

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Локтионова Оксана Геннадьевна
Должность: проректор по учебной работе
Дата подписания: 18.10.2024 12:57:17
Уникальный программный ключ:
0b817ca911e6668abb13a5d426d39e5f1c11eabbf73e943df4a4851fda56d089

МИНОБРАЗОВАНИЯ РОССИИ

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра программной инженерии

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Локтионова
«_____» _____ 2024 г.



Анализ данных в цифровой экономике

Методические указания для самостоятельной работы по
дисциплине «Математическая экономика» для студентов
направления подготовки 02.03.03 «Математическое обеспечение
и администрирование информационных систем»

Курск 2024

УДК 004

Составитель: Халин Ю.А.

Рецензент

Доктор физико-математических наук, профессор В.П. Добрица

Анализ данных в цифровой экономике: методические указания для самостоятельной работы / Юго-Зап. гос. ун-т; сост.: Ю.А. Халин. – Курск, 2024. – 9 с.: Библиогр.: с. 9.

Содержат сведения по вопросам самостоятельной работы на протяжении изучения дисциплины. Указывается порядок выполнения самостоятельных работ, содержание работ.

Предназначены для студентов направления подготовки 02.03.03 «Математическое обеспечение и администрирование информационных систем».

Текст печатается в авторской редакции

Подписано в печать *9.10.24*. Формат 60x84 1/16.

Усл. печ.л. 0,7. Уч. –изд.л. 0,47. Тираж 50 экз. Заказ *1127*.

Бесплатно.

Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

Содержание самостоятельной работы

	Тема СРС	Задание
1	Тема 1. Введение в анализ данных	Понятие данных и их основные типы. Знания и основные модели их представления: логические, семантические сети, продукционные системы, фреймы. Понятие базы данных. Модели данных. Этапы проектирования баз данных. Программные средства СУБД.
2	Тема 2. Статистические методы анализа данных	Выборка и её статистические характеристики: среднее, дисперсия, среднее квадратическое отклонение. Закон больших чисел. Проверка статистических гипотез.
3	Тема 3. Корреляционный и регрессионный анализ данных	Корреляция. Коэффициент парной корреляции. Коэффициент детерминации. Коэффициент множественной корреляции. Линейная регрессия. Определение параметров линейной регрессии методом наименьших квадратов. Оценка адекватности регрессионных моделей. Построение многофакторной регрессионной модели и проверка её адекватности. Коэффициент эластичности. Нелинейные регрессионные модели и их линейризация.
4	Тема 4. Кластерный и дискриминантный анализ	Понятие кластера. Основные метрики и их аксиомы. Алгоритмы определения центров. Примеры экономических задач кластерного анализа. Понятие и назначения дискриминантного анализа. Распознавание образов. Дискриминантная функция. Применение дискриминантного анализа в экономических задачах. Модель Альтмана. Программные средства статистического анализа данных. Электронные таблицы. SPSS. Statistica. Статистический анализ на языках R и Python.

5	Тема 5. Интеллектуальный анализ данных	Понятие и назначения интеллектуального анализа. Технология Data Mining. Классификация методов. Отличие от статистических методов. Примеры использования в цифровой экономике. Ассоциативные правила. Алгоритм Apriori. Применение алгоритмов поиска ассоциативных правил в маркетинговом анализе.
6	Тема 6. Машинное обучение	Понятие машинного обучения. Алгоритмы обучения с учителем. Алгоритмы обучения без учителя. Обучение с подкреплением. Глубокое обучение. Библиотеки Python, используемые в машинном обучении.

