|  |  |
| --- | --- |
| **Аннотация** |  |
| **Рабочей программы дисциплины** |  |
|  |  |  |  |
| Кафедра бизнес-информатики  |
| Дисциплина  | Интеллектуальный анализ данных  |
| Направление подготовки  | 09.04.03 Прикладная информатика  |
| Профиль  | Цифровая бизнес-аналитика  |
| Объем дисциплины  | 8 з.е  |
| Формы промежуточной аттестации  | Зачет |
|  |  | Экзамен, Курсовая работа |
| **Краткое** **содержание** **дисциплины**  |
| Тема | Наименование темы |
| Тема 1. | Интеллектуальный анализ данных (BDA). Основные понятия. |
| Тема 2. | Основные алгоритмы интеллектуального анализа данных. |
| Тема 3. | Основные инструментальные средства интеллектуального анализа данных. |
| Тема 4. | Понятие о принципах управления аналитической деятельностью на основе интеллектуального анализа данных. |
| Тема 5. | Комплексная работа по интеллектуальному анализу данных в цифровой экономике. |
|  |  |  |  |
| **Список** **литературы**  |
|  |  |  |  |
| **Основная литература:** |
| 1. Варфоломеева А. О., Коряковский А. В., Романов В. П. Информационные системы предприятий:учебное пособие для студентов вузов, обучающихся по направлению "Прикладная информатика" и другим экономическим специальностям. - Москва: ИНФРА-М, 2013. - 283 с. |
| 2. Вейнберг Р. Р. Интеллектуальный анализ данных и систем управления бизнес-правилами в телекоммуникациях:монография. - Москва: ИНФРА-М, 2016. - 173 с. |
| 3. Дадян Э. Г., Зеленков Ю. А. Методы, модели, средства хранения и обработки данных.:учебник. - Москва: Вузовский учебник: ИНФРА-М, 2017. - 168 с. |
| 4. Вейнберг Р. Р. Интеллектуальный анализ данных и систем управления бизнес-правилами в телекоммуникациях:монография. - Москва: ИНФРА-М, 2016. - 173 с. |
| 5. Богданов Е.П. Интеллектуальный анализ данных [Электронный ресурс]:Учебное пособие. - Волгоград: ФГБОУ ВПО Волгоградский государственный аграрный университет, 2019. - 112 – Режим доступа: https://znanium.com/catalog/product/1087885 |
| **Дополнительная литература:** |
| 1. Затонский А. В. Информационные технологии: разработка информационных моделей и систем:учебное пособие для студентов вузов, обучающихся по направлению 23.01.00 "Информатика и вычислительная техника". - Москва: РИОР: ИНФРА-М, 2014. - 344 с. |
| 2. Гурвиц Дж. С., Ньюджент А. Ф., Халпер Ф., Кауфман М. А. Просто о больших данных:перевод с английского. - Москва: Сбербанк: [Эксмо], 2015. - 395 с. |
|  |  |  |  |
| **Перечень** **информационных** **технологий,** **включая** **перечень** **лицензионного** **программного** **обеспечения** **и** **информационных** **справочных** **систем,** **онлайн** **курсов,** **используемых** **при**  |

|  |
| --- |
| **осуществлении** **образовательного** **процесса** **по** **дисциплине**  |
| **Перечень лицензионного программного обеспечения:** |
| Microsoft Office 2016.Договор № 52/223-ПО/2020 от 13.04.2020, Акт № Tr000523459 от 14.10.2020 Срок действия лицензии 30.09.2023.  |
| МойОфис стандартный. Соглашение № СК-281 от 7 июня 2017. Дата заключения - 07.06.2017. Срок действия лицензии - без ограничения срока.  |
| R Studio (среда для языка программирования R).Лицензия GNU Affero General Public License v3.Срок действия лицензии - без ограничения срока.  |
| Язык программирования R.Лицензия GNU GPL 2.Срок действия лицензии - без ограничения срока.  |
| **Перечень информационных справочных систем, ресурсов информационно- телекоммуникационной сети «Интернет»:** |
|   |
| **ИАД** https://www.intuit.ru/studies/courses/2312/612/info  |
| Аннотацию подготовил: Галактионов А.Д.  |

**Перечень тем курсовых работ**

1. Основные этапы разработки интеллектуальных систем бизнес-анализа;
2. Подготовка исходных данных к машинному обучению;
3. Создание Хранилища данных и организация доступа;
4. Оценка качества данных. Профайлинг и аудит данных;
5. Проблема пропущенных данных и ее возможные решения;
6. Редактирование выбросов и аномальных значений;
7. Поиск и редактирование дубликатов и противоречий;
8. Спектральная обработка ряда;
9. Преобразование Дата/время при подготовке временного ряда;
10. Группировка и разгруппировка данных. Задачи, решаемые группировкой;
11. Квантование данных. Назначение квантования. Параметры квантования, определение границ интервалов и задание меток;
12. Создание новых полей с помощью калькулятора;
13. Слияние наборов данных. Типы объединения;
14. Изменение структуры таблицы. Кросс-таблица и свертка столбцов;
15. Создание репрезентативного множества. Сэмплинг и Разбиение на множества. Методы выборки данных. Размеры обучающего и тестового множества;
16. Подготовка временного ряда к прогнозированию. Скользящее окно. Критерии выбора погружения в ряд;
17. Отбор переменных в модель логистической регрессии. Конечные классы. WoE-анализ. Метрики классов и значимость входных признаков;
18. Применение скриптов в сценариях обработки данных;
19. Групповая обработка данных по похожим сценариям;
20. Применение переменных в сценариях обработки данных;
21. Оценка степени периодичности ряда с помощью метода автокорреляции;
22. Выявление закономерностей между связанными событиями. Ассоциативные правила. Шаблоны покупок. Достоверность и Лифт АП;
23. Задачи, решаемые с помощью кластеризации. Кластеризация k-means, g-means. Область применения;
24. ЕМ-кластеризация. Автоматическое определение кластеров. Оценка на основе логарифмической функции правдоподобия;
25. Кластеризация транзакций. Эвристический алгоритм CLOPE. Глобальная функция стоимости. Область применения кластеризации транзакций;
26. Самоорганизующиеся карты (Self Organizing Maps - SOM). Кластеризации многомерных векторов – алгоритм проецирования с сохранением топологического подобия. Инициализация начальных весов. Область применения карт Кохонена.
27. Линейная регрессия. Коэффициенты регрессии. Область применения;
28. Логистическая регрессия. Прогнозирование бинарной переменной. Коэффициенты регрессии. ROC-анализ. Lift-анализ. Оценка качества модели;
29. Дерево решений. Нормализация полей. Обучающая выборка. Параметры обучения. Значимость атрибутов. Правила. Что-если. Таблица сопряженности;
30. Нейронная сеть. Требования к обучающей и тестовой выборке. Задание структуры нейронной сети. Скрытые слои и активационная функция. Определение числа связей и переобучение сети. Область применения нейронных сетей.