МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Воронежский государственный технический университет»

УТВЕРЖДАЮ
Декан факультета информационных технологий и компьютерной безопасности
_____ / П.Ю. Гусев /
21 декабря 2021 г.

РАБОЧАЯ ПРОГРАММА
дисциплины (модуля)
«Интеллектуальный анализ данных»

Направление подготовки 09.04.01 Информатика и вычислительная техника

Профиль (специализация) Управление программным инжинирингом

Квалификация выпускника магистр

Нормативный период обучения 2 года

Форма обучения Очная

Год начала подготовки 2022 г.

Автор(ы) программы _____ Т.Н. Недикова

Заведующий кафедрой автоматизированных и вычислительных систем _____ В.Ф. Барабанов

Руководитель ОПОП _____ О.Я. Кравец

Воронеж 2021

1. ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ

1.1. Цели дисциплины

Целью изучения дисциплины является формирование у студентов представления о типах задач, возникающих в области интеллектуального анализа данных (Data Mining) и методах их решения, которые помогут студентам выявлять, формализовать и успешно решать практические задачи анализа данных, возникающие в процессе их профессиональной деятельности.

1.2. Задачи освоения дисциплины

Задачи освоения дисциплины следующие:

- приобретение знаний по методам и моделям Data Mining;
- изучение методов, основанных на нечеткой логике;
- получение представления о преимуществах и недостатках нейросетевых технологий;
- получение представления об алгоритмах построения деревьев решений;
- освоение методов кластеризации;
- получение практических навыков в работе с существующими программными пакетами по интеллектуальному анализу данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина «Интеллектуальный анализ данных» относится к дисциплинам части, формируемой участниками образовательных отношений (дисциплина по выбору) блока Б.1 учебного плана.

3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Процесс изучения дисциплины «Интеллектуальный анализ данных» направлен на формирование следующих компетенций:

ПК-4 - Способен разрабатывать программные системы с применением современных технологий и инструментальных средств

ПК-6 - Способен применять математический аппарат в программном инжиниринге

Компетенция	Результаты обучения, характеризующие сформированность компетенции
ПК-4	Знать основные методы интеллектуального анализа данных, основные методы решения задач машинного обучения
	Уметь формулировать задачи интеллектуального анализа данных, выбирать адекватные алгоритмы их решения, оценивать качество получаемых решений

	Владеть навыками построения и обучения искусственных нейронных сетей; навыками формирования систем нечеткого логического вывода
ПК-6	Знать основные задачи Data Mining и методы, применяемые для их решения; инструментальные интегрированные программные среды разработчиков для применения к интеллектуальному анализу данных
	Уметь применять при решении прикладных задач методы и алгоритмы поиска часто встречающихся шаблонов, классификации, кластеризации, восстановления регрессии
	Владеть технологиями разработки алгоритмов и программными системами интеллектуального анализа данных

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость дисциплины «Интеллектуальный анализ данных» составляет 4 зачетных единицы.

Распределение трудоемкости дисциплины по видам занятий

Очная форма обучения

Вид учебной работы	Всего часов	Семестры			
		2			
Аудиторные занятия (всего)	72	72			
В том числе:					
Лекции	36	36			
Практические занятия (ПЗ)					
Лабораторные работы (ЛР)	36	36			
Самостоятельная работа	72	72			
Курсовой проект (работа) (есть, нет)	есть	есть			
Контрольная работа (есть, нет)	нет	нет			
Вид промежуточной аттестации (зачет, зачет с оценкой, экзамен)	Зачет с оценкой	Зачет с оценкой			
Общая трудоемкость час	144	144			
зач. ед.	4	4			

