

Документ подписан простой электронной подписью.
Информация о владельце:

ФИО: Кандрашина Елена Александровна

Должность: И.о. ректора ФГАОУ ВО «Самарский государственный экономический университет»

Дата подписания: 22.07.2024 14:06:31

Уникальный программный ключ:

2db64eb9605ce27edd3b8e8fdd32c70e0674ddd2

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Самарский государственный экономический университет»

Институт Институт экономики предприятий

Кафедра Прикладной информатики

УТВЕРЖДЕНО

Ученым советом Университета

(протокол № 10 от 30 мая 2024 г.)

РАБОЧАЯ ПРОГРАММА

Наименование дисциплины	Б1.В.ДЭ.03.01 Качество данных, подходы и инструменты
Основная профессиональная образовательная программа	09.04.03 Прикладная информатика программа Искусственный интеллект и большие данные

Квалификация (степень) выпускника магистр

Самара 2024

Содержание (рабочая программа)

Стр.

- 1 Место дисциплины в структуре ОП
- 2 Планируемые результаты обучения по дисциплине, обеспечивающие достижение планируемых результатов обучения по программе
- 3 Объем и виды учебной работы
- 4 Содержание дисциплины
- 5 Материально-техническое и учебно-методическое обеспечение дисциплины
- 6 Фонд оценочных средств по дисциплине

Целью изучения дисциплины является формирование результатов обучения, обеспечивающих достижение планируемых результатов освоения образовательной программы.

1. Место дисциплины в структуре ОП

Дисциплина Качество данных, подходы и инструменты входит в часть, формируемая участниками образовательных отношений (дисциплина по выбору) блока Б1.Дисциплины (модули)

Предшествующие дисциплины по связям компетенций: Интеллектуальное планирование, Методы машинного обучения, Методы оптимизации

Последующие дисциплины по связям компетенций: Современные методы проектирования систем искусственного интеллекта, Тестирование искусственного интеллекта, Машинное обучение на больших данных, Массово параллельные вычисления для ускорения машинного обучения, Интеллектуальные информационные системы, Анализ прикладных систем

2. Планируемые результаты обучения по дисциплине, обеспечивающие достижение планируемых результатов обучения по программе

Изучение дисциплины Качество данных, подходы и инструменты в образовательной программе направлено на формирование у обучающихся следующих компетенций:

Профессиональные компетенции (ПК):

ПК-5 - Способен выявлять новые риски, отслеживать существующие риски для понимания того, что все риски выявлены и мероприятия по работе с ними выполняются и эффективны

Планируемые результаты обучения по программе	Планируемые результаты обучения по дисциплине		
ПК-5	ПК-5.1: Знать:	ПК-5.2: Уметь:	ПК-5.3: Владеть (иметь навыки):
	типы рисков и особенности управления рисками, связанными с реализацией ИТ-проектов	Выявлять новые риски, отслеживать существующие риски для понимания того, что все риски выявлены и мероприятия по работе с ними выполняются и эффективны	навыками выявления новых рисков, отслеживания существующих рисков для понимания того, что все риски выявлены и мероприятия по работе с ними выполняются и эффективны

ПК-6 - Способен инициировать запросы на изменение (в том числе корректирующие действия, предупреждающие действия, запросы на исправление несоответствий)

Планируемые результаты обучения по программе	Планируемые результаты обучения по дисциплине		
ПК-6	ПК-6.1: Знать:	ПК-6.2: Уметь:	ПК-6.3: Владеть (иметь навыки):
	особенности инициации запросов на изменение	инициировать запросы на изменение (в том числе корректирующие действия, предупреждающие действия, запросы на исправление несоответствий)	навыками инициирования запросов на изменение (в том числе корректирующие действия, предупреждающие действия, запросы на исправление несоответствий)

3. Объем и виды учебной работы

Учебным планом предусматриваются следующие виды учебной работы по дисциплине:

Очная форма обучения

Виды учебной работы	Всего час/ з.е.
	Сем 2
Контактная работа, в том числе:	12.15/0.34
Занятия лекционного типа	4/0.11
Занятия семинарского типа	8/0.22
Индивидуальная контактная работа (ИКР)	0.15/0
Самостоятельная работа:	113.85/3.16
Промежуточная аттестация	18/0.5
Вид промежуточной аттестации: Зачет	Зач
Общая трудоемкость (объем части образовательной программы): Часы	144
Зачетные единицы	4

заочная форма

Виды учебной работы	Всего час/ з.е.
	Сем 3
Контактная работа, в том числе:	12.15/0.34
Занятия лекционного типа	4/0.11
Занятия семинарского типа	8/0.22
Индивидуальная контактная работа (ИКР)	0.15/0
Самостоятельная работа:	113.85/3.16
Промежуточная аттестация	18/0.5
Вид промежуточной аттестации: Зачет	Зач
Общая трудоемкость (объем части образовательной программы): Часы	144
Зачетные единицы	4

4. Содержание дисциплины

4.1. Разделы, темы дисциплины и виды занятий:

Тематический план дисциплины Качество данных, подходы и инструменты представлен в таблице.

Разделы, темы дисциплины и виды занятий

Очная форма обучения

№ п/п	Наименование темы (раздела) дисциплины	Контактная работа					Планируемые результаты обучения в соотношении с результатами обучения по образовательной программе
		Лекции	Занятия семинарского типа	ИКР	ГКР	Самостоятельная работа	
			Практич. занятия				
1.	Аналитические информационные системы. Компоненты обработки данных.	2	4	0.075		50	ПК-5.1, ПК-5.2, ПК -5.3, ПК-6.1, ПК- 6.2, ПК-6.3
2.	Визуализация и аналитическая отчетность	2	4	0.075		63.85	ПК-5.1, ПК-5.2, ПК -5.3, ПК-6.1, ПК- 6.2, ПК-6.3
	Контроль	18					
	Итого	4	8	0.15		113.8	

					5	
--	--	--	--	--	---	--

заочная форма

№ п/п	Наименование темы (раздела) дисциплины	Контактная работа				Самостоятельная работа	Планируемые результаты обучения в соотношении с результатами обучения по образовательной программе
		Лекции	Занятия семинарского типа	ИКР	ГКР		
			Практич. занятия				
1.	Аналитические информационные системы. Компоненты обработкиданных.	2	4	0.075		50	ПК-5.1, ПК-5.2, ПК -5.3, ПК-6.1, ПК- 6.2, ПК-6.3
2.	Визуализация и аналитическаяотчетность	2	4	0.075		63.85	ПК-5.1, ПК-5.2, ПК -5.3, ПК-6.1, ПК- 6.2, ПК-6.3
	Контроль	18					
	Итого	4	8	0.15		113.8 5	

4.2 Содержание разделов и тем

4.2.1 Контактная работа

Тематика занятий лекционного типа

№п/п	Наименование темы (раздела) дисциплины	Вид занятия лекционного типа*	Тематика занятия лекционного типа
1.	Аналитические информационные системы. Компоненты обработки данных.	лекция	Аналитические информационные системы. Компоненты обработки данных.
2.	Визуализация и аналитическая отчетность	лекция	Визуализация и аналитическая отчетность

*лекции и иные учебные занятия, предусматривающие преимущественную передачу учебной информации педагогическими работниками организации и (или) лицами, привлекаемыми организацией к реализации образовательных программ на иных условиях, обучающимся

Тематика занятий семинарского типа

№п/п	Наименование темы (раздела) дисциплины	Вид занятия семинарского типа**	Тематика занятия семинарского типа
1.	Аналитические информационные системы. Компоненты обработки данных.	практическое занятие	Очистка и предобработка данных
		практическое занятие	Трансформация данных
2.	Визуализация и аналитическая отчетность	практическое занятие	Визуализация и аналитическая отчетность
		практическое занятие	

** семинары, практические занятия, практикумы, лабораторные работы, коллоквиумы и иные аналогичные занятия

Иная контактная работа

При проведении учебных занятий СГЭУ обеспечивает развитие у обучающихся навыков командной работы, межличностной коммуникации, принятия решений, лидерских качеств (включая при необходимости проведение интерактивных лекций, групповых дискуссий, ролевых игр, тренингов, анализ ситуаций и имитационных моделей, преподавание дисциплин (модулей) в форме курсов, составленных на основе результатов научных исследований, проводимых

организацией, в том числе с учетом региональных особенностей профессиональной деятельности выпускников и потребностей работодателей).

