



УТВЕРЖДАЮ

Ректор

Искаков И.Ж.

01 июля 2020 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Анализ данных»

(наименование дисциплины)

Направление подготовки/Специальность 38.03.05 Бизнес-информатика

Квалификация выпускника Бакалавр

Направленность (профиль) Бизнес-информатика, технология блокчейн – криптовалюта

Форма обучения Очная, заочная

1. Место дисциплины в структуре образовательной программы, входные требования для освоения дисциплины (при необходимости)

Дисциплина «Анализ данных» относится к дисциплинам по выбору вариативной части Блока 1 «Дисциплины (модули)» программы бакалавриата по направлению подготовки 38.03.05 Бизнес-информатика. Для изучения дисциплины необходимы знания, умения и навыки, полученные обучающимися в результате изучения дисциплины «Теоретические основы информатики. Имитационное моделирование». Содержание дисциплины служит основой для освоения дисциплины «Анализ информационно-технологических процессов на предприятии» и прохождения производственной (преддипломной) практики.

2. Объем дисциплины в зачетных единицах

Объем дисциплины составляет 8 зачетных единиц

Объем дисциплины	Всего часов	
	Очная форма обучения	Заочная форма обучения
Общая трудоемкость дисциплины	288	288
Контактная работа (по видам учебных занятий) (всего)	128,5	48,5
Из них:		
Лекции (Лек)	32	16
Практические занятия (Пр)	56	20
Семинарские занятия (Сем)	-	-
Индивидуальные занятия (ИЗ)	16	2
Контроль самостоятельной работы (КСР)	8	8
Контактные часы на аттестацию (КА)	0,5	0,5
Самостоятельная работа под руководством преподавателя (СРП)	16	2
Самостоятельная работа студентов (СР)	124	204
Вид промежуточной аттестации	экзамен	экзамен
Объем часов, отводимых на подготовку к промежуточной аттестации (Контроль)	35,5	35,5

3. Содержание дисциплины, структурированное по темам (разделам)

Раздел 1. Основы анализа данных

Тема 1.1 Введение в анализ данных его роль в современном бизнесе

Основные понятия и определения. Специфика работы с информацией. Свойства информации. Информация и данные. Виды и методы анализа данных. Этапы решения задачи анализа данных и их взаимосвязи. Роль анализа данных в сфере бизнеса. Системы поддержки принятия решения (СППР).

Тема 1.2 Корреляционный анализ

Корреляция, коэффициент корреляции, индекс корреляции, оценка значимости коэффициента корреляции, индексов корреляции, доверительные интервалы, проверка гипотез, типы ошибок, области принятия и непринятия.

Тема 1.3 Регрессионный анализ.

Модель парной регрессии, модель множественной регрессии

Тема 1.4 Анализ данных во временных рядах.

Аддитивная и мультипликативная модели временного ряда. Тренд, сезонность, календарные эффекты. Адаптивные модели: экспоненциальное сглаживание, модели Хольта-Уинтерса и Тейла-Вейджа. Примеры прикладных задач: прогнозирование трафика, числа посещений, объемов продаж, модели оптимизации в анализе данных временных рядов.

Раздел 2. Интеллектуальный анализ данных

Тема 2.1 Анализ многомерных данных: OLAP технологии.

Многомерная модель данных. Понятия OLAP-куба, операции, выполняемые над гиперкубом. Архитектура OLAP систем. Виды OLAP систем: ROLAP (Relational OLAP), MOLAP (Multidimensional OLAP), HОLAP (Hybrid OLAP), их преимущества и недостатки. Сферы применения OLAP технологии.

Тема 2.2 DM технологии.

Понятие интеллектуального анализа данных (DM). Требования, предъявляемые к новым знаниям. Задачи Data Mining. Стадии Data Mining. Модели Data Mining: классификация, кластеризация, поиск ассоциативных правил, итоговые модели, регрессионные модели. Сферы применения технологии интеллектуального анализа данных. Методика обнаружения нового знания в хранилищах данных (KDD). Характеристика этапов KDD. Постановка задачи классификации и представление результатов. Классификационные правила. Деревья решений; алгоритмы построения деревьев решений.

Тема 2.3 Анализ текстовых данных Text Mining

Задачи текстового анализа. Этапы анализа текстов. Извлечение ключевых понятий. Современные инструменты для анализа текстов. Алгоритмы анализа текста.

4. Методические рекомендации по организации изучения учебной дисциплины

4.1 Рекомендуемые образовательные технологии

В преподавании дисциплины «Рынки информационно-коммуникационных технологий и организация продаж» используются разнообразные образовательные технологии как традиционного, так и инновационного характера, учитывающие практический и теоретический характер дисциплины:

- лекции;
- практические занятия;
- дискуссии;
- выступления с докладами и сообщениями;
- тестирование.

4.2 Методические рекомендации для обучающихся по освоению дисциплины

Теоретические занятия

Лекция 1. Тема 1.1 Введение в анализ данных его роль в современном бизнесе.

Основные понятия и определения. Специфика работы с информацией. Свойства информации. Информация и данные. Роль анализа данных в сфере бизнеса.

Лекция 2. Тема 1.2 Корреляционный анализ.

Дать определение корреляции, коэффициенту корреляции, индексу корреляции, оценки значимости коэффициента корреляции, доверительным интервалам, проверки гипотез, типом ошибок, областям принятия и непринятия решений.

Лекция 3. Тема 1.3 Регрессионный анализ.

Более подробно рассказать о модели парной регрессии и модели множественной регрессии.

