

Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Курганский государственный университет»
(КГУ)

Кафедра «Программное обеспечение автоматизированных систем»

УТВЕРЖДАЮ:
Проректор по образовательной и
международной деятельности
_____ /А.А. Кирсанкин/
«____» _____ 2026 г.

Рабочая программа учебной дисциплины

Методы и алгоритмы анализа данных
(наименование дисциплины)

образовательной программы высшего образования –
программы бакалавриата
35.03.06 - Агроинженерия

Направленность:
Информационные технологии в электроэнергетике

Формы обучения: очная

Курган 2026

Рабочая программа дисциплины «Методы и алгоритмы анализа данных» составлена в соответствии с учебным планом по программе бакалавриата Агроинженерия (Информационные технологии в электроэнергетике), утвержденным:

- для очной формы обучения «26» июня 2026 года.

Рабочая программа дисциплины одобрена на заседании кафедры «Программное обеспечение автоматизированных систем» «29» июня 2026 года, протокол № 12.

ст. преподаватель _____ Ю.В. Адаменко

Согласовано:

Заведующий
кафедрой ПОАС _____ С.В. Косовских

Заведующий кафедрой
«Цифровая энергетика» _____ Ж.В. Нечеухина

Специалист по учебно-методической
работе учебно-методического отдела _____ Г.В. Казанкова

Начальник управления образовательной
деятельности _____ И.В. Григоренко

1. ОБЪЕМ ДИСЦИПЛИНЫ

Всего: 3 зачетные единицы трудоемкости (108 академических часа)

Очная форма обучения

Вид учебной работы	На всю дисциплину	Семестр
		5
Аудиторные занятия (контактная работа с преподавателем), всего часов	64	64
в том числе:		
Лекции	32	32
Лабораторные работы	32	32
Самостоятельная работа, всего часов	44	44
в том числе:		
Подготовка к зачету	18	18
Другие виды самостоятельной работы (самостоятельное изучение тем (разделов) дисциплины)	26	26
Вид промежуточной аттестации	Зачет	Зачет
Общая трудоемкость дисциплины и трудоемкость по семестрам, часов	108	108

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «Методы и алгоритмы анализа данных» к дисциплинам части, формируемой участниками образовательных отношений блока 1.

Изучение дисциплины базируется на результатах обучения, сформированных при изучении следующих дисциплин: «Информатика», «Основы программирования», «Математика».

Результаты обучения по дисциплине необходимы при выполнении научно-исследовательской работы, а также при подготовке выпускной квалификационной работы.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ

Цель освоения дисциплины: изучение современных математических методов и получение практических навыков анализа исходной информации, применительно к задачам, возникающим при исследовании сложных систем.

Задачи дисциплины: используя ресурсы образовательной программы, университетского образовательного пространства, профессионального сообщества способствовать формированию у студентов навыков решения прикладных задач, связанных с анализом данных, обработкой результатов эксперимента с использованием современных программных средств и технологий.

Компетенции, формируемые в результате освоения дисциплины.

- Способен использовать программное обеспечение для моделирования, анализа, расчета, обработки и защиты информации в электроэнергетике (ПК-3).

В результате изучения дисциплины обучающийся должен

знать:

- принципы работы современных информационных технологий и программных средств для моделирования и анализа информации в электроэнергетике (для ПК-3);

уметь:

- применять современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности (для ПК-3);

владеть:

- современными методами и программными средствами для разработки приложений и создания программных прототипов решения прикладных задач в электроэнергетике (для ПК-3).

Индикаторы и дескрипторы части соответствующей компетенции, формируемой в процессе изучения дисциплины «Методы и алгоритмы анализа данных», оцениваются при помощи оценочных средств.

Планируемые результаты обучения по дисциплине «Методы и алгоритмы анализа данных»,
индикаторы достижения компетенций ПК-3 перечень оценочных средств

№ п/п	Код индикатора достижения компетенции	Наименование индикатора достижения компетенции	Код планируемого результата обучения	Планируемые результаты обучения	Наименование оценочных средств
1.	ИД-1 _{ПК-3}	Знать: принципы работы современных информационных технологий и программных средств для моделирования и анализа информации в электроэнергетике	З (ИД-1 _{ПК-3})	Знает: принципы работы современных информационных технологий и программных средств для моделирования и анализа информации в электроэнергетике	Отчеты по лабораторным работам Вопросы для сдачи зачета
2.	ИД-2 _{ПК-3}	Уметь: применять современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности	У (ИД-2 _{ПК-3})	Умеет: применять современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности	Отчеты по лабораторным работам Вопросы для сдачи зачета
3.	ИД-3 _{ПК-3}	Владеть: средствами организации поиска, анализа, синтеза, системного подхода при решении поставленных задач	В (ИД-3 _{ПК-3})	Владеет: средствами организации поиска, анализа, синтеза, системного подхода при решении поставленных задач	Отчеты по лабораторным работам Вопросы для сдачи зачета

