Deductor. Создание отчетов.
43. Классификация ПО Data Mining.
44. Обзор и анализ ПО по бизнес-характеристикам и техническим характеристикам.
45. Подбор ПО для решения задач оценивания и прогнозирования.
46. Подбор ПО для решения задач классификации.
47. Подбор ПО для решения задач кластеризации и сегментации.
48. Подбор ПО для поиска и ассоциативных правил.
49. Из каких частей состоит Deductor?
50. Какие варианты поставки Deductor существуют?
51. Чем отличается версия Professional от Academic?
52. Имеются ли ограничения по количеству обрабатываемых записей в версии Deductor Academic?
53. Сколько категорий пользователей Deductor можно выделить?
54. Перечислите функции аналитика.
55. Кто обычно занимается проектированием и наполнением хранилища данных?
56. Каким образом лицензируется Deductor?
57. Что такое ассоциативные правила? Приведите пример.
58. В чём отличие поиска ассоциативных правил и секвенциального анализа?
59. Как можно модифицировать алгоритм Apriori?
60. Для чего используются ассоциативные правила при анализе данных?
61. Что такое достоверность правила?
62. Что такое поддержка правила?
63. Какой алгоритм генерации ассоциативных правил имеется в Deductor?
64. Какие входные поля набора данных необходимы для запуска обработчика Ассоциативные правила в Deductor?

65. Какие специализированные визуализаторы предлагаются к узлу обработчику Ассоциативные правила?

66. Какой визуализатор используется для формирования предложений клиенту?
67. С помощью, каких кнопок можно отсортировать правила по убыванию лифта?
68. Дайте определение понятия «кластер».
69. Дайте определение понятия «кластеризация».
70. В чём заключается задача кластеризации? Приведите пример.
71. Как представляются результаты при кластеризации?
72. На какие два больших класса можно поделить алгоритмы кластеризации.
73. Охарактеризуйте два любых алгоритма кластеризации.
74. Назовите входные данные алгоритма k-средних.
75. Назовите выходные данные алгоритма k - средних.
76. Приведите последовательность шагов в алгоритме k - средних.
77. Приведите порядок кластерного анализа с помощью ПО kmeans.

Критерии и шкала оценивания по оценочному средству собеседование (устный или письменный опрос)

Шкала оценивания (интервал баллов)	Критерий оценивания
5	собеседование (устный или письменный опрос) прошел на высоком уровне (студент в полном объеме осветил рассматриваемый вопрос, владеет профильным понятийным (категориальным) аппаратом и т.п.)
4	собеседование (устный или письменный опрос) прошел на среднем уровне (студент в целом осветил рассматриваемый вопрос, владеет профильным понятийным (категориальным) аппаратом и т.п.)
3	собеседование (устный или письменный опрос) на низком уровне (студент допустил существенные неточности, изложил материал с ошибками, не владеет в достаточной степени профильным категориальным аппаратом и т.п.)
2	собеседование (устный или письменный опрос) прошел на неудовлетворительном уровне или не представлен (студент не готов, не выполнил задание и т.п.)

Тесты

1. Какие науки включены в Data Mining?
 - a. статистика, базы данных, искусственный интеллект;
 - b. информатика, базы данных, статистика;
 - c. искусственный интеллект, базы данных, базы знаний;
 - d. информатика, базы данных, хранилища данных.
2. Каким образом технология Data Mining используется в интернет?
 - a. для создания сайтов;
 - b. для организации поисковых систем;
 - c. для отображения web-страниц.
3. Какие задачи решаются в СППР?
 - a. ввод данных, преобразование данных, вывод данных;
 - b. ввод данных, модификация данных, передача данных;
 - c. ввод данных, хранение данных, анализ данных.