Лекция 4. Тема 1.4 Анализ данных во временных рядах.

Аддитивная и мультипликативная модели временного ряда. Тренд, сезонность, календарные эффекты. Адаптивные модели.

Лекция 5. Тема 2.1 Анализ многомерных данных: OLAP технологии.

Рассказать о многомерной модели данных. Дать определение OLAP-кубу, операции, выполняемые над гиперкубом. Познакомиться с архитектурой OLAP систем.

Лекция 6. Тема 2.2 DM технологии.

Описать интеллектуальный анализ данных (DM). Требования, предъявляемые к новым знаниям. Задачи Data Mining. Стадии Data Mining. Модели Data Mining: классификация, кластеризация, поиск ассоциативных правил, итоговые модели, регрессионные модели.

Лекция 7. Тема 2.3 Анализ текстовых данных Text Mining.

Более детально разобрать задачи текстового анализа. Этапы анализа текстов. Извлечение ключевых понятий.

Практические занятия

Тема 1.1 Введение в анализ данных его роль в современном бизнесе.

Решение задач на анализ данных и его взаимосвязи.

Тема 1.2 Корреляционный анализ.

Решение задач на нахождение коэффициента корреляции, индекса корреляции, оценки значимости коэффициента корреляции, индексов корреляции, доверительных интервалов, проверки гипотез, типа ошибки.

Тема 1.3 Регрессионный анализ.

Уметь применять формулы парной регрессии и множественной регрессии для решения задач.

Тема 1.4 Анализ данных во временных рядах.

Уметь применять модель Хольта-Уинтерса и Тейла-Вейджа. Решить задачи: прогнозирование трафика, числа посетителей, объемов продаж, модели оптимизации в анализе данных временных рядов.

Тема 2.1 Анализ многомерных данных: OLAP технологии.

Построить одну из моделей по ROLAP (Relational OLAP), MOLAP (Multidimensional OLAP), HOLAP (Hybrid OLAP). Применение OLAP технологии.

Тема 2.2 DM технологии.

Научиться строить алгоритмы построения деревьев решений.

Тема 2.3 Анализ текстовых данных Text Mining.

Научиться строить алгоритмы анализа текста.

4.3 Методические рекомендации для обеспечения самостоятельной работы обучающихся по дисциплине

Самостоятельная работа студентов включает усвоение теоретического материала, подготовку к практическим занятиям, выполнение самостоятельных заданий, изучение

литературных источников, использование Internet-данных, подготовку к текущему контролю знаний, к промежуточной аттестации.

Контроль за выполнением самостоятельной работы ведется в процессе изучения курса преподавателем на практических занятиях, а так же при проверке индивидуальных заданий и письменных работ.

Управление самостоятельной работой студента

Формы управления самостоятельной работой:

- консультирование;
- проверка части выполненной работы;
- предложение списка рекомендованной литературы;

План самостоятельной работы:

- повторение материала, подготовка к практическим (семинарским) занятиям.

Вопросы для самоконтроля

1. Дайте определение понятию «данные».
2. Назовите типы шкал с примерами?
3. Чем дискретные данные отличаются от непрерывных?
4. Что такое анализ данных?
5. Назовите этапы анализа данных.
6. Чем качественный этап анализа данных отличается от количественного?
7. Опишите основе задачи при качественном анализе данных.
8. Что такое доверительный интервал?
9. Из чего состоит алгоритм проверки гипотез?
10. Что такое нулевая гипотеза?
11. Перечислите типы ошибок?
12. Что такое сводная таблица?
13. Назовите причины использования регрессионного анализа.
14. На основе какого метода вычисляются коэффициенты уравнения регрессии?
15. Что такое величина достоверности аппроксимации?
16. Из чего состоит проверка достоверности модели регрессии?
17. В чем специфика построения моделей по временным рядам данных
18. Дайте сравнительный анализ OLAP и OLTP систем. Сферы их применения.
19. В чем отличие информационного хранилища от баз данных?
20. Принципы построения информационных хранилищ. Классификация информационных хранилищ.
21. Модели информационных хранилищ. Многомерная модель данных. Нормальная форма.
22. Правила Кодда. Зачем применяется денормализация моделей?
23. Дайте характеристику стандартам Data Mining.
24. Дайте характеристику этапа ETL (Extracting Transforming and Loading).
25. Какие задачи решаются Data Mining?
26. Каково предназначение и средства разведочный анализ данных? Дайте характеристику диаграммы «ящик с усами»
27. Назовите какие операции выполняются при агрегировании данных.
28. Приведите примеры использования статистических пакетов для разведочного анализа.

29. Назовите и выполните сравнительный анализ графических средств анализа. Дайте характеристику биржевых диаграмм.

30. Дайте определение задачи классификации. Какие методы решения задачи классификации Вы знаете?

31. Особенности решения задач классификации с обучением.

32. Деревья классификации и их свойства.

5. Оценочные средства для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине

5.1 Перечень оценочных средств

Оценочные средства представляют собой задания, обязательные для выполнения студентом, позволяющие ему приобрести теоретические знания, практически умения (навыки) и опыт, а также решать задачи, связанные с будущей профессиональной деятельностью. Включают в себя задания для текущего контроля уровня успеваемости, оценивающие ход освоения учащимися дисциплины, и задания для промежуточной аттестации обучающихся, обеспечивающие оценивание промежуточных и окончательных результатов обучения по дисциплине.