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1. Учебно-тематический план

Очная форма обучения

Рубеж	Номер раздела, темы	Наименование раздела, темы	Количество часов контактной работы с преподавателем	
			Лекции	Лабораторные работы
Рубеж 1	1.	Интеллектуальный анализ данных	2	
	2.	Классификация и регрессия	4	4
	3.	Поиск ассоциативных правил	4	4
	4.	Кластеризация	4	4
		Рубежный контроль № 1		2
Рубеж 2	5.	Визуальный анализ данных (Visual Mining)	2	4
	6.	Анализ текстовой информации (Text Mining)	4	4
	7.	Стандарты Data Mining	2	
	8.	Data Mining в реальном времени	4	
	9.	Извлечение знаний из Web (Web Maining)	4	4
	10.	Средства анализа процессов (Process Maining)	2	4
		Рубежный контроль № 2		2
Всего			32	32
Итого			32	32

4.2. Содержание лекционных занятий

Тема 1 Интеллектуальный анализ данных

Добыча данных (Data Mining). Задачи и практическое применение Data Mining. Модели и методы Data Mining. Процесс обнаружения знаний. Управление знаниями. Средства Data Mining.

Тема 2 Классификация и регрессия

Постановка задачи. Представление результатов: правила классификации, деревья решений, математические функции. Методы построения правил классификации: алгоритм построения 1-правил, метод Naïve Bayes. Методы построения деревьев решений: методика "разделяй и властвуй", алгоритм покрытия. Методы построения математических функций: метод наименьших квадратов, нелинейные методы, метод опорных векторов (SVM). Прогнозирование временных рядов.

Тема 3 Поиск ассоциативных правил

Постановка задачи. Секвенциальный анализ. Представление результатов. Алгоритмы: алгоритм Apriori, разновидности алгоритма Apriori.

Тема 4 Кластеризация

Постановка задачи кластеризации. Меры близости, основанные на расстояниях, используемые в алгоритмах кластеризации. Представление результатов. Базовые алгоритмы кластеризации. Классификация алгоритмов. Алгоритм k-means. Адаптивные методы кластеризации.

Тема 5 Визуальный анализ данных (Visual Mining)

Выполнение визуального анализа данных. Характеристики средств визуализации данных. Методы визуализации: методы геометрических преобразований, отображение иконок, методы, ориентированные на пиксели, иерархические образы.

Тема 6 Анализ текстовой информации (Text Mining)

Задача анализа текстов: этапы анализа текстов, предварительная обработка текста, задачи Text Mining. Извлечение ключевых понятий из текста: общее описание процесса извлечения понятий из текста, стадия локального анализа, стадия интеграции и вывода понятий. Классификация текстовых документов. Методы классификации текстовых документов: представление текстовых документов, иерархические методы кластеризации текстов, бинарные методы кластеризации текстов. Задача аннотирования текстов. Средства анализа текстовой информации

Тема 7 Стандарты Data Mining

Области стандартизации Data Mining. Стандарт CWM. Структура и состав CWM. Пакет Data Mining. Стандарт CRISP. Структура стандарта CRISP. Стандарт PMML. Другие стандарты Data Mining: стандарт SQL/MM, стандарт Microsoft Data Mining extensions (DMX), стандарт Java Data Mining.

Тема 8 Data Mining в реальном времени

Использование Data Mining в реальном времени. Адаптивная добыча данных. Применение Data Mining в реальном времени. Рекомендательные машины: классификация рекомендательных машин, подход на основе содержания, совместное фильтрование. Анализ рыночной корзины и секвенциальный анализ. Усиление обучения и агенты. Инструменты Data Mining в реальном времени.

Тема 9 Извлечение знаний из Web (Web Mining)

Проблемы анализа информации из Web. Этапы Web Mining. Web Mining и другие интернет-технологии. Категории Web Mining. Методы извлечения Web-контента. Извлечение Web-контента для формирования баз данных. Извлечение Web-структур. Поиск Web-документов с учетом гиперссылок. Кластеризация Web-структур. Исследование использования Web-ресурсов.

Тема 10 Средства анализа процессов (Process Mining)

Автоматизация выполнения бизнес-процессов: понятие бизнес-процесса, формализация бизнес-процессов, Workflow-системы, проектирование бизнес-процессов. Анализ про-

цессов: технология Process Mining, анализ протоколов, стандарт записи протоколов MXML, задачи Process Mining. Методы Process Mining. Библиотека алгоритмов Process Mining – ProM.