4. Какой класс задач анализа отвечает за выполнение запросов?
 - a. информационно-поисковый;
 - b. оперативно-аналитический;
 - c. интеллектуальный.
5. Какой класс задач анализа отвечает за обобщение данных?
 - a. информационно-поисковый;
 - b. оперативно-аналитический;
 - c. интеллектуальный.
6. Какой класс задач анализа отвечает за построение моделей?
 - a. информационно-поисковый;
 - b. оперативно-аналитический;
 - c. интеллектуальный.
7. Какой класс задач анализа отвечает за поиск закономерностей в данных?
 - a. информационно-поисковый;
 - b. оперативно-аналитический;
 - c. интеллектуальный.
8. Какая подсистема СППР отвечает за ввод данных?
 - a. OLTP;
 - b. хранилище данных;
 - c. SQL;
 - d. OLAP;
 - e. Data Mining.
9. Какая подсистема СППР отвечает за хранение данных?
 - a. OLTP;
 - b. хранилище данных;
 - c. SQL;
 - d. OLAP;
 - e. Data Mining.
10. Какая подсистема СППР отвечает за информационно-поисковый анализ данных?
 - a. OLTP;
 - b. хранилище данных;
 - c. SQL;
 - d. OLAP;
 - e. Data Mining.
11. Какая подсистема СППР отвечает за оперативный анализ данных?
 - a. OLTP;
 - b. хранилище данных;
 - c. SQL;
 - d. OLAP;
 - e. Data Mining.
12. Какая подсистема СППР отвечает за интеллектуальный анализ данных?
 - a. OLTP;
 - b. хранилище данных;
 - c. SQL;
 - d. OLAP;
 - e. Data Mining.
13. Как реализуется подсистема ввода данных?
 - a. с помощью технологии Data Mining;
 - b. с помощью базы данных;
 - c. с помощью СУБД;
 - d. с помощью хранилища данных;
 - e. с помощью витрины данных.

14. Какие данные могут храниться в системе анализа?
 - a. детализированные;
 - b. обобщенные;
 - c. детализированные и обобщенные.
15. Могут ли в системе анализа храниться данные в разных форматах?
 - a. могут;
 - b. не могут;
 - c. могут только данные в текстовых форматах.
16. Какие данные могут храниться в системе анализа?
 - a. нормализованные;
 - b. избыточные (денормализованные);
 - c. частично нормализованные.
17. Как называются структуры данных, предназначенные для ввода, модификации и поиска?
 - a. оперативные источники данных;
 - b. хранилища данных;
 - c. базы данных;
 - d. файлы данных.
18. Как называются структуры данных, предназначенные для анализа?
 - a. оперативные источники данных;
 - b. хранилища данных;
 - c. базы данных;
 - d. витрины данных.
19. Предметно-ориентированный, интегрированный, неизменяемый, поддерживающий хронологию набор данных, организованный для целей поддержки принятия решений – это
 - a. оперативный источник данных;
 - b. хранилище данных;
 - c. база данных;
 - d. файл данных.
20. Что является главным недостатком виртуального хранилища данных?
 - a. большое время обработки запросов;
 - b. значительные ресурсы компьютера;
 - c. разные форматы и кодировки данных в разных ОИД;
 - d. практическая невозможность получения данных за длительный период времени.
21. Как называются данные, непосредственно переносимые из ОИД?
 - a. метаданные;
 - b. агрегированные данные;
 - c. детальные.
22. Какие данные отражают сущность события?
 - a. измерения;
 - b. метаданные;
 - c. факты.
23. Какие данные описывают события?
 - a. измерения;
 - b. метаданные;
 - c. факты.
24. Как называются числовые фактические данные, которые могут быть просуммированы по всем измерениям?
 - a. аддитивные;
 - b. полуаддитивные;

с. неаддитивные.

25. Как называются числовые фактические данные, которые могут быть просуммированы только по определенным измерениям?

- а. аддитивные;
- б. полуаддитивные;
- с. неаддитивные.

26. Как называются фактические данные, которые не могут быть просуммированы ни по одному измерению?

- а. аддитивные;
- б. полуаддитивные;
- с. неаддитивные.

27. На какие вопросы должны отвечать метаданные?

- а. что, кто, где, как, когда, почему;
- б. что, кто, зачем, как, когда, почему;
- с. что, кто, где, как, по какой причине, почему;
- д. что, кто, где, как, когда, сколько.

28. Репозиторий – это

- а. словарь терминов;
- б. хранилище метаданных;
- с. каталог с файлами.