Оценочные средства для проведения текущего контроля успеваемости

Темы рефератов

1. Информационные ресурсы общества и предприятий
2. Информационные технологии как средство доступа к информационным ресурсам
3. Информационные системы и их классификация
4. Системы обработки данных
5. Принципы проектирования, внедрения и эксплуатации ЭИС
6. Состав и структура экономических информационных систем
7. Понятие об информационном поиске
8. Понятие об информационно-поисковой системе.
9. Информационно-поисковые системы научно-технической информации
10. Системы научно-технической информации
11. Основы Интернет. Интернет-технологии, используемые в сфере бизнеса
12. Поисковые системы Интернет
13. Использование Интернет в корпоративных информационных системах
14. Виды использования сети Интернет в экономической деятельности
15. Организация электронной коммерции
16. Технология и методы проектирования ЭИС
17. Стадии и этапы процесса разработки ЭИС
18. Реинжиниринг бизнес-процессов
19. Современные средства проектирования ИС
20. Методы защиты информации.

Темы докладов-презентаций

1. Инструментарий OLAP-технологий в бизнесе.
2. Анализ данных, знаний и структур в системах искусственного интеллекта .

3. Искусственный интеллект в современных бизнес-проектах.
4. Визуализация как средство анализа информации в бизнес-проектах .
5. Применение ассоциативных правил в бизнес-проектах.
6. Перспективные средства анализа бизнес-информации.
7. Методы анализа структурных объектов.
8. IT-решения в визуализации данных. Обзор бизнес-проектов .
9. Анализ текстовой информации в бизнес-проектах.
10. Сравнительная характеристика методов анализа данных
11. Распределенная обработка данных в бизнес-проектах.
12. Классификационный анализ с обучением. Деревья решений
13. Классификационный анализ без обучения. Кластерный анализ
14. Системы мобильных агентов. Использование мобильных агентов для анализа данных.
15. Аннотирование текстов. Современные инструменты для анализа текстовой информации в бизнес-проектах.
16. Задачи интеллектуального анализа данных в системах поддержки принятия решений

Практические задания

Задание 1 «Корреляционный анализ».

На основании данных для соответствующего варианта (табл. 6,8) :

1. Вычислить линейный коэффициент парной корреляции r_{xy} и индекс корреляции R .
2. Проверить значимость коэффициента парной корреляции r_{xy} и индекса корреляции R при заданном уровне значимости α .
3. Построить доверительный интервал для значимого линейного коэффициента парной корреляции r_{xy} .

Таблица 6

Вариант	Графы из таблицы (x,y)	Уровень значимости	Зависимость
1	1,2	0,05	$-65+45 \cdot \ln x$
2	1,3	0,025	$3,5 \cdot x^{0,9}$
3	1,4	0,01	$5+4 \cdot \ln x$
4	1,5	0,05	$e^{0,55+0,08 \cdot x}$
5	1,6	0,025	$5,9 \cdot x^{0,6}$
6	1,7	0,01	$15+85/x$
7	1,8	0,05	$22 \cdot x^{0,4}$
8	2,3	0,025	$-200+70 \cdot \ln x$
9	2,4	0,01	$7 \cdot x^{0,24}$
10	2,5	0,05	$0,008 \cdot x^2$
11	2,6	0,025	$13 \cdot e^{0,013x}$
12	2,7	0,01	$40-5 \cdot \ln x$
13	2,8	0,05	$75-35 \cdot \ln x$
14	3,4	0,025	$1,5+4 \cdot \ln x$
15	3,5	0,01	$0,055 \cdot x^{1,6}$
16	3,6	0,05	$1,5 \cdot x^{0,8}$
17	3,7	0,025	$35-4 \cdot \ln x$
18	3,8	0,01	$40-30 \cdot \ln x$

19	4,5	0,05	$12 \cdot e^{0,07x}$
20	4,6	0,025	$19 \cdot e^{0,04x}$
21	4,7	0,01	$35 \cdot x - 0,2$
22	4,8	0,05	$25 \cdot x^{0,4}$
23	5,6	0,025	$9 \cdot x^{0,4}$
24	5,7	0,01	$25 - 1,7 \cdot \ln x$
25	5,8	0,05	$14 + 18 \cdot \ln x$

В таблице 6 представлены исходные данные для практической работы по корреляционному анализу.

Задание 2 «Регрессионный анализ»

На основании данных таблиц 7, 9, 10 :

1. Построить уравнения регрессии: линейную и нелинейную регрессию в соответствии с вариантом.
2. Вычислить показатели качества и точности для каждого уравнения.
3. Проверить значимость уравнений регрессии при уровнях значимости 0,05 и 0,01.
4. Определить лучшее уравнение регрессии на основе средней ошибки аппроксимации.
5. Проверить значимость коэффициентов линейной регрессии и построить доверительные интервалы для точных значений параметров \tilde{a} и \tilde{b} уравнения линейной регрессии с уровнем значимости 0,05.
6. Построить точечный и интервальный прогноз для значения $x = x_{\max}$ по уравнению линейной регрессии с уровнем значимости 0,05.
7. Определить средний коэффициент эластичности по уравнению линейной регрессии.
8. Графически представить результаты моделирования.