4.3. Лабораторные занятия

Номер раздела, темы	Наименование раз-дела, темы	Наименование лабораторной работы	Норматив вре-мени, час.
			Очная форма обучения
2	Классификация и ре-грессия	Лабораторная работа № 1. Решение задач класси-фикации и регрессии	4
3	Поиск ассоциативных правил	Лабораторная работа № 2. Нахождение взаимосвязей в массивах данных.	4
4	Кластеризация	Лабораторная работа № 3. Кластеризация исходных данных	4
	Рубежный контроль 1.		2
5	Визуальный анализ данных (Visual Mining)	Лабораторная работа № 4. Элементы визуального анализа данных	4
6	Анализ текстовой ин-формации (Text Mining)	Лабораторная работа № 5. Элементы анализа тексто-вой информации	4
9	Извлечение знаний из Web (Web Maining)	Лабораторная работа № 6. Работа с Web	4
10	Средства анализа процессов (Process Maining)	Лабораторная работа № 7. Сети Петри	4
	Рубежный контроль 2.		2
Всего			32
Итого			32

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Лекционный курс базируется на пассивном методе обучения, реализующем традиционную объяснительно-иллюстративную образовательную технологию, в рамках которой обучающиеся выступают в роли слушателей, воспринимающих учебный материал, и участвующих в дискуссиях и экспресс-опросах.

При прослушивании лекций рекомендуется в конспекте отмечать все важные моменты, на которых заостряет внимание преподаватель, в частности те, которые направлены на качественное выполнение соответствующей лабораторной работы.

Преподавателем запланировано использование при чтении лекций технологии учебной дискуссии. Поэтому рекомендуется фиксировать для себя интересные моменты с целью их активного обсуждения на дискуссии в конце лекции.

Залогом качественного выполнения лабораторных работ является самостоятельная подготовка путем повторения материалов лекций. Рекомендуется подготовить вопросы по неясным моментам и обсудить их с преподавателем в начале каждого занятия.

Преподавателем запланировано применение на лабораторных занятиях технологий развивающейся кооперации, коллективного взаимодействия, разбора конкретных ситуаций.

Для текущего контроля успеваемости по очной форме обучения преподавателем используется балльно-рейтинговая система контроля и оценки академической активности. Поэтому настоятельно рекомендуется тщательно прорабатывать материал дисциплины при самостоятельной работе, участвовать во всех формах обсуждения и взаимодействия, как на лекциях, так и на лабораторных работах в целях лучшего освоения материала и получения высокой оценки по результатам освоения дисциплины.

Выполнение самостоятельной работы подразумевает самостоятельное изучение разделов дисциплины, подготовку к лабораторным работам, к рубежным контролям (для очной формы обучения), подготовку к зачету.

Рекомендуемая трудоемкость самостоятельной работы представлена в таблице.

Наименование вида самостоятельной работы	Рекомендуемая трудоемкость, акад. час.
	Очная форма обучения
Самостоятельное изучение тем дисциплины:	8
1. Визуальный анализ данных (Visual Mining). Характеристики средств визуализации данных. Методы визуализации: методы геометрических преобразований, отображение иконок, методы, ориентированные на пиксели, иерархические образы.	2
2 Анализ текстовой информации (Text Mining). Этапы анализа текстов, предварительная обработка текста. Извлечение ключевых понятий из текста.	2
3. Извлечение знаний из Web. Этапы Web Mining. Категории Web Mining. Методы извлечения Web-контента.	2
4. Автоматизация выполнения бизнес-процессов: понятие бизнес-процесса, формализация бизнес-процессов, Workflow-системы, проектирование бизнес-процессов. Анализ процессов	2
Подготовка к лабораторным занятиям (по 1 часу на каждое занятие (очная форма обучения))	16
Подготовка к рубежным контролям (по 1 часу на каждый рубеж)	2
Подготовка к зачету	18
Всего:	44

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ

6.1. Перечень оценочных средств

1. Балльно-рейтинговая система контроля и оценки академической активности обучающихся в КГУ (для очной формы обучения).
2. Отчеты обучающихся по лабораторным занятиям.
3. Банк тестовых заданий к рубежным контролям № 1, 2 (для очной формы обучения).
4. Список вопросов к зачету.