29. Если поток образуется данными, копируемыми из ОИД, то он называется

- а. входным потоком;
- б. потоком обобщения;
- с. архивным потоком;
- д. потоком метаданных;
- е. обратным потоком.

30. Если поток образуется агрегированными данными, то он называется

- а. входным потоком;
- б. потоком обобщения;
- с. архивным потоком;
- д. потоком метаданных;
- е. обратным потоком.

31. Если поток образуется детальными данными, количество обращений к которым снизилось, то он называется

- а. входным потоком;
- б. потоком обобщения;
- с. архивным потоком;
- д. потоком метаданных;
- е. обратным потоком.

32. Если поток образуется данными, переносимыми в репозиторий, то он называется

- а. входным потоком;
- б. потоком обобщения;
- с. архивным потоком;
- д. потоком метаданных;
- е. обратным потоком.

34. Если поток образуется очищенными данными, записываемыми в ОИД, то он называется

- а. входным потоком;
- б. потоком обобщения;
- с. архивным потоком;
- д. потоком метаданных;
- е. обратным потоком.

35. Многомерный анализ – это:
- а. одновременный анализ по нескольким измерениям;
 - б. одновременный анализ по нескольким параметрам;
 - с. одновременный анализ по нескольким данным
36. Формирование подмножества многомерного массива данных, соответствующего единственному значению одного или нескольких элементов измерений, не входящих в это подмножество, называется
- а. гиперкубом;
 - б. срезом гиперкуба;
 - с. базой данных;
 - д. витриной данных.
37. Технология оперативной аналитической обработки данных, использующая методы и средства для сбора, хранения и анализа многомерных данных в целях поддержки процессов принятия решений, называется
- а. OLTP;
 - б. OLAP;
 - с. Data Mining.
38. Если для реализации многомерной модели используют многомерные базы данных, то способ реализации гиперкуба называется
- а. MOLAP;
 - б. ROLAP;
 - с. HOLAP;
 - д. DOLAP;
 - е. JOLAP.
39. Если для реализации многомерной модели используют реляционные базы данных, то способ реализации гиперкуба называется
- а. MOLAP;
 - б. ROLAP;
 - с. HOLAP;
 - д. DOLAP;
 - е. JOLAP.
40. Если для реализации многомерной модели используют и многомерные, и реляционные базы данных, то способ реализации гиперкуба называется
- а. MOLAP;
 - б. ROLAP;
 - с. HOLAP;
 - д. DOLAP;
 - е. JOLAP.
41. Настольная OLAP, предназначенная для локального анализа и представления данных, называется
- а. MOLAP;
 - б. ROLAP;
 - с. HOLAP;
 - д. DOLAP;
 - е. JOLAP.
42. OLAP, предназначенная для создания и управления данными и метаданными, называется
- а. MOLAP;
 - б. ROLAP;
 - с. HOLAP;
 - д. DOLAP;
 - е. JOLAP.

43. В каком отношении находятся таблица фактов и таблица измерений?

- a. «один-к-одному»;
- b. «один-ко-многим»;
- c. «многие-ко-многим».

44. Исследование и обнаружение машиной (алгоритмами, средствами искусственного интеллекта) в сырых данных скрытых знаний, которые: ранее не были известны, нетривиальны, практически полезны, доступны для интерпретации человеком, называется

- a. OLTP;
- b. хранилищем данных;
- c. OLAP;
- d. Data Mining.

45. Какие операции над данными включены в ETL-процесс?

- a. ввод, модификация, вывод;
- b. чтение, изменение, запись;
- c. извлечение, преобразование, загрузка;
- d. получение, хранение, анализ.

46. Гиперкуб – это ...

- a. объект, все измерения которого имеют одинаковую размерность;
- b. поликуб;
- c. объект, все измерения которого имеют разную размерность;
- d. многомерный куб;
- e. многомерная база данных.

47 Многомерный просмотр данных основан на ...

- a. многомерной базе данных;
- b. технологии мультимедиа;
- c. многослойной базе;
- d. сетевой технологии;
- e. гипертекстовой технологии.