Таблица 7

№ варианта	Графы из табл. (x, y)	Виды кривых выравнивания
1	1,14	Степенная
2	4,14	Показательная
3	6,14	Логарифмическая
4	11,14	Степенная
5	2,15	Показательная
6	3,15	Логарифмическая
7	8,15	Степенная
8	1,17	Показательная
9	2,17	Логарифмическая
10	6,17	Степенная
11	11,17	Показательная
12	12,17	Логарифмическая
13	2,19	Степенная
14	6,19	Показательная
15	9,19	Логарифмическая

Таблица 8

Парная корреляция	1	2	3	4	5	6	7	8
1	3	14	11	12	6	9	32	45
2	6	24	18	14	5	16	29	55
3	6	35	18	11	11	12	22	64

4	9	53	23	17	11	16	25	63
5	12	67	32	13	18	29	19	67
6	13	61	33	11	14	31	18	47
7	13	55	34	15	6	28	19	50
8	14	47	38	20	7	34	22	62
9	15	38	35	19	21	33	19	63
10	19	66	45	19	36	30	20	54
11	24	79	62	25	24	32	20	70
12	26	79	66	26	62	33	18	63
13	26	78	61	19	33	54	18	86
14	26	95	66	21	17	42	18	78
15	27	99	65	12	45	45	18	78
16	28	79	73	15	51	52	19	80
17	30	80	74	26	28	49	18	80
18	31	91	69	24	67	55	18	74
19	31	88	70	16	39	52	17	95
20	33	77	86	17	53	42	19	80
21	34	74	79	21	48	40	19	79
22	35	107	82	18	57	64	19	80
23	66	172	162	16	327	112	17	117
24	36	80	79	19	66	43	18	100
25	36	110	91	20	51	43	17	81
26	37	88	97	25	40	48	19	76
27	67	166	145	14	388	90	18	139
28	38	109	98	26	100	58	18	95
29	39	93	98	29	84	67	18	106
30	40	120	85	18	86	65	17	80
31	40	92	101	29	54	77	17	82
32	40	110	95	22	120	74	17	102
33	41	93	101	24	61	50	17	83
34	44	127	104	26	79	50	18	102
35	47	109	117	18	149	48	18	99
36	48	126	102	17	101	43	18	97
37	48	133	116	16	129	85	18	93
38	49	132	123	18	157	70	18	103
39	52	136	133	25	148	46	17	99
40	54	114	131	24	125	50	18	101

Таблица 9

1	2	3	4	5	6	7	8	9	10	11	12	13
0.16	0.11	2.40	0.16	14.99	0.80	0.57	12.01	0.81	74.96	3.35	2.39	50.46
0.80	0.19	5.44	0.24	14.72	4.01	0.96	27.20	1.20	73.61	16.83	4.02	114.26
0.94	0.20	5.87	0.32	4.55	4.72	1.00	29.33	1.58	22.77	19.80	4.19	123.17
0.26	0.15	9.65	0.48	11.57	1.29	0.75	48.27	2.39	57.86	5.44	3.13	202.72
0.27	0.05	8.11	0.13	1.64	1.36	0.23	40.53	0.66	8.18	5.73	0.99	170.22
0.47	0.05	8.23	0.10	16.75	2.35	0.25	41.17	0.50	83.73	9.87	1.04	172.90
0.34	0.16	3.89	0.22	17.56	1.71	0.80	19.47	1.12	87.81	7.17	3.35	81.79
0.31	0.06	7.97	0.13	18.92	1.55	0.28	39.87	0.66	94.60	6.53	1.18	167.45
0.65	0.12	0.97	0.11	6.11	3.27	0.61	4.83	0.53	30.57	13.74	2.55	20.29
0.06	0.08	9.05	0.30	14.92	0.31	0.40	45.24	1.48	74.62	1.28	1.66	190.01
0.14	0.05	9.20	0.18	19.72	0.70	0.25	46.02	0.90	98.58	2.95	1.05	193.29
0.10	0.09	5.00	0.14	19.90	0.52	0.46	25.01	0.68	99.52	2.20	1.94	105.02
0.81	0.15	5.69	0.24	9.36	4.04	0.75	28.47	1.18	46.81	16.95	3.14	119.58
0.19	0.11	7.32	0.36	3.61	0.93	0.53	36.59	1.81	18.07	3.89	2.22	153.66
0.39	0.05	2.99	0.17	1.02	1.97	0.26	14.93	0.85	5.08	8.26	1.10	62.71
0.91	0.11	3.05	0.19	3.11	4.55	0.55	15.27	0.94	15.53	19.12	2.32	64.15
0.64	0.02	4.17	0.48	1.89	3.20	0.11	20.87	2.39	9.45	13.44	0.46	87.66
0.87	0.01	9.54	0.03	9.53	4.33	0.06	47.72	0.16	47.64	18.17	0.24	200.41
0.33	0.14	6.88	0.15	0.01	1.64	0.69	34.41	0.76	0.03	6.90	2.92	144.51
0.92	0.16	8.51	0.04	11.08	4.62	0.78	42.57	0.20	55.39	19.39	3.26	178.80
0.49	0.11	0.94	0.09	2.98	2.44	0.53	4.69	0.45	14.92	10.24	2.24	19.70
0.17	0.01	7.51	0.10	1.88	0.83	0.03	37.56	0.52	9.39	3.48	0.14	157.75
0.47	0.07	0.81	0.17	7.65	2.33	0.35	4.06	0.84	38.25	9.80	1.45	17.06
0.79	0.15	5.16	0.44	0.02	3.97	0.75	25.80	2.18	0.08	16.67	3.17	108.35
0.22	0.08	6.21	0.33	4.25	1.08	0.38	31.06	1.66	21.25	4.53	1.60	130.45
0.39	0.04	9.38	0.47	0.60	1.95	0.21	46.88	2.34	3.01	8.20	0.86	196.90
0.57	0.04	4.28	0.10	2.30	2.84	0.22	21.38	0.52	11.51	11.94	0.93	89.80
0.24	0.05	3.42	0.30	10.11	1.20	0.24	17.10	1.52	50.53	5.03	1.00	71.83
0.08	0.20	3.90	0.06	0.10	0.39	1.00	19.52	0.28	0.48	1.64	4.19	81.98
0.53	0.08	4.38	0.11	17.98	2.66	0.40	21.90	0.56	89.90	11.18	1.69	91.97
0.24	0.11	5.30	0.28	1.34	1.19	0.55	26.51	1.41	6.70	5.01	2.29	111.35
0.12	0.13	1.63	0.39	6.40	0.58	0.65	8.15	1.95	32.01	2.45	2.74	34.25
0.76	0.09	5.71	0.47	1.86	3.78	0.44	28.53	2.37	9.32	15.87	1.85	119.84
0.61	0.01	7.65	0.45	3.49	3.07	0.03	38.25	2.24	17.47	12.90	0.12	160.64
0.85	0.13	0.82	0.41	15.02	4.23	0.63	4.10	2.05	75.10	17.78	2.63	17.24
0.39	0.20	4.50	0.38	10.15	1.95	0.99	22.50	1.89	50.76	8.17	4.18	94.51
0.23	0.17	1.17	0.09	14.31	1.17	0.85	5.87	0.47	71.55	4.90	3.58	24.66
0.77	0.10	5.71	0.28	6.39	3.85	0.51	28.53	1.42	31.96	16.15	2.15	119.82
0.29	0.15	8.93	0.48	11.19	1.46	0.76	44.64	2.38	55.95	6.13	3.20	187.48
0.99	0.15	1.63	0.12	0.30	4.97	0.76	8.17	0.59	1.51	20.89	3.19	34.32
0.83	0.17	8.58	0.08	17.06	4.17	0.86	42.91	0.42	85.30	17.50	3.61	180.22
0.44	0.13	4.19	0.46	1.50	2.19	0.65	20.96	2.28	7.50	9.20	2.73	88.02
0.03	0.16	6.48	0.34	12.22	0.14	0.81	32.40	1.69	61.08	0.60	3.38	136.08
0.07	0.11	2.37	0.34	5.36	0.33	0.57	11.85	1.71	26.79	1.41	2.40	49.77
0.75	0.16	2.80	0.10	3.24	3.74	0.80	14.00	0.50	16.22	15.69	3.36	58.78