6.2. Система балльно-рейтинговой оценки работы обучающихся по дисциплине

№	Наименование	Содержание					
		Распределение баллов для зачета					
		Посещение лекций	Посещение лабораторных занятий	Выполнение лабораторных работ	Рубежный контроль № 1	Рубежный контроль № 2	Зачет
1	Распределение баллов за семестры по видам учебной работы, сроки сдачи учебной работы (доводятся до сведения обучающихся на первом учебном занятии)	До 16 баллов (16 * 1 балл = 16 б.)	До 8 баллов (16 * 0,5 балла = 8 б.)	До 28 баллов (7 * 4 балла = 28 б.)	До 9 баллов	До 9 баллов	До 30 баллов

№	Наименование	Содержание
2	Критерий пересчета баллов в традиционную оценку по итогам работы в семестре и зачета	- 60 и менее баллов – не зачтено (компетенции не освоены); - 61...73 – зачтено (пороговый уровень освоения компетенций); - 74...90 – зачтено (базовый уровень освоения компетенций); - 91...100 – зачтено (продвинутый уровень освоения компетенций).
3	Критерии допуска к промежуточной аттестации, возможности получения автоматического зачета (экзаменационной оценки) по дисциплине, возможность получения бонусных баллов	Для допуска к промежуточной аттестации по дисциплине за семестр обучающийся должен набрать по итогам текущего и рубежного контролей не менее 51 балла. В случае если обучающийся набрал менее 51 балла, то к аттестационным испытаниям он не допускается. Для получения зачета без проведения процедуры промежуточной аттестации обучающемуся необходимо набрать в ходе текущего и рубежных контролей не менее 61 балла. В этом случае итог балльной оценки, получаемой обучающимся, определяется по количеству баллов, набранных им в ходе текущего и рубежных контролей. При этом, на усмотрение преподавателя, балльная оценка обучающегося может быть повышена за счет получения дополнительных баллов за академическую активность. Обучающийся, имеющий право на получение оценки без проведения процедуры промежуточной аттестации, может повысить ее путем сдачи аттестационного испытания. В случае получения обучающимся на аттестационном испытании 0 баллов итог балльной оценки по дисциплине не снижается. За академическую активность в ходе освоения дисциплины, участие в учебной, научно-исследовательской, спортивной, культурно-творческой и общественной деятельности обучающемуся могут быть начислены дополнительные баллы. Максимальное количество дополнительных баллов за академическую активность составляет 30. Основанием для получения дополнительных баллов являются: - выполнение дополнительных заданий по дисциплине; дополнительные баллы начисляются преподавателем; - участие в течение семестра в учебной, научно-исследовательской, спортивной, культурно-творческой и общественной деятельности КГУ.
4	Формы и виды учебной работы для неуспевающих (восстановившихся на курсе обучения) обучающихся для получения недостающих баллов в конце семестра	В случае если к промежуточной аттестации зачету набрана сумма менее 51 балла, обучающемуся необходимо набрать недостающее количество баллов за счет выполнения дополнительных заданий, до конца последней (зачетной) недели семестра. Ликвидация академических задолженностей, возникших из-за разности в учебных планах при переводе или восстановлении, проводится путем выполнения дополнительных заданий, форма и объем которых определяется преподавателем

6.3. Процедура оценивания результатов освоения дисциплины

Рубежные контроли проводятся в форме письменных ответов на вопросы. Каждому обучающемуся предлагается по 2 вопроса, за ответ на который обучающийся может получить до 4,5 баллов (рубежи №1, 2).

Перед проведением каждого рубежного контроля преподаватель прорабатывает с обучающимися основной материал соответствующих разделов дисциплины в форме краткой лекции-дискуссии. Примерные варианты заданий для 1, 2 рубежного контроля приведены ниже. На каждый рубежный контроль обучающиеся отводится 2 академических часа.

Преподаватель оценивает в баллах результаты каждого обучающегося и заносит в ведомость учета текущей успеваемости.

Билеты на зачет состоят из 2 вопросов и практического задания. Ответы на каждый вопрос оцениваются до 10 баллов, выполнение практического задания оценивается до 10

баллов. Время, отводимое обучающемуся на подготовку к ответу на билет, составляет 1 астрономический час.

Результаты текущего контроля успеваемости (зачета) заносятся преподавателем в зачетную ведомость, которые сдаются в организационный отдел института в день зачета, а также выставляются в зачетную книжку обучающегося.