Формы и методы проведения иной контактной работы приведены в Методических указаниях по основной профессиональной образовательной программе.

4.2.2 Самостоятельная работа

№п/п	Наименование темы (раздела) дисциплины	Вид самостоятельной работы ***
1.	Аналитические информационные системы. Компоненты обработки данных.	- подготовка доклада - подготовка электронной презентации - тестирование
2.	Визуализация и аналитическая отчетность	- подготовка доклада - подготовка электронной презентации - тестирование

*** самостоятельная работа в семестре, написание курсовых работ, докладов, выполнение контрольных работ

5. Материально-техническое и учебно-методическое обеспечение дисциплины

5.1 Литература:

Основная литература

1. Миркин, Б. Г. Базовые методы анализа данных : учебник и практикум для вузов / Б. Г. Миркин. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2024. — 303 с. — (Высшее образование). — ISBN 978-5-534-18842-4. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/551786>

Дополнительная литература

1. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2024. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/536007>

5.2. Перечень лицензионного программного обеспечения

1. Astra Linux Special Edition «Смоленск», «Орел»; РедОС
2. МойОфис Стандартный 2, МойОфис Образование, Р7-Офис Профессиональный

5.3 Современные профессиональные базы данных, к которым обеспечивается доступ обучающихся

1. Профессиональная база данных «Информационные системы Министерства экономического развития Российской Федерации в сети Интернет» (Портал «Официальная Россия» - <http://www.gov.ru/>)
2. Государственная система правовой информации «Официальный интернет-портал правовой информации» (<http://pravo.gov.ru/>)
3. Профессиональная база данных «Финансово-экономические показатели Российской Федерации» (Официальный сайт Министерства финансов РФ - <https://www.minfin.ru/ru/>)
4. Профессиональная база данных «Официальная статистика» (Официальный сайт Федеральной службы государственной статистики - <http://www.gks.ru/>)

5.4. Информационно-справочные системы, к которым обеспечивается доступ обучающихся

1. Справочно-правовая система «Консультант Плюс»
2. Справочно-правовая система «ГАРАНТ-Максимум»

Учебные аудитории для проведения занятий лекционного типа	Комплекты ученической мебели Мультимедийный проектор
---	---

	Доска Экран
Учебные аудитории для проведения практических занятий (занятий семинарского типа)	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ
Учебные аудитории для групповых и индивидуальных консультаций	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ
Учебные аудитории для текущего контроля и промежуточной аттестации	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ
Помещения для самостоятельной работы	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ
Помещения для хранения и профилактического обслуживания оборудования	Комплекты специализированной мебели для хранения оборудования

5.6 Лаборатории и лабораторное оборудование

Лаборатория информационных технологий в профессиональной деятельности	Комплекты ученической мебели Мультимедийный проектор Доска Экран Компьютеры с выходом в сеть «Интернет» и ЭИОС СГЭУ Лабораторное оборудование
---	--

6. Фонд оценочных средств по дисциплине Качество данных, подходы и инструменты:

6.1. Контрольные мероприятия по дисциплине

Вид контроля	Форма контроля	Отметить нужное знаком « + »
Текущий контроль	Оценка докладов	+
	Устный/письменный опрос	+
	Тестирование	+
	Практические задачи	+
	Оценка контрольных работ (для заочной формы обучения)	+
Промежуточный контроль	Зачет	+

Порядок проведения мероприятий текущего и промежуточного контроля определяется Методическими указаниями по основной профессиональной образовательной

программе высшего образования; Положением о балльно-рейтинговой системе оценки успеваемости обучающихся по основным образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры в федеральном государственном автономном образовательном учреждении высшего образования «Самарский государственный экономический университет».