48. Какую технологию используют большинство хранилища данных?

- a. концептуальную БД;
- b. реляционную БД;
- c. иерархическую БД;
- d. физическую БД.

49. Какое конструирование у хранилищ данных?

- a. физическое;
- b. логическое;
- c. логическое и физическое;
- d. иерархическое.

50. Концептуальную модель хранилища данных можно представить в виде...

- a. таблицы;
- b. графического рисунка;
- c. геометрических фигур;
- d. схемы.

51. К классу описательных задач относятся:

- a. кластеризация и классификация;
- b. кластеризация и поиск ассоциативных правил;
- c. классификация и регрессия;
- d. классификация и поиск ассоциативных правил.

52. К классу предсказательных задач относятся:

- a. кластеризация и классификация;
- b. кластеризация и поиск ассоциативных правил;

- с. классификация и регрессия;
 - d. классификация и поиск ассоциативных правил.
- 53. К классу задач supervised learning (обучение с учителем) относятся:
 - a. кластеризация и классификация;
 - b. кластеризация и поиск ассоциативных правил;
 - с. классификация и регрессия;
 - d. классификация и поиск ассоциативных правил.
- 54. К классу задач unsupervised learning (обучение без учителя) относятся:
 - a. кластеризация и классификация;
 - b. кластеризация и поиск ассоциативных правил;
 - с. классификация и регрессия;
 - d. классификация и поиск ассоциативных правил.
- 55. Задача классификации сводится к ...
 - a. нахождения частых зависимостей между объектами или событиями;
 - b. определения класса объекта по его характеристиками;
 - с. определение по известным характеристиками объекта значение некоторого его параметра;
 - d. поиска независимых групп и их характеристик в всем множестве анализируемых данных.
- 56. Задача регрессии сводится к ...
 - a. нахождения частых зависимостей между объектами или событиями;
 - b. определению класса объекта по его характеристиками;
 - с. определению по известным характеристиками объекта значения некоторого его параметра;
 - d. поиска независимых групп и их характеристик в всем множестве анализируемых данных.
- 57. Задача кластеризации заключается в ...
 - a. нахождения частых зависимостей между объектами или событиями;
 - b. определения класса объекта по его характеристиками;
 - с. определение по известным характеристиками объекта значение некоторого его параметра;
 - d. поиска независимых групп и их характеристик в всем множестве анализируемых данных.
- 8. Целью поиска ассоциативных правил является ...
 - a. нахождения частых зависимостей между объектами или событиями;
 - b. определения класса объекта по его характеристиками;
 - с. определение по известным характеристиками объекта значение некоторого его параметра;
 - d. поиска независимых групп и их характеристик в всем множестве анализируемых данных.
- 59. В описательных моделях относятся следующие модели данных:
 - a. модели классификации и последовательностей;
 - b. регрессивные, кластеризации, исключений, итоговые и ассоциации;
 - с. классификации, кластеризации, исключений, итоговые и ассоциации;
 - d. модели классификации, последовательностей и исключений.
- 60. Модели классификации описывают ...
 - a. правила или набор правил в соответствии с которыми можно отнести описание любого нового объекта к одному из классов;
 - b. функции, которые позволяют прогнозировать изменения непрерывных числовых параметров;
 - с. функциональные зависимости между зависимыми и независимыми показателями и переменными в понятной человеку форме;

d. группы, на которые можно разделить объекты, данные о которых подвергаются анализу.

61. Модели последовательностей описывают ...

a. правила или набор правил в соответствии с которыми можно отнести описание любого нового объекта к одному из классов;

b. функции, которые позволяют прогнозировать изменения непрерывных числовых параметров;

c. функциональные зависимости между зависимыми и независимыми показателями и переменными в понятной человеку форме;

d. группы, на которые можно разделить объекты, данные о которых подвергаются анализу.

62. Регрессивные модели описывают ...

a. правила или набор правил в соответствии с которыми можно отнести описание любого нового объекта к одному из классов;

b. функции, которые позволяют прогнозировать изменения непрерывных числовых параметров;

c. функциональные зависимости между зависимыми и независимыми показателями и переменными в понятной человеку форме;

d. группы, на которые можно разделить объекты, данные о которых подвергаются анализу.