Таблица 10

14	15	16	17	18	19
21.1	15.0	-0.6	20.7	23.0	12.7
20.6	10.3	43.0	20.5	24.4	45.7
20.7	4.5	49.2	20.6	25.7	51.0
22.1	-0.5	51.7	21.0	28.5	101.7
20.2	-4.5	27.6	19.9	20.3	49.5
19.8	2.9	41.7	19.6	19.4	62.4
21.3	14.2	26.2	20.8	24.8	34.8
20.2	4.6	23.7	20.0	20.4	41.4
20.0	13.6	24.2	19.9	20.6	5.9
21.4	0.9	36.8	20.6	23.4	80.1
20.7	2.4	42.4	20.2	21.1	83.9
20.9	11.8	32.0	20.6	22.3	51.6
20.3	6.3	36.4	20.2	22.7	38.9
21.6	-0.7	17.4	20.7	24.8	42.5
20.1	5.6	33.0	19.8	20.2	30.8
19.6	7.7	45.1	19.6	20.2	29.7
20.3	3.0	31.6	19.8	18.8	38.6
18.5	-4.1	70.8	18.4	16.8	94.3
20.9	-1.0	26.3	20.6	23.0	44.8
19.5	1.6	71.4	19.5	19.4	88.9
20.1	11.8	23.8	20.0	20.7	6.2
20.0	-3.9	34.6	19.8	19.5	64.1
20.1	13.6	4.7	19.9	20.4	3.6
20.9	2.7	25.3	20.7	25.8	30.4
21.2	1.2	31.1	20.4	22.9	57.3
21.0	-7.6	46.4	20.1	21.0	83.4
19.5	3.5	19.9	19.4	18.9	23.3
20.8	9.2	27.5	20.1	21.2	36.4
21.6	6.2	1.0	21.5	24.0	18.0
19.9	11.8	19.6	19.8	20.1	24.5
21.2	2.2	30.2	20.6	23.9	50.3
22.0	12.5	1.4	20.9	26.7	14.7
20.6	1.3	40.0	20.3	22.6	49.6
20.2	-3.4	54.4	19.6	17.9	81.0
20.5	18.4	29.6	20.4	23.8	10.7
21.9	10.1	22.8	21.3	29.1	36.9
21.2	18.2	-2.8	21.0	23.4	4.0
20.1	3.8	50.4	19.9	21.5	57.1
22.1	0.8	49.5	21.0	28.5	93.8
19.6	9.9	37.4	19.6	20.2	12.4
20.0	4.8	44.1	19.9	20.9	47.8
21.5	5.0	28.8	20.8	26.2	40.9
22.2	6.4	18.1	21.3	27.9	48.8
21.8	10.2	20.0	20.8	25.5	27.2
20.1	9.2	14.5	20.0	21.2	11.5