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

5.1 Содержание разделов дисциплины и распределение трудоемкости по видам занятий

очная форма обучения

№ п/п	Наименование темы	Содержание раздела	Лекц	Прак зан.	Лаб. зан.	СРС	Всего, час
1	Задачи Data Mining.	Понятие Data Mining, возникновение, перспективы, проблемы Data mining. Взгляд на технологию Data Mining как на часть рынка информационных технологий. Классификация задач Data Mining. От данных к решениям. Сопоставление и сравнение понятий "информация", "данные", "знание". Сферы применения Data Mining	4			8	12
2	Данные. Методы и стадии Data Mining	Понятие данных. Объясняется значение понятий объект и атрибут, выборка, зависимая и независимая переменная. Подробно обсуждаются типы шкал. Приводятся различные типы наборов данных. Понятия базы данных и СУБД. Стадии Data Mining. Известные классификации методов Data Mining.	4			8	12
3	Классификация и кластеризация	Задача классификации Процесс классификации Методы, применяемые для решения задач классификации Точность классификации: оценка уровня ошибок. Оценивание классификационных методов. Задача кластеризации. Оценка качества кластеризации. Процесс кластеризации. Применение кластерного анализа. Кластерный анализ в маркетинговых исследованиях. Практика применения кластерного анализа в маркетинговых исследованиях.	4		6	8	18
4	Прогнозирование и визуализация	Задача прогнозирования Сравнение задач прогнозирования и классификации Прогнозирование и временные ряды Тренд, сезонность и цикл Точность прогноза Виды прогнозов Методы прогнозирования Задача визуализации	4		6	8	18
5	Методы классификации и прогнозирования. Ансамблирование.	Деревья решений. Метод опорных векторов. Метод "ближайшего соседа". Байесовская классификация. Нейронные сети. Построение ансамблей классификаторов: случайный лес, бустинг, бэггинг.	4		6	8	18
6	Методы кластерного анализа. Отбор значимых признаков.	Иерархические методы (агломеративные и дивизимные, дендрограмма). Итеративные методы (к-средних). Метрики расстояния между объектами и кластерами. Методы отбора признаков (метод главных компонент)	4		6	8	18
7	Методы поиска ассоциативных правил	Часто встречающиеся приложения с применением ассоциативных правил: Введение в ассоциативные правила. Часто встречающиеся шаблоны или образцы. Характеристики ассоциативных правил. Границы поддержки и достоверности ассоциативного правила. Методы поиска ассоциативных правил Разновидности алгоритма Apriori .	4		6	8	18
8	Способы визуального представления данных. Методы визуализации.	Визуализация инструментов Data Mining Визуализация Data Mining моделей. Методы визуализации. Качество визуализации.	4		6	8	18

9	Рынок инструментов Data Mining. Существующие программные продукты.	Инструменты Data Mining, их развитие, поставщики инструментов, классификация инструментов. Критерии сравнения и выбора инструмента Data Mining. Существующие программные продукты. Data Mining-услуги.	4			8	12
Итого			**36**		**36**	**72**	**144**

5.2 Перечень лабораторных работ

1. Классификация и кластеризация.

Анализ практических задач. Определение типа задачи (классификация, регрессия, кластеризация). Подготовка исходных данных.

2. Прогнозирование и визуализация Разработка модели МО для решения задач регрессии. Предварительный анализ исходных данных. Визуализация признаков. Базовый статистический анализ средствами Python.

3. Методы классификации и прогнозирования. Ансамблирование.Разработка моделей МО для решения задачи классификации: Деревья решений. Метод опорных векторов. Метод "ближайшего соседа". Байесовская классификация. Нейронные сети. Построение ансамблей классификаторов: случайный лес, бустинг, бэггинг.

4. Методы кластерного анализа. Отбор значимых признаков. Разработка моделей МО для решения задачи классификации: К-средних, агломеративная кластеризация, DBSCAN. Расчет критериев качества кластеризации. Использование метода главных компонент для снижения размерности признаков.

5. Методы поиска ассоциативных правил Разработка моделей МО для решения задачи поиска ассоциативных правил: алгоритм Apriori.

6. Способы визуального представления данных. Методы визуализации. Визуализация инструментов Data Mining. Методы визуализации.

6. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ ПРОЕКТОВ (РАБОТ) И КОНТРОЛЬНЫХ РАБОТ

В соответствии с учебным планом освоение дисциплины предусматривает выполнение курсового проекта во 2 семестре.