6.4. Примеры оценочных средств для рубежных контролей изачета

Примерные вопросы для рубежных контролей

Рубежный контроль 1: Список вопросов

1. Основные задачи Data Mining
2. . Модели и методы Data Mining
3. Средства Data Mining.
4. Классификация и регрессия. Краткая характеристика. Общее и различие.
5. Классификация и регрессия. Правила классификации
6. Классификация и регрессия. Деревья решений
7. Классификация и регрессия. Методы построения правил классификации: алгоритм построения 1-правил.
8. Классификация и регрессия. Методы построения правил классификации: метод Naïve Bayes.
9. Классификация и регрессия. Методы построения математических функций: метод наименьших квадратов.
10. Методы построения математических функций: нелинейные методы.
11. Методы построения математических функций: метод опорных векторов (SVM).
12. Поиск ассоциативных правил, общие положения.
13. Поиск ассоциативных правил. Секвенциальный анализ.
14. Поиск ассоциативных правил Алгоритм Apriori.
15. Постановка задачи кластеризации.
16. Кластеризация. Меры близости, основанные на расстояниях, используемые в алгоритмах кластеризации.
17. Кластеризация. Базовые алгоритмы кластеризации.
18. Кластеризация. Алгоритм k-means.
19. Кластеризация. Адаптивные методы кластеризации.

Рубежный контроль 2: Список вопросов

1. Визуальный анализ данных. Характеристики средств визуализации данных.
2. Визуальный анализ данных. Методы визуализации: методы геометрических преобразований
3. Визуальный анализ данных. Методы визуализации: отображение иконок.
4. Визуальный анализ данных. Методы визуализации: методы, ориентированные на пиксели, иерархические образы.
5. Задача анализа текстов: этапы анализа текстов, предварительная обработка текста, задачи Text Mining.
6. Извлечение ключевых понятий из текста: общее описание процесса извлечения понятий из текста, стадия локального анализа, стадия интеграции и вывода понятий.
7. Классификация текстовых документов.
8. Методы классификации текстовых документов: представление текстовых документов.
9. Методы классификации текстовых документов: иерархические методы кластеризации текстов.
10. Методы классификации текстовых документов: бинарные методы кластеризации текстов.
11. Задача аннотирования текстов.
12. Средства анализа текстовой информации.

13. Стандарты Data Mining. Стандарт CWM. Структура и состав CWM.
14. Стандарты Data Mining. Стандарт CRISP. Структура стандарта CRISP.
15. Использование Data Mining в реальном времени.
16. Рекомендательные машины.
17. Проблемы анализа информации из Web. Этапы Web Mining.
18. Категории Web Mining.
19. Методы извлечения Web-контента.
20. Извлечение Web-структур.
21. Поиск Web-документов с учетом гиперссылок.
22. Кластеризация Web-структур.
23. Автоматизация выполнения бизнес-процессов: понятие бизнес-процесса, формализация бизнес-процессов
24. Workflow-системы, проектирование бизнес-процессов.
25. Анализ процессов: технология Process Mining, анализ протоколов.
26. Методы Process Mining.

Примерный список вопросов к зачету.

1. Добыча данных (Data Mining). Задачи и практическое применение Data Mining.
2. Модели и методы Data Mining. Процесс обнаружения знаний.
3. Постановка задачи классификации и регрессии. Представление результатов (правила классификации, деревья решений, математические функции).
4. Методы построения правил классификации (алгоритм построения 1-правил, метод Naïve Bayes).
5. Методы построения деревьев решений (методика "разделяй и властвуй", алгоритм покрытия).
6. Методы построения математических функций (метод наименьших квадратов, нелинейные методы, метод опорных векторов (SVM)).
7. Прогнозирование временных рядов.
8. Управление знаниями. Средства Data Mining.
9. Постановка задачи поиска ассоциативных правил. Секвенциальный анализ.
10. Алгоритм Apriori, его разновидности.
11. Постановка задачи кластеризации.
12. Меры близости, основанные на расстояниях, используемые в алгоритмах кластеризации. Представление результатов.
13. Базовые алгоритмы кластеризации.
14. Классификация алгоритмов. Алгоритм k-means.
15. Адаптивные методы кластеризации.
16. Выполнение визуального анализа данных. Характеристики средств визуализации данных.
17. Методы визуализации (методы геометрических преобразований, отображение иконок, методы, ориентированные на пиксели, иерархические образы)
18. Задача анализа текстов: этапы анализа текстов, предварительная обработка текста, задачи Text Mining.
19. Извлечение ключевых понятий из текста: общее описание процесса извлечения понятий из текста, стадия локального анализа, стадия интеграции и вывода понятий.
20. Классификация текстовых документов. Методы классификации текстовых документов: представление текстовых документов, иерархические методы кластеризации текстов, бинарные методы кластеризации текстов.
21. Задача аннотирования текстов. Средства анализа текстовой информации
22. Области стандартизации Data Mining. Стандарт CWM. Структура и состав CWM.
23. Области стандартизации Data Mining. Стандарт CRISP. Структура стандарта CRISP.
24. Области стандартизации Data Mining. Стандарт PMML.
25. Области стандартизации Data Mining. Другие стандарты Data Mining: стандарт SQL/MM, стандарт Microsoft Data Mining extensions (DMX), стандарт Java Data Mining.
26. Использование Data Mining в реальном времени. Адаптивная добыча данных.