63. Какова основная цель создания и использования хранилищ данных:

a. анализ данных для принятия управленческих решений;

b. надежное хранение, накопленных данных;

c. резервное копирование данных.

64. OLAP - это:

a. технология онлайн-быстрой аналитической обработки сложных запросов к базе данных;

b. технология онлайн-обработки небольших по размерам, но идущих большим потоком транзакций, требующих максимально быстрого ответа.

65. Какие из вариантов ответов являются характерными требованиями к хранению данных для принятия решений в хранилищах данных?

a. данные ориентированы на приложения;

b. данные управляются транзакциями;

c. данные обобщены либо очищены.

66. Перечислите основные этапы работы с хранилищами данных?

a. этап очистки данных;

b. этап обновления;

c. этап нормализации.

67. Что называют кубом OLAP?

a. структуру, в которой хранятся совокупности данных, полученные путем всех возможных сочетаний измерений в таблице измерений;

b. структуру, в которой хранятся совокупности данных, полученные путем всех возможных сочетаний измерений в таблице фактов;

c. таблицу размерностей.

68. Информационные хранилища созданы для удобства ...

a. руководителей всех уровней для принятия решений;

b. стратегического планирования;

c. реорганизации бизнеса;

d. предметных приложений;

e. редактирования данных.

69. Информационные хранилища размещаются на ...

a. библиотеках-автоматах;

- b. сетевых серверах;
 - c. мейнфреймах;
 - d. серверах и кластерах серверов;
 - e. файл-серверах.
70. При слиянии данных в информационное хранилище из внутренних и внешних источников обеспечивается ...
- a. предметная ориентация данных;
 - b. выбор требуемых сведений из предметных приложений по наименованиям;
 - c. гипертекстовый просмотр данных;
 - d. согласование данных по наименованию;
 - e. хранение данных по предметным областям.
71. Интеллектуальный выбор данных из информационного хранилища – это ...
- a. реализация методов искусственного интеллекта;
 - b. выбор по заданному алгоритму;
 - c. реализация самообучающихся систем;
 - d. реализация экономико-статистических методов.
72. В процессе погружения в информационное хранилище данные ...
- a. очищаются от ненужной для анализа информации;
 - b. агрегируются;
 - c. преобразуются из разных типов данных предметных приложений в единую структуру хранения;
 - d. индексируются;
 - e. синхронизируются.
73. Какова основная цель создания и использования хранилищ данных:
- a. анализ данных для принятия управленческих решений;
 - b. надежное хранение накопленных данных;
 - c. резервное копирование данных.
74. Классификация — ...
- a. некоторый набор операций над базой данных, который рассматривается как единственное завершённое, с точки зрения пользователя, действие над некоторой информацией, обычно связано с обращением к базе данных;
 - b. разновидность систем хранения, ориентированная на поддержку процесса анализа данных, обеспечивающая непротиворечивость и хронологию данных, а также высокую скорость выполнения аналитических запросов;
 - c. высокоуровневые средства отражения информационной модели и описания структуры данных;
 - d. это установление зависимости дискретной выходной переменной от входных переменных.
75. Регрессия — ...
- a. это установление зависимости непрерывной выходной переменной от входных переменных;
 - b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;
 - c. выявление закономерностей между связанными событиями;
 - d. это установление зависимости дискретной выходной переменной от входных переменных.
76. Кластеризация — ...
- a. это установление зависимости непрерывной выходной переменной от входных переменных;
 - b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;
 - c. выявление закономерностей между связанными событиями;

d. это установление зависимости дискретной выходной переменной от входных переменных.

77. Ассоциация — ...

a. это установление зависимости непрерывной выходной переменной от входных переменных;

b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;

c. выявление закономерностей между связанными событиями;

d. это установление зависимости дискретной выходной переменной от входных переменных.

78. Машинное обучение — ...

a. специализированное программное решение (или набор решений), который включает в себя все инструменты для извлечения закономерностей из сырых данных;

b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;

c. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат;

d. подразделение искусственного интеллекта, изучающее методы построения алгоритмов, способных обучаться на данных.