КОНТРОЛЬНЫЕ ВОПРОСЫ

– Тема 1: Основные понятия анализа данных

1. Основную цель анализа данных?
2. Приведите наиболее распространенные определения информации.
3. Приведите формализованное описание факта.
4. Что называется фактографическими данными?
5. Перечислите основные составляющие экспертной системы?
6. Назовите основные задачи, решаемые средствами ИАД?
7. Перечислите принципы анализа данных, предложенные Дж. Тьюки?
8. Назовите три уровня анализа информации. Чем они отличаются?
9. Чем отличается классификация от кластеризации?
10. Назовите различие между данными и знаниями.
11. Понятие моделей объектов анализа данных.
12. Виды моделей для анализа и прогнозирования.
13. Сущность и условия применения методов прогнозирования.
14. Цели и задачи прогнозирования.
15. Методологические основы прогнозов и планов и их сопряжения
16. Система методов прогнозирования и планирования.
17. Принципы выбора методов прогнозирования и планирования.
18. Классификация прогнозов и методов прогнозирования.
19. Характеристики методов анализа и прогнозирования.
20. Современные методы анализа данных.
21. Современные методы анализа данных.

– Тема: 2 Кластерный анализ

22. Как реализуются методы классификации, основанные на оценках плотностей распределения значений признаков?
23. Как реализуются логические методы классификации?
24. Перечислите методы оценивания информативности признаков
25. Перечислите методы способности признаков разделять объекты.
26. В чем состоит метод сравнения расстояний?
27. В чем состоит метод накопленных частот?
28. Назовите наиболее распространенные методы оценки информативности энтропийном подходе.
29. Опишите метод Шеннона.
30. Что определяет индекс Джини?
31. то называется дивергенцией Кульбака?

32. Сравните методы анализа информативности признаков.
33. Какие методы для выделения кластеров информативных признаков существуют?
34. В чем заключается идея простого алгоритма построения кластеров?
35. Что такое пороговое значение, в чем заключается смысл?
36. Как влияет выбор величины порогового значения на количество кластеров?
37. Сформулируйте, от чего зависит работа простого алгоритма построения кластеров?
38. В чем заключается идея алгоритма максимального расстояния для выделения кластеров?
39. Какой критерий оценки расстояний используется для анализа работы алгоритма максиминного расстояния?
40. В чем заключается идея алгоритма K-внутригрупповых средних для построения кластеров и выделения их центров?
41. Как осуществляется коррекция назначенных центров кластеров алгоритма K - внутригрупповых средних?
42. Сформулировать критерий окончания процесса итераций алгоритма K-внутригрупповых средних для построения кластеров и выделения их центров.

– **Тема 3: Метод нейронных сетей**

43. Понятие нейрона. Построение и настройка нейросети.
44. Прогнозирование на основе нейросетевого моделирования.
45. Приведите примеры для каких областей знаний нейрокомпьютеры применяются.
46. Сформулируйте преимущества нейрокомпьютеров?
47. Что собой представляют нейронные сети?
48. Приведите пример нейронные сети прямого распространения.
49. Какие черты живых нейронных сетей положены в основу искусственных нейронных сетей?
50. В чем заключаются особенности коннекционистских моделей обработки информации?
51. Что значит задать нейронную сеть для решения конкретной задачи?
52. В чем заключаются особенности символично-алгоритмической парадигмы.
53. Из чего состоят биологические нейроны?
54. Приведите основную схему искусственного нейрона.
55. Какие требования предъявляются к входным сигналам, поступающим в искусственный нейрон?

56. Какие требования предъявляются к весовым коэффициентам, поступающим в память
57. нейрона?
58. Что представляет собой сумматор в схеме искусственного нейрона?
59. В чем заключается назначение функция активации?

– **Тема 4: Основы регрессионного анализа**

60. Методы изучения взаимосвязей.
61. Определение наличия взаимосвязи факторов.
62. Коэффициент корреляции.
63. Корреляционно-регрессионный метод анализа данных.
64. Построение модели регрессии.
65. Виды моделей регрессии.
66. Оценка значимости параметров модели.
67. Оценка значимости уравнения регрессии.
68. Модель парной линейной регрессии и порядок ее построения .
69. Сущность и условия построения модели.
70. Прогнозирование на основе модели регрессии.
71. Показатели оценки точности модели регрессии.
72. Доверительный интервал прогноза.
73. Принятие решений на основе уравнения регрессии.
74. Проверка значимости каждого коэффициента регрессии осуществляется с помощью. t-критерия Стьюдента.