27. Применение Data Mining в реальном времени. Рекомендательные машины: классификация рекомендательных машин, подход на основе содержания, совместное фильтрование.
28. Применение Data Mining в реальном времени. Анализ рыночной корзины и секвенциальный анализ.
29. Применение Data Mining в реальном времени. Усиление обучения и агенты. Инструменты Data Mining в реальном времени.
30. Проблемы анализа информации из Web. Этапы Web Mining. Web Mining и другие интернет-технологии.
31. Категории Web Mining. Методы извлечения Web-контента. Извлечение Web-контента для формирования баз данных.
32. Извлечение Web-структур. Поиск Web-документов с учетом гиперссылок.
33. Кластеризация Web-структур. Исследование использования Web-ресурсов.
34. Автоматизация выполнения бизнес-процессов: понятие бизнес-процесса, формализация бизнес-процессов.
35. Workflow-системы, проектирование бизнес-процессов.
36. Анализ процессов: технология Process Mining, анализ протоколов, стандарт записи протоколов MXML.
37. Задачи и методы Process Mining.

Примерный список заданий к зачету.

Реализовать метод k-средних и выполнить кластеризацию образов.

Вариант 1

Образы: четырехугольники.

Признаки: четыре длины сторон четырехугольника.

Класс «маленький четырехугольник»: (2,2,2,2), (2,2,3,2), (2,3,2,2), (3,2,2,2), (3,3,2,2), (2,3,3,2), (3,2,3,2), (3,3,3,2), (2,2,2,3), (2,2,3,3), (2,3,2,3), (3,2,2,3), (3,3,2,3), (2,3,3,3), (3,2,3,3), (3,3,3,3).

Класс «большой четырехугольник»: (9,9,9,9), (9,9,11,9), (9,11,9,9), (11,9,9,9), (11,11,9,9), (11,9,11,9), (9,11,11,9), (11,11,11,9), (9,9,9,11), (9,9,11,11), (9,11,9,11), (11,9,9,11), (11,11,9,11), (11,9,11,11), (9,11,11,11).

Класс «очень маленький четырехугольник»: (0.1,0.1,0.1,0.1).

Вариант 2

Образы: книги.

Признаки: количество страниц, количество картинок, количество таблиц, средняя длина слова.

Класс «детская книга»: (200,50,0,4), (150,37,0,5).

Класс «учебник»: (400,15,90,8), (500,3,105,7).

Класс «журнал»: (80,200,5,6), (100,198,4,7).

Вариант 3

Образы: погода на сутки (летом).

Признаки: количество часов с дождем, количество солнечных часов, средняя температура воздуха, количество детей во дворе днем, скорость ветра.

Класс «хорошая погода»: (0,17,25,15,2), (1,16,24,13,3), (1,17,23,15,2).

Класс «плохая погода»: (20,0,18,1,7), (17,2,18,0,8), (18,1,19,2,7).

Вариант 4

Образы: погода на сутки (зимой).

Признаки: количество часов со снегом, количество солнечных часов, средняя температура воздуха, количество детей во дворе днем, скорость ветра.

Класс «хорошая погода»: (0,13,-10,15,2), (1,12,-12,13,3), (1,13,-13,15,2).

Класс «плохая погода»: (20,0,-21,1,8), (17,2,-24,0,9), (18,1,-23,2,8).

Вариант 5

Образы: животные.

Признаки: количество полосок, рост (см), длина хвоста (см).

Класс «пчела»: (7,1,0), (8,2,0), (6,1,0).

Класс «зебра»: (30,150,50), (35,160,60), (33,155,50).
Класс «тигр»: (60,100,100), (70, 110,110), (55,100,110).
Класс «муравей»: (0,0.5,0), (0,0.6,0).

Вариант 6

Образы: одежда.

Признаки: длина рукава, количество пуговиц, длина изделия.

Класс «футболка»: (15,0,60), (20,0,70).

Класс «рубашка»: (50,8,60), (60,9,60), (55,8,70).

Класс «летняя рубашка»: (15,6,60), (20,8,70), (20,7,65).

Класс «халат»: (15,9,120), (20,10,130), (25,9,115).

Вариант 7

Образы: школьники.