79. Аналитическая платформа — ...

a. специализированное программное решение (или набор решений), который включает в себя все инструменты для извлечения закономерностей из сырых данных;

b. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;

c. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат;

d. подразделение искусственного интеллекта, изучающее методы построения алгоритмов, способных обучаться на данных.

80. Обучающая выборка — ...

a. это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;

b. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр и соответствующий ему правильный выходной результат;

c. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат;

d. выявление в сырых данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

81. Ошибка обучения — ...

a. это ошибка, допущенная моделью на учебном множестве;

b. это ошибка, полученная на тестовых примерах, то есть вычисляется по тем же формулам, но для тестового множества;

c. имена, типы, метки и назначения полей исходной выборки данных;

d. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат.

82. Ошибка обобщения — ...

- a. это ошибка, допущенная моделью на учебном множестве;
- b. это ошибка, полученная на тестовых примерах, то есть вычисляется по тем же формулам, но для тестового множества;
- c. имена, типы, метки и назначения полей исходной выборки данных;
- d. набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат.

Критерии и шкала оценивания по оценочному средству тесты

Шкала оценивания (интервал баллов)	Критерий оценивания
5	Тесты выполнены на высоком уровне (правильные ответы даны на 90-100% тестов)
4	Тесты выполнены на среднем уровне (правильные ответы даны на 75-89% тестов)
3	Тесты выполнены на низком уровне (правильные ответы даны на 50-74% тестов)
2	Тесты выполнены на неудовлетворительном уровне (правильные ответы даны менее чем на 50% тестов)

Оценочные средства для промежуточной аттестации (зачет)

Вопросы к зачету

1. Какие науки включены в Data Mining?
 - a. статистика, базы данных, искусственный интеллект;
 - b. информатика, базы данных, статистика;
 - c. искусственный интеллект, базы данных, базы знаний;
 - d. информатика, базы данных, хранилища данных.
2. Какой класс задач анализа отвечает за обобщение данных?
 - a. информационно-поисковый;
 - b. оперативно-аналитический;
 - c. интеллектуальный.
3. Какая подсистема СППР отвечает за хранение данных?
 - a. OLTP;
 - b. хранилище данных;
 - c. SQL;
 - d. OLAP;
 - e. Data Mining.
4. Как реализуется подсистема ввода данных?
 - a. с помощью технологии Data Mining;
 - b. с помощью базы данных;
 - c. с помощью СУБД;
 - d. с помощью хранилища данных;
 - e. с помощью витрины данных.
5. Как называются структуры данных, предназначенные для ввода, модификации и поиска?
 - a. оперативные источники данных;
 - b. хранилища данных;
 - c. базы данных;
 - d. файлы данных.
6. Как называются данные, непосредственно переносимые из ОИД?

- a. метаданные;
 - b. агрегированные данные;
 - c. детальные.
7. Как называются числовые фактические данные, которые могут быть просуммированы только по определенным измерениям?
- d. аддитивные;
 - e. полуаддитивные;
 - f. неаддитивные.
8. Если поток образуется данными, копируемыми из ОИД, то он называется
- a. входным потоком;
 - b. потоком обобщения;
 - c. архивным потоком;
 - d. потоком метаданных;
 - e. обратным потоком.
9. Многомерный анализ – это:
- a. одновременный анализ по нескольким измерениям;
 - b. одновременный анализ по нескольким параметрам;
 - c. одновременный анализ по нескольким данным.
10. Если для реализации многомерной модели используют реляционные базы данных, то способ реализации гиперкуба называется
- a. MOLAP;
 - b. ROLAP;
 - c. HOLAP;
 - d. DOLAP;
 - e. JOLAP.
11. В каком отношении находятся таблица фактов и таблица измерений?
- a. «один-к-одному»;
 - b. «один-ко-многим»;
 - c. «многие-ко-многим».
12. Многомерный просмотр данных основан на ...
- a. многомерной базе данных;
 - b. технологии мультимедиа;
 - c. многослойной базе;
 - d. сетевой технологии;
 - e. гипертекстовой технологии.
13. К классу описательных задач относятся:
- a. кластеризация и классификация;
 - b. кластеризация и поиск ассоциативных правил;
 - c. классификация и регрессия;
 - d. классификация и поиск ассоциативных правил.
14. Задача классификации сводится к ...
- a. нахождения частых зависимостей между объектами или событиями;
 - b. определения класса объекта по его характеристикам;
 - c. определение по известным характеристикам объекта значения некоторого его параметра;
 - d. поиска независимых групп и их характеристик в всем множестве анализируемых данных.
15. В описательных моделях относятся следующие модели данных:
- a. модели классификации и последовательностей;
 - b. регрессивные, кластеризации, исключений, итоговые и ассоциации;
 - c. классификации, кластеризации, исключений, итоговые и ассоциации;
 - d. модели классификации, последовательностей и исключений.