Примерная тематика курсового проекта (работы):

«Интеллектуальный анализ данных (по варианту)»

Задачи, решаемые при выполнении курсового проекта:

• провести анализ предметной области, поставить задачу, обеспечить достаточный объем данных для анализа (датасет), выполнить предварительную подготовку данных, используя соответствующие методы.

• определить тип задачи машинного обучения и выбрать необходимые виды моделей для ее решения; построить модели машинного обучения для решения поставленной задачи; осуществить настройку гиперпараметров.

• используя технологию разработки алгоритмов и программ, на одном из алгоритмических процедурных языков программирования высокого уровня провести реализацию;

• проанализировать качество получившихся моделей, проверить корректность и эффективность проектного решения;

• при составлении пояснительной записки руководствоваться основными стандартами Единой системы программной документации.

Курсовой проект включает в себя графическую часть и расчетно-пояснительную записку.

Учебным планом по дисциплине «Интеллектуальный анализ данных» не предусмотрено выполнение контрольной работы (контрольных работ) во 2 семестре.

7. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

7.1. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

7.1.1 Этап текущего контроля

Результаты текущего контроля знаний и межсессионной аттестации оцениваются по следующей системе:

«аттестован»;

«не аттестован».

Компетенция	Результаты обучения, характеризующие сформированность компетенции	Критерии оценивания	Аттестован	Не аттестован
ПК-4	Знать основные методы интеллектуального анализа данных, основные методы решения задач машинного обучения	Активная работа на занятиях, ответы на теоретические вопросы при защите лабораторных работ	Выполнение работ в срок, предусмотренный в рабочих программах	Невыполнение работ в срок, предусмотренный в рабочих программах
	Уметь формулировать задачи интеллектуального анализа данных, выбирать адекватные алгоритмы их решения, оценивать качество получаемых решений	Решение стандартных практических задач при выполнении лабораторных работ, отвечает на теоретические вопросы при защите лабораторных работ, написание курсового проекта	Выполнение работ в срок, предусмотренный в рабочих программах	Невыполнение работ в срок, предусмотренный в рабочих программах
	Владеть навыками построения и обучения искусственных нейронных сетей; навыками формирования систем нечеткого логического вывода	Решение прикладных задач в конкретной предметной области, выполнение плана работ по разработке курсового проекта	Выполнение работ в срок, предусмотренный в рабочих программах	Невыполнение работ в срок, предусмотренный в рабочих программах
ПК-6	Знать основные задачи Data Mining и методы, применяемые для их решения; инструментальные	Активная работа на занятиях, ответы на теоретические вопросы при защите лаборатор-	Выполнение работ в срок, предусмотренный в рабочих програм-	Невыполнение работ в срок, предусмотренный в рабочих про-

	интегрированные программные среды разработчиков для применения к интеллектуальному анализу данных	ных работ	мах	граммах
	Уметь применять при решении прикладных задач методы и алгоритмы поиска часто встречающихся шаблонов, классификации, кластеризации, восстановления регрессии	Решение стандартных практических задач при выполнении лабораторных работ, отвечает на теоретические вопросы при защите лабораторных работ, написание курсового проекта	Выполнение работ в срок, предусмотренный в рабочих программах	Невыполнение работ в срок, предусмотренный в рабочих программах
	Владеть технологиями разработки алгоритмов и программными системами интеллектуального анализа данных	Решение прикладных задач в конкретной предметной области, выполнение плана работ по разработке курсового проекта	Выполнение работ в срок, предусмотренный в рабочих программах	Невыполнение работ в срок, предусмотренный в рабочих программах

7.1.2 Этап промежуточного контроля знаний

Результаты промежуточного контроля знаний оцениваются в 2 семестре для очной формы обучения по системе:

«отлично»;

«хорошо»;

«удовлетворительно»;

«неудовлетворительно».