Признаки: оценка по алгебре, оценка по геометрии, оценка по физике, оценка по истории, оценка по литературе, оценка по иностранному языку, оценка по физкультуре.

Класс «отличник»: (5,5,5,5,5,5,5), (4,5,5,5,5,5,5), (5,4,5,5,5,5,5), (5,5,5,4,5,5,5), (5,5,5,5,4,5,5), (5,5,5,5,5,4,5), (5,5,5,5,5,5,4).

Класс «троечник»: (3,3,3,3,3,3,3), (4,3,3,3,3,3,3), (3,4,3,3,3,3,3), (3,3,3,4,3,3,3), (3,3,3,3,4,3,3), (3,3,3,3,3,4,3), (3,3,3,3,3,3,4).

Класс «математик»: (5,5,5,3,3,3,4), (5,5,5,3,4,3,3), (5,5,5,4,3,3,4), (5,5,5,3,4,4,3), (5,5,5,4,4,3,4), (5,5,5,4,3,4,3).

Класс «гуманитарий»: (3,3,3,5,5,5,4), (3,4,3,5,5,5,3), (4,3,3,5,5,5,4), (3,4,4,5,5,5,3), (4,4,3,5,5,5,4), (4,3,4,5,5,5,3).

Вариант 8

Образы: ягоды.

Признаки: количество семечек, радиус сечения (см), количество на одном растении.

Класс «вишня»: (1,2,300), (1,1.5,400), (1,1,350).

Класс «арбуз»: (50,20,1), (60,25,1), (40,21,1).

Класс «облепиха»: (6,0.3,500), (8,0.2,550), (8,0.3,600).

Класс «виноград»: (6,1,1500), (8,1.5,1400), (8,0.7,1700).

Вариант 9

Образы: пятиугольники.

Признаки: пять длин сторон пятиугольника.

Класс «маленький пятиугольник»: (2,2,2,2,2), (2,2,3,2,2), (2,3,2,2,3), (3,3,2,2,2), (2,3,3,2,3), (3,2,3,2,2), (3,3,3,2,3), (2,2,2,3,2), (3,3,3,3,3).

Класс «большой пятиугольник»: (9,9,9,9,9), (9,9,11,9,11), (9,11,9,9,9), (11,11,9,9,9), (11,9,11,9,11), (9,11,11,9,9), (11,11,11,9,9).

Класс «очень маленький пятиугольник»: (0.1,0.1,0.1,0.1,0.1).

Вариант 10

Образы: треугольники.

Признаки: три длины сторон треугольника.

Класс «маленький треугольник»: (2,2,2), (2,2,3), (2,3,2), (3,2,2), (3,3,2), (3,2,3), (3,3,3).

Класс «большой треугольник»: (9,9,9), (9,9,11), (9,11,9), (11,9,9), (11,11,9), (11,9,11), (9,11,11), (11,11,11).

Класс «очень маленький треугольник»: (0.1,0.1,0.1).

6.5. Фонд оценочных средств

Полный банк заданий для текущего, рубежных контролей и промежуточной аттестации по дисциплине, показатели, критерии, шкалы оценивания компетенций, методические материалы, определяющие процедуры оценивания образовательных результатов, приведены в учебно-методическом комплексе дисциплины.

7. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ УЧЕБНАЯ ЛИТЕРАТУРА

7.1. Основная учебная литература

1 Богданов, Е. П. Интеллектуальный анализ данных : практикум для магистрантов направления 09.04.03 «Прикладная информатика» профиль подготовки «Информационные системы и технологии корпоративного управления» / Е. П. Богданов. - Волгоград : ФГБОУ ВО Волгоградский ГАУ, 2019. - 112 с. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1087885>. – Режим доступа: по подписке.

2 Целых, А. Н. Современные методы прикладной информатики в задачах анализа данных : учебное пособие по курсу "Методы интеллектуального анализа данных" / А. Н. Целых, А. А. Целых, Э. М. Котов ; Южный федеральный университет. - Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2021. - 130 с. - ISBN 978-5-9275-3783-9. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1894428>. – Режим доступа: по подписке.

3 Зарова, Е. В. Методы Data mining в обработке и анализе статистических данных (решения в R) : монография / Е.В. Зарова. — Москва : ИНФРА-М, 2021. — 232 с. : ил. - ISBN 978-5-16-016814-2. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1240276>. – Режим доступа: по подписке.

7.2 Дополнительная литература

1. Барсегян А.А., Куприянов М.С., Холод И.И., Тесс М.Д., Елизаров С.И. Анализ данных и процессов 3-е издание. Учебная литература для вузов. — БХВ-Петербург, 2009 — 512 с. : ил.