16. Какова основная цель создания и использования хранилищ данных:
- анализ данных для принятия управленческих решений;
 - надежное хранение, накопленных данных;
 - резервное копирование данных.
17. Что называют кубом OLAP?
- структуру, в которой хранятся совокупности данных, полученные путем всех возможных сочетаний измерений в таблице измерений;
 - структуру, в которой хранятся совокупности данных, полученные путем всех возможных сочетаний измерений в таблице фактов;
 - таблицу размерностей.
18. Интеллектуальный выбор данных из информационного хранилища – это ...
- реализация методов искусственного интеллекта;
 - выбор по заданному алгоритму;
 - реализация самообучающихся систем;
 - реализация экономико-статистических методов.
19. Регрессия — ...
- это установление зависимости непрерывной выходной переменной от входных переменных;
 - это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;
 - выявление закономерностей между связанными событиями;
 - это установление зависимости дискретной выходной переменной от входных переменных.
20. Аналитическая платформа — ...
- специализированное программное решение (или набор решений), который включает в себя все инструменты для извлечения закономерностей из сырых данных;
 - это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;
 - набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат;
 - подразделение искусственного интеллекта, изучающее методы построения алгоритмов, способных обучаться на данных.
21. Каким образом технология Data Mining используется в интернет?
- для создания сайтов;
 - для организации поисковых систем;
 - для отображения web-страниц.
22. Какой класс задач анализа отвечает за построение моделей?
- информационно-поисковый;
 - оперативно-аналитический;
 - интеллектуальный.
23. Какая подсистема СППР отвечает за информационно-поисковый анализ данных?
- OLTP;
 - хранилище данных;
 - SQL;
 - OLAP;
 - Data Mining.
24. Какие данные могут храниться в системе анализа?
- детализированные;
 - обобщенные;
 - детализированные и обобщенные.

25. Как называются структуры данных, предназначенные для анализа?
- а. оперативные источники данных;
 - б. хранилища данных;
 - в. базы данных;
 - г. витрины данных.
26. Какие данные отражают сущность события?
- а. измерения;
 - б. метаданные;
 - в. факты.
27. Как называются фактические данные, которые не могут быть просуммированы ни по одному измерению?
- а. аддитивные;
 - б. полуаддитивные;
 - в. неаддитивные.
28. Если поток образуется агрегированными данными, то он называется
- а. входным потоком;
 - б. потоком обобщения;
 - в. архивным потоком;
 - г. потоком метаданных;
 - д. обратным потоком.
29. К классу предсказательных задач относятся:
- а. кластеризация и классификация;
 - б. кластеризация и поиск ассоциативных правил;
 - в. классификация и регрессия;
 - г. классификация и поиск ассоциативных правил.
30. Задача регрессии сводится к ...
- а. нахождения частых зависимостей между объектами или событиями;
 - б. определения класса объекта по его характеристиками;
 - в. определение по известным характеристиками объекта значение некоторого его параметра;
 - г. поиска независимых групп и их характеристик в всем множестве анализируемых данных.
31. Модели классификации описывают ...
- а. правила или набор правил в соответствии с которыми можно отнести описание любого нового объекта к одному из классов;
 - б. функции, которые позволяют прогнозировать изменения непрерывных числовых параметров;
 - в. функциональные зависимости между зависимыми и независимыми показателями и переменными в понятной человеку форме;
 - г. группы, на которые можно разделить объекты, данные о которых подвергаются анализа.
32. OLAP - это:
- а. технология онлайн-быстрой аналитической обработки сложных запросов к базе данных;
 - б. технология онлайн-обработки небольших по размерам, но идущих большим потоком транзакций, требующих максимально быстрого ответа.
33. Информационные хранилища созданы для удобства ...
- а. руководителей всех уровней для принятия решений;
 - б. стратегического планирования;
 - в. реорганизации бизнеса;
 - г. предметных приложений;
 - д. редактирования данных.