Компетенция	Результаты обучения, характеризующие сформированность компетенции	Критерии оценивания	Отлично	Хорошо	Удовл	Неудовл
ПК-4	Знать основные методы интеллектуального анализа данных, основные методы решения задач машинного обучения	Тест	Выполнение теста на 90-100%	Выполнение теста на 80-90%	Выполнение теста на 70-80%	В тесте менее 70% правильных ответов
	Уметь формулировать задачи интеллектуального анализа данных, выбирать адекватные алгоритмы их решения, оценивать качество получаемых решений	Решение стандартных практических задач	Задачи решены в полном объеме и получены верные ответы	Продемонстрирован верный ход решения, но не получен верные ответы во всех задачах	Продемонстрирован верный ход решения в большинстве задач	Задачи не решены
	Владеть навыками построения и обучения искусственных нейронных сетей; навыками формирования систем нечеткого логического вывода	Решение прикладных задач в конкретной предметной области	Задачи решены в полном объеме и получены верные ответы	Продемонстрирован верный ход решения, но не получен верные ответы во всех задачах	Продемонстрирован верный ход решения в большинстве задач	Задачи не решены
ПК-6	Знать основные задачи Data Mining и методы, применяемые для их решения; инструментальные интегрированные программные среды разработчиков для применения к интеллектуальному ана-	Тест	Выполнение теста на 90-100%	Выполнение теста на 80-90%	Выполнение теста на 70-80%	В тесте менее 70% правильных ответов

	лизу данных					
	Уметь применять при решении прикладных задач методы и алгоритмы поиска часто встречающихся шаблонов, классификации, кластеризации, восстановления регрессии	Решение стандартных практических задач	Задачи решены в полном объеме и получены верные ответы	Продемонстрирован верный ход решения, но не получен верные ответы во всех задачах	Продемонстрирован верный ход решения в большинстве задач	Задачи не решены
	Владеть технологиями разработки алгоритмов и программными системами интеллектуального анализа данных	Решение прикладных задач в конкретной предметной области	Задачи решены в полном объеме и получены верные ответы	Продемонстрирован верный ход решения, но не получен верные ответы во всех задачах	Продемонстрирован верный ход решения в большинстве задач	Задачи не решены

7.2 Примерный перечень оценочных средств (типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности)

7.2.1 Примерный перечень заданий для подготовки к тестированию

Задание #1
Data Mining — это процесс обнаружения в сырых данных
(1) ранее сформулированных гипотез
+(2) неочевидных закономерностей
+(3) практических закономерностей
+(4) объективных закономерностей
(5) большого количества закономерностей

Задание #2
В приведенном списке отметьте задачи интеллектуального анализа данных, относящиеся к классу "обучение с учителем":
+(1) классификация
(2) кластеризация
+(3) регрессия
(4) поиск ассоциативных правил

Задание #3
В приведенном списке отметьте задачи интеллектуального анализа данных, относящиеся к классу "обучение без учителя":
(1) классификация
(2) регрессия
+(3) кластеризация
+(4) поиск ассоциативных правил

Задание #4
В приведенном списке выберите корректное утверждение относительно структур и моделей интеллектуального анализа
+(1) структура содержит модели, которые используются для анализа ее данных
(2) модель содержит структуры, которые используются для анализа ее данных
(3) модели и структуры – это несвязанные типы объектов, ни один из них не может включать (содержать) другой

Задание #5
Выберите корректное продолжение фразы: "В структуре интеллектуального анализа данных можно ..."
+(1) выделить обучающий и проверочный набор данных
(2) указать алгоритм интеллектуального анализа
(3) определить фильтры для данных

Задание #6
Процедуру "очистки" исходных данных рекомендуется производить:
+(1) до создания модели интеллектуального анализа
(2) после создания и обработки (обучения) модели интеллектуального анализа
(3) после проведения тестирования (оценки точности) модели интеллектуального анализа

Задание #7
Ниже перечислены рекомендуемые этапы проведения интеллектуального анализа данных Расставьте их в правильной последовательности. Ответ укажите в виде последовательности чисел, например 654321
(1) изучение данных
(2) подготовка данных
(3) постановка задачи
(4) развертывание и обновление моделей
(5) построение моделей
(6) исследование и проверка моделей
Ответ: 321564