6.2. Планируемые результаты обучения по дисциплине, обеспечивающие достижение планируемых результатов обучения по программе

Профессиональные компетенции (ПК):

ПК-5 - Способен выявлять новые риски, отслеживать существующие риски для понимания того, что все риски выявлены и мероприятия по работе с ними выполняются и эффективны

Планируемые результаты обучения по программе	Планируемые результаты обучения по дисциплине		
	ПК-5.1: Знать:	ПК-5.2: Уметь:	ПК-5.3: Владеть (иметь навыки):
	типы рисков и особенности управления рисками, связанными с реализацией ИТ-проектов	Выявлять новые риски, отслеживать существующие риски для понимания того, что все риски выявлены и мероприятия по работе с ними выполняются и эффективны	навыками выявления новых рисков, отслеживания существующих рисков для понимания того, что все риски выявлены и мероприятия по работе с ними выполняются и эффективны
Пороговый	систему экономических процессов и явлений в условиях риска, основные теоретические и эконометрические модели, положения применения эконометрических моделей, основные нормативно- правовые документы	пользоваться современными техническими средствами и информационными технологиями	навыками и современными техническими средствами для самостоятельного, методически правильного решения аналитических и исследовательских заданий и задач,
Стандартный (в дополнение к пороговому)	основные методы решения аналитических и исследовательских задач	принимать адекватные решения при построении эконометрических моделей в условиях риска	навыками построения стандартных эконометрических моделей в условиях риска
Повышенный (в дополнение к пороговому, стандартному)	современные технические средства и информационные технологии, используемые при решении исследовательских задач в условиях риска	обосновывать все виды рисков и анализировать проведённые расчеты	методами анализа и содержательно интерпретировать полученные результаты

ПК-6 - Способен инициировать запросы на изменение (в том числе корректирующие действия, предупреждающие действия, запросы на исправление несоответствий)

Планируемые результаты обучения по	Планируемые результаты обучения по дисциплине
------------------------------------	---

программе			
	ПК-6.1: Знать:	ПК-6.2: Уметь:	ПК-6.3: Владеть (иметь навыки):
	особенности инициации запросов на изменение	инициировать запросы на изменение (в том числе корректирующие действия, предупреждающие действия, запросы на исправление несоответствий)	навыками инициирования запросов на изменение (в том числе корректирующие действия, предупреждающие действия, запросы на исправление несоответствий)
Пороговый	особенности работы со специализированными программными библиотеками языков программирования для анализа данных и решения задач машинного обучения	составлять композиции моделей, проводить отбор признаков	навыками использования профессиональных программ для решения типовых задач
Стандартный (в дополнение к пороговому)	основные понятия теории вероятностей и математической статистики; методик и расчетов, используемые при анализе данных	использовать инструменты описательной статистики и визуализации данных, вероятностные и статистические методы для решения типовых задач	Навыками анализа данных на примере решения задач сегментации, классификации, прогнозирования
Повышенный (в дополнение к пороговому, стандартному)	вероятностные и статистические методы, принципы предварительной обработки, визуализации и анализа данных, вероятностные и статистические методы:	использовать инструменты описательной статистики и визуализации данных, вероятностные и статистические методы для решения прикладных задач	навыками оценивания качества данных при работе с проектами среднего и малого уровня сложности

6.3. Паспорт оценочных материалов

№ п/п	Наименование темы (раздела) дисциплины	Контролируемые планируемые результаты обучения в соотношении с результатами обучения по программе	Вид контроля/используемые оценочные средства	
			Текущий	Промежуточный
1.	Аналитические информационные системы. Компоненты обработки данных.	ПК-5.1, ПК-5.2, ПК- 5.3, ПК-6.1, ПК-6.2, ПК-6.3	Оценка контрольных работ (для заочной формы обучения) Тестирование	Зачет
2.	Визуализация и аналитическая отчетность	ПК-5.1, ПК-5.2, ПК- 5.3, ПК-6.1, ПК-6.2, ПК-6.3	Оценка контрольных работ (для заочной формы	Зачет

			обучения) Тестирование	
--	--	--	---------------------------	--

6.4. Оценочные материалы для текущего контроля

Примерная тематика докладов

Раздел дисциплины	Темы
Аналитические информационные системы. Компоненты обработки данных.	Сопровождение данных Этапы и задачи курирования данных Подходы к оценке качества данных Подходы к очистке данных Происхождение данных Качество открытых данных
Визуализация и аналитическая отчетность	Визуализация данных Основные принципы визуализации Способы визуализации данных Виды диаграмм Инструменты и сервисы для визуализации