Оценочные средства для проведения промежуточной аттестации по дисциплине

Список экзаменационных вопросов

1. Понятие анализа данных. Методология анализа данных.
2. Подходы к моделированию.
3. Принципы анализа данных.
4. Типы данных. Методы анализа данных.
5. Этапы статистического исследования.
6. Задачи и методы статистического анализа данных. Дисперсионный анализ данных.
7. Корреляционный анализ данных. Ковариационный анализ данных.
8. Описательная статистика. Экспоненциальное сглаживание.
9. Регрессия. Выборка.
10. Понятие СППР. Задачи.
11. Базы данных и СППР.
12. Базы данных и базы знаний
13. Понятие хранилища данных. Свойства хранилища данных.
14. Требования, предъявляемые к системам OLTP и СППР.
15. Концепция хранилища данных.
16. Организация хранилища данных.
17. Очистка данных.
18. Многомерная модель данных.
19. Определение OLAP-систем. Основные правила Кодда.
20. Архитектура OLAP-систем.
21. MOLAP.
22. ROLAP.
23. HOLAP.
24. Добыча данных.
25. Классификация задач Data Mining.
26. Задача классификации и регрессии в интеллектуальном анализе данных.
27. Задача поиска ассоциативных правил в интеллектуальном анализе данных.
28. Задача кластеризации в интеллектуальном анализе данных.
29. Задачи текстового анализа.
30. Этапы анализа текстов.
31. Современные инструменты для анализа текстов.
32. Алгоритмы анализа текста .
33. Практическое применение Data Mining.
34. Модели Data Mining.
35. Методы Data Mining.
36. Нечеткая логика. Генетические алгоритмы.
37. Нейронные сети. Этапы процесса обнаружения знаний.
38. Подготовка к анализу. Средства Data Mining.
39. Правила классификации. Деревья решений.
40. Методы построения правил классификации.
41. Методы построения деревьев решений.
42. Прогнозирование временных рядов.
43. Поиск ассоциативных правил.
44. Кластеризация.

Тест

1. *Полезность и ценность информации проявляется в том, что данные только тогда ценны для руководителя предприятия, когда...*

1. данные верны; с. руководитель удовлетворён быстродействием ИСУП;
2. данные сохранены на сервере
3. информация используется для принятия управленческих решений;

2. *Система класса ERP это*

1. Фактографическая система
2. Информационно-поисковая система
3. Документальная система с полнотекстовым доступом
4. Информационно-аналитическая система
5. Информационная система управления ресурсами

6. *Информационно-логическая система*

7. *Партнерская система*

3. *Консультант + относится к классу*

1. Фактографических систем
2. Информационно-поисковых систем
3. Документальных систем с полнотекстовым доступом
4. Информационно-аналитических систем управления ресурсами
5. Информационно-логических систем
6. Партнерских систем

4. *Понятие _____ было введено аналитиками Gartner Group в конце 80-х годов, и означало*

- ✓ «бизнес-разведка»,
- ✓ «бизнес-аналитика»,
- ✓ «бизнес-интеллект»

5. *Инструменты для анализа исторических данных, помогающие аналитикам синтезировать из данных значимую информацию, используя интуицию и выявляя скрытые закономерности*

1. OLTP
2. CRM
3. Data Mining
4. Olap
5. Технология гипертекста
6. Построения воронки продаж
7. Полнотекстовый доступ к документам
6. Средства построения хранилищ и витрин данных

7. *Средства интеллектуальной добычи данных*

8. *Автор концепции данной технологии дал ей следующее определение « предметно-ориентированные, интегрированные, неизменяемые, поддерживающие хронологию наборы данных, организованные для целей поддержки управления», обеспечивающие менеджеров и аналитиков достоверной информацией, необходимой для принятия решения»*

9. Этот класс информационных систем выполняет обработку с использованием данных, полученных из разнородных источников:

- логический вывод,
- поиск фактов и ситуаций,
- выявление похожих объектов и ситуаций,
- контекстный анализ,
- поиск неявных связей

Это информационные системы _____

10. Этот класс информационных систем предназначен для агрегирования и отображения информации, полученной из различных источников реализуют функции поиска информации, формирования специализированных отчетов

Это информационные системы _____

11. Хранилище данных ориентированно/ не ориентированно на решение какой-либо определенной функциональной аналитической задачи

12. Небольшое функционально-ориентированное хранилище, оптимизированного для запросов к данным конкретной предметной области с целью для передачи этих данных для последующей аналитической обработки

13. Коэффициент парной корреляции характеризует тесноту _____ связи между _____ переменными.

- | | | | |
|---|--------------------------|-------------|----------------------|
| 1 | линейной ... несколькими | 2 | нелинейной ... |
| | | несколькими | |
| 3 | линейной ... двумя | 4 | нелинейной ... двумя |

14. Установите соответствие между наименованиями элементов уравнения $Y=b_0+b_1X+e$ и их буквенными обозначениями:

- | | | | |
|------------------------|---------------------------|---------------------------|-------------------------|
| 1. параметры регрессии | 2. объясняющая переменная | 3. объясняемая переменная | 4. случайные отклонения |
| 1. А) | 2. Y | 3. В) | 4. b_0, b_1 |
| 5. С) | 6. X | 7. D) | 8. e |

15. Для линейного уравнения регрессии $y = a + bx + \varepsilon$ метод наименьших квадратов используется при оценивании параметров...

- | | | | |
|----|---|----|---|
| 1) | b | 2) | y |
| 3) | x | 4) | a |

16. Сколько параметров содержит парное линейное уравнение регрессии?

- 1 1
- 2 2
- 3 3
- 4 4

17. При выполнении предпосылок МНК оценки параметров регрессии обладают свойствами:

(-выберите несколько вариантов ответа)

- | | | | |
|----|---------------|----|-------------------|
| 1) | достоверность | 2) | эффективность |
| 3) | несмещенность | 4) | несостоятельность |

18. Как влияет увеличение объема выборки на величину остаточной дисперсии случайной величины?