Задание #8
Выберите наиболее корректное продолжение фразы "При проведении интеллектуального анализа данных на этапе изучения данных ..."
+(1) может проводиться поиск минимальных и максимальных значений параметров, анализ статистических характеристик, сравнение полученных результатов с представлениями о предметной области

(2) может проводиться удаление "выбросов", обработка отсутствующих значений параметров, численное преобразование (например, нормализация)
(3) может определяться, нужно ли будет делать прогнозы на основании модели интеллектуального анализа данных или просто найти содержательные закономерности и взаимосвязи

Задание #9
Выберите наиболее точное продолжение фразы: "Перекрёстная проверка может использоваться для оценки точности модели ..."
(1) только в случае решения предсказательных задач интеллектуального анализа
(2) только в случае решения описательных задач интеллектуального анализа
+(3) при решении как предсказательных, так и описательных задач интеллектуального анализа

Задание #10
Среди перечисленных средств оценки точности, выберите те, что плохо применимы в случае решения описательных задач интеллектуального анализа:
+(1) диаграмма точности прогнозов
+(2) матрица классификации
(3) перекрестная проверка

7.2.2 Примерный перечень заданий для решения стандартных задач

Задание #1
Выберите корректное утверждение относительно упрощенного алгоритма Байеса:
(1) предполагается, что все входные атрибуты взаимосвязаны
+(2) предполагается, что все входные атрибуты взаимно независимы
(3) никаких предположений относительно степени взаимной зависимости входных атрибутов не делается

Задание #2
Среди представленных в списке алгоритмов интеллектуального анализа выберите наиболее быстрый (требующий наименьшего количества вычислений):
(1) Microsoft Clustering
+(2) Microsoft Naive Bayes
(3) Microsoft Neural Network

Задание #3

Деревья решений позволяют решать задачи (отметьте все подходящие варианты):
+(1) классификации
(2) кластеризации
+(3) регрессии
+(4) поиска взаимосвязей

Задание #4
Явление, заключающееся в том, что модель отлично классифицирует данные из обучающего набора, но имеет невысокую точность прогнозирования для новых данных, принято называть:
+(1) переобучение
(2) недостаточное обучение
(3) избыточность
(4) зацикленность

Задание #5
В алгоритмах регрессии "регрессором" называется:
+(1) независимая переменная (аргумент)
(2) зависимая переменная
(3) выявленная алгоритмом зависимость

Задание #6
При анализе временных рядов, как правило, предполагается, что предсказываемый атрибут:
(1) текстовый
+(2) числовой непрерывный
(3) числовой дискретный

Задание #7
В приведенном списке отметьте два основных формата представления временных рядов, применяемых в ходе интеллектуального анализа данных:
+(1) столбчатый
+(2) чередующийся
(3) непрерывный
(4) перекрестный

Задание #8
Алгоритм "к-средних" (c-means) относится к алгоритмам ...
(1) классификации
+(2) "жесткой" кластеризации
(3) "мягкой" кластеризации
(4) поиска взаимосвязей

Задание #9
Термин "жесткая кластеризация" означает что:
+(1) каждый вариант будет отнесен только к одному кластеру
(2) каждый вариант будет отнесен к одному или нескольким кластерам с разными вероятностями
(3) каждый кластер будет содержать ровно один вариант

Задание #10
Определение методами интеллектуального анализа значения непрерывного числового параметра на основании значений независимых переменных производится при решении задачи:
(1) классификации
(2) кластеризации
+(3) регрессии

7.2.3 Примерный перечень заданий для решения прикладных задач

Задание #1
Постановка задачи, построение оптимальной модели, понимание модели, применение результатов. Перечисленные выше этапы являются этапами:
(1) традиционного процесса Data Mining
+(2) подхода KXEN
(3) и того, и другого
(4) ни того, ни другого

Задание #2
Проект в PolyAnalyst объединяет в себе:
(1) только дерево проекта
(2) только графики
(3) только правила
+(4) все объекты исследования