Вопросы для устного/письменного опроса

Раздел дисциплины	Вопросы
Аналитические информационные системы. Компоненты обработки данных.	1 Анализ данных. Основные понятия и определения 2 Бизнес-аналитика. Основные понятия и определения 3 Концепция хранилища данных. Понятие хранилища данных 4 Многомерная модель данных 5-6 Интеграция данных и бизнес-аналитика 7-8 Интеграция данных 9 Хранилища данных 10 Процессы информативной корпоративной фабрики 11 Базовые архитектуры корпоративной информационной фабрики 12 Визуализация бизнес-аналитики 13 Технологии бизнес-аналитики 14 Аналитические платформы. Инструменты бизнес-аналитики 15 Большие данные. 16. Наука о данных
Визуализация и аналитическая отчетность	17. Методы визуализации. 18. Цели и задачи визуализации 19. Визуализация Источников данных 20. Методы визуализации 21. Визуализаторы общего назначения 22. Сложные визуализаторы общего назначения 23. OLAP-анализ 24. Манипуляции с измерениями

Задания для тестирования по дисциплине для оценки сформированности компетенций (min 20, max 50 + ссылку на ЭИОС с тестами) <https://lms2.sseu.ru/course/index.php?categoryid=1863>

1. Аналитик это ...
 - а) специалист в области анализа и моделирование
 - б) специалист в предметной области;
 - в) человек, решающий определенные задачи;
 - г) человек, который имеет опыт в программировании.
- 2 Эксперт это ...
 - а) специалист в области анализа и моделирование;
 - б) специалист в предметной области;
 - в) человек, решать определенные задачи;

г) человек, который имеет опыт в программировании.

3 Задача классификации 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.

4 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.

5 Задача кластеризации заключается в ...

- 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.

6 Целью поиска ассоциативных правил является ...

- 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.

7 До предполагаемых моделей относятся такие модели данных:

- а) модели классификации и последовательностей;
- б) регрессивные, кластеризации, исключений, итоговые и ассоциации;
- в) классификации, кластеризации, исключений, итоговые и ассоциации;
- г) модели классификации, последовательностей и исключений.

8 В описательных моделей относятся следующие модели данных:

- а) модели классификации и последовательностей;
- б) регрессивные, кластеризации, исключений, итоговые и ассоциации;
- в) классификации, кластеризации, исключений, итоговые и ассоциации;
- г) модели классификации, последовательностей и исключений.

9 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.

10 Модели последовательностей 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.

11 Регрессивные модели 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.

12. Виды лингвистической неопределенности:

- а) неточность измерений значений определенной величины, выполняемых

физическими приборами;

б) неопределенность значений слов (Многозначность, размытость,

непонятность, нечеткость); неоднозначность смысла фраз (Синтаксическая и семантическая);

в) случайность (или наличие в внешней среде нескольких возможностей, каждая из которых случайным образом может стать действительностью); неопределенность значений слов (многозначность, размытость, неясность, нечеткость)

г) неоднозначность смысла фраз (Синтаксическая и семантическая).

13. Модели исключений описывают ...

а) исключительные ситуации в записях, которые резко отличаются произвольной признаку от основной множества записей;

б) ограничения на данные анализируемого массива;

в) закономерности между связанными событиями;

г) группы, на которые можно разделить объекты, данные о которых подвергаются анализа.

14 Итоговые модели обнаружат ...

а) исключительные ситуации в записях, которые резко отличаются произвольной признаку от основной множества записей;

б) ограничения на данные анализируемого массива;

в) закономерности между связанными событиями;

г) группы, на которые можно разделить объекты, данные о которых подвергаются анализа.

15 Модели ассоциации проявляют ...

а) исключительные ситуации в записях, которые резко отличаются произвольной признаку от основной множества записей;

б) ограничения на данные анализируемого массива;

в) закономерности между связанными событиями;

г) группы, на которые можно разделить объекты, данные о которых подвергаются анализа.