(- выберите один вариант ответа)

- 1 Никак.
- 2 Остаточная дисперсия увеличивается.
- 3 Остаточная дисперсия уменьшается.
- 4 Результат зависит от конкретного вида случайной величины.

19. При каком значении параметра x оценка случайной величины y , полученная в рамках парной линейной регрессионной модели, будет наиболее точной?

(- выберите один вариант ответа)

1. При $x = (x_{\min} + x_{\max})/2$, где x_{\min} , x_{\max} - минимальное и максимальное значения параметра x из обследованного интервала.

2. При $x = \sqrt{x_{\min} x_{\max}}$

3. При $x = \bar{x}$, где \bar{x} - среднее значение параметра x из обследованного интервала.

4. Точность одинакова при всех x .

20. Рассматривается парная линейная регрессионная модель. Как изменится ширина доверительного интервала для условного математического ожидания случайной величины $\bar{y}(x)$ при увеличении объема выборки в 4 раза?

(- выберите один вариант ответа)

1. Увеличится в 4 раза.
2. Уменьшится в 4 раза.
3. Увеличится в 2 раза.
4. Уменьшится в 2 раза.

Ключ:

№ вопроса	1	2	3	4	5	6	7	8	9	10
Ответ	3	5	2, 3	Business Intelligence BI	3, 4	data warehouse	OLAP	Data Warehouse	анализа структурированной информации	визуализации структурированной информации

№ вопроса	11	12	13	14	15	16	17	18	19	20
Ответ	не ориентированно	OLAP	1	A-3, B-1, C-2, D-4	1,4	2	2.3	3	1	4

5.2 Планируемые результаты обучения, соотнесенные с индикаторами достижения компетенций

В процессе изучения дисциплины у обучающихся должны быть сформированы следующие компетенции:

- способность создавать новые бизнес-проекты на основе инноваций в сфере ИКТ (ПК-28).

Код и формулировка компетенции	Индикаторы достижения компетенций
ПК-28 способность создавать новые бизнес-проекты на основе инноваций в сфере ИКТ	Знает о мировых информационных ресурсах и основных участниках рынка мировых информационных ресурсов; закономерностей развития рынка мировых информационных ресурсов; практических основ работы с информационными ресурсами, понятий и закономерностей использования информации в бизнесе; особенности функционирования рынка информационных ресурсов: особенности спроса, предложения, формирования рыночного равновесия и продвижения продукции
	Умеет анализировать тенденции развития отечественного и зарубежного рынков ИКТ; организовывать взаимодействие с клиентами и партнерами на рынке ИКТ; разрабатывать маркетинговую стратегию продвижения ИКТ
	Владет навыками оценки заполненности ниш рынка ИКТ и анализа тенденций его развития; использования информации, позволяющей прогнозировать потребительский спрос для формирования предложения ИКТ; разработки мероприятий по продвижению продукции на рынке ИКТ

5.3. Система оценивания результатов и критерии выставления оценок в ходе промежуточной аттестации

Для оценивания результатов промежуточной аттестации применяется система оценивания, включающая следующие оценки: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Экзамен. Критерии оценивания

На экзамен выносятся вопросы, охватывающие все содержание учебной дисциплины.

Знания обучающихся оцениваются путем выставления по результатам ответа обучающегося итоговой оценки «отлично», либо «хорошо», либо «удовлетворительно», либо «неудовлетворительно».

Оценка «отлично» при приеме экзамена выставляется в случае:

- полного, правильного и уверенного изложения обучающимся учебного материала по каждому из вопросов билета;
- уверенного владения обучающимся понятийно-категориальным аппаратом учебной дисциплины;
- логически последовательного, взаимосвязанного и правильно структурированного изложения обучающимся учебного материала, умения устанавливать и прослеживать причинно-следственные связи между событиями, процессами и явлениями, о которых идет речь в вопросах билета;
- приведения обучающимся надлежащей аргументации, наличия у обучающегося логически и нормативно обоснованной точки зрения при освещении проблемных, дискуссионных аспектов учебного материала по вопросам билета;
- лаконичного и правильного ответа обучающегося на дополнительные вопросы преподавателя.

Оценка «хорошо» при приеме экзамена выставляется в случае:

- недостаточной полноты изложения обучающимся учебного материала по отдельным (одному или двум) вопросам билета при условии полного, правильного и уверенного изложения учебного материала по, как минимум, одному вопросу билета;
- допущения обучающимся незначительных ошибок и неточностей при изложении учебного материала по отдельным (одному или двум) вопросам билета;
- допущения обучающимся незначительных ошибок и неточностей при использовании в ходе ответа отдельных понятий и категорий дисциплины;
- нарушения обучающимся логической последовательности, взаимосвязи и структуры изложения учебного материала по отдельным вопросам билета, недостаточного умения обучающегося устанавливать и прослеживать причинно-следственные связи между событиями, процессами и явлениями, о которых идет речь в вопросах билета;
- приведения обучающимся слабой аргументации, наличия у обучающегося недостаточно логически и нормативно обоснованной точки зрения при освещении проблемных, дискуссионных аспектов учебного материала по вопросам билета;
- допущения обучающимся незначительных ошибок и неточностей при ответе на дополнительные вопросы преподавателя.

Любой из указанных недостатков или их определенная совокупность могут служить основанием для выставления обучающемуся оценки «хорошо».