38. В процессе погружения в информационное хранилище данные ...
- очищаются от ненужной для анализа информации;
 - агрегируются;
 - преобразуются из разных типов данных предметных приложений в единую структуру хранения;
 - индексируются;
 - синхронизируются.
39. Кластеризация — ...
- это установление зависимости непрерывной выходной переменной от входных переменных;
 - это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;
 - выявление закономерностей между связанными событиями;
 - это установление зависимости дискретной выходной переменной от входных переменных.
40. Обучающая выборка — ...
- это группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов;
 - набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр и соответствующий ему правильный выходной результат;
 - набор данных, каждая запись которого представляет собой учебный пример, содержащий заданный входной параметр, которому отвечает правильный выходной результат;
 - выявление в сырых данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

Критерии и шкала оценивания итоговой аттестации

Характеристика знания предмета и ответов	Зачеты
Студент глубоко и в полном объёме владеет программным материалом. Грамотно, исчерпывающе и логично его излагает в устной или письменной форме. При этом знает рекомендованную литературу, проявляет творческий подход в ответах на вопросы и правильно обосновывает принятые решения, хорошо владеет умениями и навыками при выполнении практических задач.	зачтено
Студент знает программный материал, грамотно и по сути излагает его в устной или письменной форме, допуская незначительные неточности в утверждениях, трактовках, определениях и категориях или незначительное количество ошибок. При этом владеет необходимыми умениями и навыками при выполнении практических задач.	
Студент знает только основной программный материал, допускает неточности, недостаточно чёткие формулировки, непоследовательность в ответах, излагаемых в устной или письменной форме. При этом недостаточно владеет умениями и навыками при выполнении практических задач. Допускает до 30% ошибок в излагаемых ответах.	
Студент не знает значительной части программного материала. При этом допускает принципиальные ошибки в доказательствах, в трактовке понятий и категорий, проявляет низкую культуру знаний, не владеет основными умениями и навыками при выполнении практических задач. Студент отказывается от ответов на дополнительные вопросы.	не зачтено

9. Особенности организации обучения для лиц с ограниченными возможностями здоровья и инвалидов

При необходимости рабочая программа учебной дисциплины может быть адаптирована для обеспечения образовательного процесса инвалидов и лиц с ограниченными возможностями здоровья, в том числе с применением электронного обучения и дистанционных образовательных технологий.

Для этого требуется заявление студента (его законного представителя) и заключение психолого-медико-педагогической комиссии (ПМПК). В случае необходимости обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося), а для инвалидов также в соответствии с индивидуальной программой реабилитации инвалида могут предлагаться следующие варианты восприятия учебной информации с учетом их индивидуальных психофизических особенностей:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников, например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения – аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной (модулем), за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение электронного обучения, дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение электронного обучения, дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи:
 - продолжительность сдачи зачёта или экзамена, проводимого в письменной форме, – не более чем на 90 минут;
 - продолжительность подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, – не более чем на 20 минут; – продолжительность выступления обучающегося при защите курсовой работы – не более чем на 15 минут.

Лист изменений и дополнений

№ п/п	Виды дополнений и изменений	Дата и номер протокола заседания кафедры (кафедр), на котором были рассмотрены и одобрены изменения и дополнения	Подпись (с расшифровкой) заведующего кафедрой (заведующих кафедрами)