16 Виды физической неопределенности данных:

а) неточность измерений значений определенной величины, выполняемых

физическими приборами; случайность (или наличие в внешней среде нескольких возможностей, каждая из которых случайным образом может стать действительностью)

б) неопределенность значений слов (Многозначность, размытость, непонятность, нечеткость); неоднозначность смысла фраз (Синтаксическая и семантическая);

в) случайность (или наличие в внешней среде нескольких возможностей, каждая из которых случайным образом может стать действительностью); неопределенность значений слов (многозначность, размытость, неясность, нечеткость);

г) неоднозначность смысла фраз (Синтаксическая и семантическая).

17 Очистка данных — ...

а) комплекс методов и процедур, направленных на устранение причин, мешающих корректной обработке: аномалий, пропусков, дубликатов, противоречий, шумов и т.д.

б) процесс дополнения данных некоторой информацией, позволяющей повысить эффективность развязку аналитических задач

в) объект, содержащий структурированные данные, которые могут оказаться полезными для развязку аналитического задачи

г) комплекс методов и процедур, направленных на извлечение данных из различных источников, обеспечение необходимого уровня их информативности и качества, преобразования в единый формат, в котором они могут быть загружены в хранилище данных или аналитическую систему

18 Обогащение — ...

а) комплекс методов и процедур, направленных на устранение причин, мешающих корректной обработке: аномалий, пропусков, дубликатов, противоречий, шумов и т.д.

б) процесс дополнения данных некоторой информацией, позволяющей повысить эффективность развязку аналитических задач

в) объект, содержащий структурированные данные, которые могут оказаться полезными для развязку аналитического задачи

г) комплекс методов и процедур, направленных на извлечение данных из различных источников, обеспечение необходимого уровня их информативности и качества, преобразования в единый формат, в котором они могут быть загружены в хранилище данных или аналитическую систему.

19 Консолидация — ...

а) комплекс методов и процедур, направленных на устранение причин, мешающих корректной обработке: аномалий, пропусков, дубликатов, противоречий, шумов и т.д.

б) процесс дополнения данных некоторой информацией, позволяющей повысить эффективность развязку аналитических задач

в) объект, содержащий структурированные данные, которые могут оказаться полезными для развязку аналитического задачи

г) комплекс методов и процедур, направленных на извлечение данных из различных источников, обеспечение необходимого уровня их информативности и качества, преобразования в единый формат, в котором они могут быть загружены в хранилище данных или аналитическую систему

20 Транзакция — ...

а) некоторый набор операций над базой данных, который рассматривается как единственное завершено, с точки зрения пользователя, действие над некоторой информацией, обычно связано с обращением к базе данных

б) разновидность систем хранения, ориентирована на поддержку процесса анализа данных целостность, обеспечивает, непротиворечивость и хронологию данных, а также высокую скорость выполнения аналитических запросов

в) высокоуровневые средства отражения информационной модели и описания структуры данных

г) это установление зависимости дискретной выходной переменной от входных переменных

Практические задачи (min 20, max 50 + ссылку на ЭИОС с электронным изданием, если имеется)

Раздел дисциплины	Задачи
Аналитические информационные системы. Компоненты обработки данных.	Необходимо построить сценарий в аналитической платформе Deductor/Loginom, который ежедневно формирует отчет, показывающий 10 лидеров продаж по сумме продаж по итогам последних 10 дней. Результат необходимо визуализировать в виде упорядоченной по убыванию столбчатой диаграммы. Необходимо произвести визуализацию в двух видах и сформировать соответствующие отчеты, а именно, с отображением сумм продаж, с отображением долей продаж в общей сумме продаж этих товаров. Исходные данные находятся в файле «Продажи», характеристики товаров в файле «Товары».
Визуализация и аналитическая отчетность	Вы работаете в небольшой туристической фирме и планируете массовую рассылку рекламного буклета. Ваши средства ограничены, поэтому вы хотите послать ее тем, кто готов тратить на путешествия и отдых в большей степени. В файле P5_1.XLS содержатся данные о случайной выборке клиентов размером 925 (пол, возраст, суммы, затраченные на путешествия и отдых в предыдущем году). Используйте данные, чтобы понять, насколько пол и возраст влияют на объем затрат. Сформулируйте обоснованные рекомендации относительно контингента для рассылки рекламной брошюры. 2) Основываясь на данных о продажах из файла «Продажи» и других сопутствующих справочниках сформировать сценарий ежедневного отчета по 5 лидерам товарных групп по суммам продаж за последние 15 дней. Отчет визуализировать с помощью столбчатых диаграмм с информацией о сумме продаж и названии товарных групп.