2. Замятин А. В. Интеллектуальный анализ данных : учеб. пособие / А. В. Замятин. – Томск : Томск. гос. ун-т, 2016. – 119 с.

3. Журавлёв, Ю. И. Распознавание. Математические методы. Программная система. Практические применения / Ю. И. Журавлёв, В. В. Рязанов, О. В. Сенько. — М. : Фазис, 2006. — 176 с.

8. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

1. Кластерный анализ. StatSoft / [Электронный ресурс]. — URL: <http://statsoft.ru/home/textbook/modules/stcluan.html#k>.

2. Дюк В., Самойленко А. Data Mining: учебный курс (+CD). – СПб.: Изд. Питер, 2001. – 368 с.

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

9. РЕСУРСЫ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫЕ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

1. Сайт дистанционного обучения в НОУ (Национальный Открытый Университет) «ИНТУИТ» содержит бесплатные курсы, программы повышения квалификации и профессиональной переподготовки, интересные доклады и другую полезную информацию <http://www.intuit.ru>.

2. Федеральный портал «Российское образование» <http://www.edu.ru/>

3. Информационный сайт, содержащий справочные материалы по информатике, которые включают в себя курс лекций, схемы, презентации, рефераты и др. informatikaplus.narod.ru.

4. Материалы сайта об интеллектуальном анализе данных: <http://www.rdatamining.com/>

5. База данных временных рядов <http://irafm.osu.cz/cif/main.php>

6. Сайт кафедры ПОАС КГУ «Информатика и программирование: шаг за шагом» <http://it.kgsu.ru/>.

10. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ И ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ

1. ЭБС «Лань»
2. ЭБС «Консультант студента»
3. ЭБС «Znaniyum.com»
4. Гарант – справочно-правовая система

11. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Материально-техническое обеспечение по реализации дисциплины осуществляется в соответствии с требованиями ФГОС ВО по данной образовательной программе.

12. ДЛЯ ОБУЧАЮЩИХСЯ С ИСПОЛЬЗОВАНИЕМ ДИСТАНЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

При использовании электронного обучения и дистанционных образовательных технологий (далее ЭО и ДОТ) занятия полностью или частично проводятся в режиме онлайн. Объем дисциплины и распределение нагрузки по видам работ соответствует п. 4.1. Распределение баллов соответствует п. 6.2 либо может быть изменено в соответствии с решением кафедры, в случае перехода на ЭО и ДОТ в процессе обучения. Решение кафедры об используемых технологиях и системе оценивания достижений обучающихся принимается с учетом мнения ведущего преподавателя и доводится до обучающихся.

Аннотация к рабочей программе дисциплины
«МЕТОДЫ И АЛГОРИТМЫ АНАЛИЗА ДАННЫХ»

образовательной программы высшего образования –
программы бакалавриата

35.03.06 - Агроинженерия

Направленность:
Информационные технологии в электроэнергетике

Трудоемкость дисциплины: 3 з.е. (108 академических часов)

Семестр: 5

Форма промежуточной аттестации: зачет

Содержание дисциплины

Понятие Data Mining. Задачи и практическое применение Data Mining. Модели и методы Data Mining.

Классификация и регрессия. Поиск ассоциативных правил. Кластеризация. Базовые алгоритмы кластеризации.

Визуальный анализ данных (Visual Mining). Методы визуализации: методы геометрических преобразований, отображение иконок, методы, ориентированные на пикселы, иерархические образы.

Анализ текстовой информации (Text Mining). Этапы анализа текстов. Извлечение ключевых понятий из текста. Классификация текстовых документов.

Стандарты Data Mining.

Data Mining в реальном времени.

Извлечение знаний из Web. Этапы Web Mining. Категории Web Mining. Методы извлечения Web-контента. Извлечение Web-структур. Поиск Web-документов с учетом гиперссылок.

ЛИСТ
регистрации изменений (дополнений) в рабочую программу
учебной дисциплины
«МЕТОДЫ И АЛГОРИТМЫ АНАЛИЗА ДАННЫХ»

Изменения / дополнения в рабочую программу
на 20__ / 20__ учебный год:

Ответственный преподаватель _____ / _____ /

Изменения утверждены на заседании кафедры «__» _____ 20__ г.,
Протокол № ____

Заведующий кафедрой _____ «__» _____ 20__ г.

Изменения / дополнения в рабочую программу
на 20__ / 20__ учебный год:

Ответственный преподаватель _____ / _____ /

Изменения утверждены на заседании кафедры «__» _____ 20__ г.,
Протокол № ____

Заведующий кафедрой _____ «__» _____ 20__ г.