– **Тема 5 : Дисперсионный анализ**

75. Каково назначение дисперсионного анализа?
76. Как формулируются гипотезы в дисперсионном анализе?
77. Какие ограничения имеет дисперсионный анализ?
78. Однофакторный дисперсионный анализ с одинаковым числом испытаний
79. на различных уровнях.
80. Однофакторный дисперсионный анализ с различным числом испытаний на различных уровнях.
81. Двухфакторный дисперсионный анализ
82. .и многофакторный дисперсионный анализ.
83. Критерий адекватности Фишера.

– **Тема 6: Когнитивный анализ**

84. Основные положения когнитивного анализа. Типы концептов
85. Категоризация, дифференциальные и классификационные когнитивные признаки

86. Задачи семантико-когнитивного анализа

87. Методы и приемы семантико-когнитивного исследования концептов

– **Тема 7: Методы классификации**

88. Поясните значение термина «классификация».

89. Что называется кластеризацией?

90. Когда используется термин таксономия?

91. Что называется признаками?

92. Приведите формализованную постановку задачи классификации?

93. Что такое признаковое пространство?

94. Что называется решающим правилом?

95. Приведите алгоритм типовой системы классификации.

96. Что называется разбиением? Покрытием? неполным покрытием?

97. Приведите классификацию методов распознавания по степени формализации решающих правил.

98. Приведите классификацию методов распознавания по используемому математическому аппарату.

– **Тема 8: Интеллектуальный анализ данных**

99. Какие системы интеллектуального анализа существуют

100. Что такое OLAP-технологии и системы OLAP-анализа. Где применяются.

101. Чем отличается хранилище данных от базы данных?

102. Что такое многомерное представление данных?

103. Что такое кубы данных?

ПЕРЕЧЕНЬ ЛИТЕРАТУРЫ И ИНФОРМАЦИОННЫХ РЕСУРСОВ

1. Григорьев, А.А. Методы и алгоритмы обработки данных : учебное пособие для студентов экономических специальностей / А. А. Григорьев, Е. А. Исаев. - 2-е изд., перераб. и доп. - Москва : ИНФРА-М, 2022. - 383 с. - Текст : непосредственный.
2. Крутиков, В. Н. Анализ данных : учебное пособие / В. Н. Крутиков, В. В. Мешечкин. - Кемерово : Кемеровский государственный университет, 2014. - 138 с. - URL: <https://biblioclub.ru/index.php?page=book&id=278426> (дата обращения: 17.02.2024) . - Режим доступа: по подписке. - Текст : электронный.
3. Балдин, К. В. Эконометрика : учебное пособие / К. В. Балдин, О. Ф. Быстров, М. М. Соколов. - 2-е изд., перераб. и доп. - Москва : Юнити-Дана, 2017. - 255 с. - URL: <http://biblioclub.ru/index.php?page=book&id=684636> (дата обращения: 17.02.2024) . - Режим доступа: по подписке. - Текст : электронный.
4. Дятлов, А. В. Анализ данных в социологии : учебник / А. В. Дятлов, Д. А. Гугуева ; Министерство науки и высшего образования Российской Федерации ; Южный федеральный университет. - Ростов-на-Дону ; Таганрог : Южный федеральный университет, 2018. - 227 с. - URL: <https://biblioclub.ru/index.php?page=book&id=560998> (дата обращения: 17.02.2024) . - Режим доступа: по подписке. - Текст : электронный.
5. Кийко, П. В. Эконометрика. Продвинутый уровень : учебное пособие / П. В. Кийко, Н. В. Щукина. - Москва ; Берлин : Директ-Медиа, 2015. - 61 с. - URL: <http://biblioclub.ru/index.php?page=book&id=279003> (дата обращения: 17.02.2024) . - Режим доступа: по подписке. - Текст : электронный.
6. Агалаков, С. А. Анализ данных в среде R : практикум / С. А. Агалаков. - Омск : Омский государственный университет им. Ф.М. Достоевского (ОмГУ), 2020. - 52 с. - URL: <https://biblioclub.ru/index.php?page=book&id=614033> (дата обращения: 17.02.2024) . - Режим доступа: по подписке. - Текст : электронный.