Тематика контрольных работ

Раздел дисциплины	Темы
Аналитические информационные системы. Компоненты обработки данных.	Разработчики аналитических систем и платформ.
Визуализация и аналитическая отчетность	Аналитические платформы.

6.5. Оценочные материалы для промежуточной аттестации

Фонд вопросов для проведения промежуточного контроля в форме зачета

Раздел дисциплины	Вопросы
Аналитические информационные системы. Компоненты обработки данных.	1 Факторный анализ. 2 Дискриминантный анализ. 3 Кластерный анализ. 4 Многомерное шкалирование. 5 Методы контроля качества. 6 Основные направления развития методов обработки и хранения данных. 7 Volume. 8 Закон Мура. 9 Velocity. Variety. 10 Фреймворк Hadoop. 11 Проблема хранения неструктурированных данных. 12 Проблема преобразования данных. 13 Семантические анализаторы. 14 Самообучающиеся автоматы. 15 Языки для Big Data: R. 16 Языки для Big Data: Python. 17 Языки для Big Data: Julia. 18 Языки для Big Data: Java. 19 Языки для Big Data: Scala. 20 Языки для Big Data: MATLAB. 21 Языки для Big Data: Kafka. 22 Языки для Big Data: Hadoop. 23 Языки для Big Data: Go. 24 Фреймворки для Big Data: Hadoop. 25 Фреймворки для Big Data: Spark. 26 Фреймворки для Big Data: Storm. 27 Базы данных для Big Data: Hive. 28 Базы данных для Big Data: Impala. 29 Базы данных для Big Data: Presto. 30 Базы данных для Big Data: Drill. 31 Аналитические платформы для Big Data: Rapid Miner 32 Аналитические платформы для Big Data: IBM SPSSModeler . 33 Аналитические платформы для Big Data: KNIME. 34 Аналитические платформы для Big Data: Qlik Analytics Platform . 35 Аналитические платформы для Big Data: STATISTICA 36 Data Miner. 37 Аналитические платформы для Big Data: Informatica 38 Intelligent Data Platform. 39 Аналитические платформы для Big Data: WorldProgramming

	System. 33 Аналитические платформы для Big Data: Deductor.34 Аналитические платформы для Big Data: SAS Enterprise Miner. 35 Zookeeper. 36 Flume. 37 IBM Watson Analytics. 38 Dell EMC Analytic Insights Module.39 Windows Azure HDInsight. 40 Microsoft Azure Machine Learning.41 Pentaho Data Integration. 42 Teradata Aster Analytics. 43 SAP BusinessObjects Predictive Analytics.44 Oracle Big Data Preparation. 45 Аналитика Big Data — реалии и перспективы в России и мире. 46 Data Mining. 47 Краудсорсинг. 48 Смешение и интеграция данных.
Визуализация и аналитическая отчетность	49 Машинное обучение. 50 Искусственные нейронные сети.51 Распознавание образов. 52 Прогнозная аналитика. 53 Имитационное моделирование. 54 Пространственный анализ. 55 Статистический анализ. 56 Визуализация аналитических данных. 57 Big data: применение и возможности. 58 Решения на основе Big data. 59 Рынок Big data в России. 60 Big data в банках. 61 Big data в бизнесе. 62 Big data в маркетинге. 63. Loginom

6.6. Шкалы и критерии оценивания по формам текущего контроля и промежуточной аттестации

Шкала и критерии оценивания

Оценка	Критерии оценивания для мероприятий контроля с применением 2-х балльной системы
«зачтено»	ПК-5, ПК-6
«не зачтено»	Результаты обучения не сформированы на пороговом уровне