Оценка «удовлетворительно» при приеме экзамена выставляется в случае:

- невозможности изложения обучающимся учебного материала по любому из вопросов билета при условии полного, правильного и уверенного изложения учебного материала по как минимум одному из вопросов билета;
- допущения обучающимся существенных ошибок при изложении учебного материала по отдельным (одному или двум) вопросам билета;
- допущении обучающимся ошибок при использовании в ходе ответа основных понятий и категорий учебной дисциплины;
- существенного нарушения обучающимся или отсутствия у обучающегося логической последовательности, взаимосвязи и структуры изложения учебного материала, неумения обучающегося устанавливать и прослеживать причинно-следственные связи между событиями, процессами и явлениями, о которых идет речь в вопросах билета;
- отсутствия у обучающегося аргументации, логически и нормативно обоснованной точки зрения при освещении проблемных, дискуссионных аспектов учебного материала по вопросам билета;
- невозможности обучающегося дать ответы на дополнительные вопросы преподавателя.

Любой из указанных недостатков или их определенная совокупность могут служить основанием для выставления обучающемуся оценки «удовлетворительно».

Оценка «неудовлетворительно» при приеме экзамена выставляется в случае:

- отказа обучающегося от ответа по билету с указанием, либо без указания причин;
- невозможности изложения обучающимся учебного материала по двум или всем вопросам билета;
- допущения обучающимся существенных ошибок при изложении учебного материала по двум или всем вопросам билета;
- скрытое или явное использование обучающимся при подготовке к ответу нормативных источников, основной и дополнительной литературы, конспектов лекций и иного вспомогательного материала, кроме случаев специального указания или разрешения преподавателя;
- невладения обучающимся понятиями и категориями данной дисциплины;
- невозможность обучающегося дать ответы на дополнительные вопросы преподавателя;

Любой из указанных недостатков или их совокупность могут служить основанием для выставления обучающемуся оценки «неудовлетворительно».

Обучающийся имеет право отказаться от ответа по выбранному билету с указанием, либо без указания причин и взять другой билет. При этом с учетом приведенных выше критериев оценка обучающемуся должна быть выставлена на один балл ниже заслуживаемой им.

Дополнительные вопросы могут быть заданы обучающемуся в случае:

- необходимости конкретизации и изложенной обучающимся информации по вопросам билета с целью проверки глубины знаний отвечающего по связанным между собой темам и проблемам;
- необходимости проверки знаний обучающегося по основным темам и проблемам курса при недостаточной полноте его ответа по вопросам билета.

При проведении промежуточной аттестации в форме тестирования, оценивание результата проводится следующим образом:

«Отлично» - получают обучающиеся в том случае, если верные ответы составляют от 80% до 100% от общего количества

«Хорошо» - получают обучающиеся в том случае, если верные ответы составляют от 71 до 79% от общего количества;

«Удовлетворительно»- получают обучающиеся в том случае, если верные ответы составляют 50 –70 % правильных ответов;

«Неудовлетворительно» - работа, содержащая менее 50% правильных ответов.

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

Основная литература

1. Зюзин, А. С. Мировые информационные ресурсы : учебное пособие / А. С. Зюзин, К. В. Мартиросян. — Ставрополь : Северо-Кавказский федеральный университет, 2016. — 139 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/66030.html> — Режим доступа: для авторизир. пользователей

Дополнительная литература

1. Владимирова И. С., Александровская Ю. П., Филиппова Н. К., Гадельпина Г. А. Информационные технологии в экономике и управлении: учебное пособие. – Казань: Издательство КНИТУ, 2014 г. - 112 с.
https://lib.biblioclub.ru/book_428687_informatsionnyie_tehnologii_v_ekonomike_i_upravlenii/

2. Вылегжанина, А.О. Информационно-технологическое и программное обеспечение управления проектом : учебное пособие / А.О. Вылегжанина. - Москва ; Берлин : Директ-Медиа, 2015. - 429 с. : ил., схем., табл. - Библиогр. в кн. - ISBN 978-5-4475-4462-1; [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=362892>

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", справочных систем и профессиональных баз данных, необходимых для освоения дисциплины

1. ЭБС «Университетская библиотека онлайн» <https://biblioclub.ru>
2. ЭБС IPR BOOKS www.iprbookshop.ru
3. СПС «Консультант Плюс»

8. Лицензионное программное обеспечение

- Notepad++ 7.5.8
- Oracle Java SE 8u181
- Visual Studio Community 2017
- MS Windows 7 Профессиональная
- MS Windows 10 Pro

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

В зависимости от вида проводимых занятий используется следующее материально-техническое обеспечение дисциплины:

- лекционные аудитории (оборудованные видеопроекционным оборудованием для презентаций, средствами звуковоспроизведения, экраном и имеющие выход в Интернет) - аудитория № 720, аудитория № 737.

- специализированные помещения для проведения практических занятий по дисциплине – компьютерные классы с демонстрационно-обучающими и обучающе-контролирующими возможностями, доступом к базам данных и Интернет: № 728 (компьютерный класс № 4), № 729 (компьютерный класс № 5);

- библиотека (имеющая читальные залы и рабочие места для студентов, оснащенные компьютерами с доступом к базам данных и Интернет).

Для инвалидов и лиц с ограниченными возможностями здоровья форма проведения занятий по дисциплине устанавливается образовательной организацией с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья. При определении формы проведения занятий с обучающимся-инвалидом образовательная организация должна учитывать рекомендации, данные по результатам медико-социальной экспертизы, содержащиеся в индивидуальной программе реабилитации инвалида, относительно рекомендованных условий и видов труда. При необходимости для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья создаются специальные рабочие места с учетом нарушенных функций и ограничений жизнедеятельности.