Задание #3
Пакет SAS Enterprise Miner особенно удобен для осуществления анализа данных в …
+(1) масштабах крупных организаций
(2) масштабах средних организаций
(3) масштабах средних и небольших организаций

Задание #4
Существуют следующие варианты решений по внедрению инструментов Data Mining:
+(1) покупка готового программного обеспечения Data Mining

+(2) покупка программного обеспечения Data Mining, адаптированного под конкретный бизнес
+(3) комбинация этих вариантов, в т.ч. использование различных библиотек, компонентов и инструментальные наборы для разработчиков создания встроенных приложений Data Mining

Задание #5
На рынке инструментов Data Mining в последние годы наблюдается:
(1) спад
+(2) рост
+(3) ситуация на рынке за последние годы почти не меняется

Задание #6
Построение моделей Data Mining осуществляется с целью:
+(1) исследования или изучения моделируемого объекта и получения новых знаний, необходимых для принятия решений
(2) выбора наиболее быстродействующей модели
(3) исследования всех возможных свойств и характеристик изучаемого объекта

Задание #7
Прогнозирующие модели Data Mining ...
+(1) позволяют на основе выявленных закономерностей предсказывать будущее поведение объекта
(2) описывают общие закономерности предметной области
(3) решают задачи кластеризации, группировки, обобщения
(4) все ответы неверны

Задание #8
Классификационные модели Data Mining ...
+(1) осуществляют прогнозирование класса объекта
(2) описывают общие закономерности предметной области
(3) решают задачи кластеризации, группировки, обобщения
(4) все ответы верны

Задание #9
Создание каких моделей Data Mining означает поиск правил, которые объясняют зависимость выходных параметров от входных?
+(1) моделей классификации и прогнозирования
(2) моделей кластеризации и классификации
(3) моделей правил ассоциаций

Задание #10
Экзогенные переменные — это переменные, которые ...
+(1) задаются вне модели, они известны заранее

(2) определяются по ходу расчетов в модели, они не задаются извне
(3) задаются внутри модели, они известны заранее

7.2.4 Примерный перечень вопросов для подготовки к зачету

1. Классификация задач Data Mining. Сферы применения Data Mining.

2. Задача классификации. Процесс классификации. Методы, применяемые для решения задач классификации. Точность классификации: оценка уровня ошибок

3. Задача кластеризации Оценка качества кластеризации Процесс кластеризации 4. Применение кластерного анализа Кластерный анализ в маркетинговых исследованиях Практика применения кластерного анализа в маркетинговых исследованиях

5. Задача прогнозирования Сравнение задач прогнозирования и классификации 6. 6. Прогнозирование и временные ряды Тренд, сезонность и цикл

7. Точность прогноза Виды прогнозов Методы прогнозирования

8. Деревья решений.

9. Метод опорных векторов.

10. Метод "ближайшего соседа".

11.Байесовская классификация.

12. Классификация с помощью Нейронных сетей

13. Иерархические методы кластеризации.

14. Итеративные методы кластеризации.

15. Ассоциативные правила. Часто встречающиеся шаблоны или образцы. Поддержка. Характеристики ассоциативных правил. Границы поддержки и достоверности ассоциативного правила.

16. Методы поиска ассоциативных правил Разновидности алгоритма Apriori .

17. Методы визуализации Представление данных в одном, двух и трех измерениях

18. Параллельные координаты. "Лица Чернова". Качество визуализации. Представление пространственных характеристик.

7.2.5 Примерный перечень вопросов для подготовки к экзамену

Не предусмотрено учебным планом

7.2.6 Методика выставления оценки при проведении промежуточной аттестации

Зачет проводится по билетам, включающим по два вопроса. Допуском к зачету является выполнение всех лабораторных работ и положительное текущее тестирование.

Зачет ставится, если студент выполнил все лабораторные работы, прошел тестирование по темам теоретического материала и ответил на один или два вопроса.

Зачет не ставится, если студент не выполнил лабораторные работы и не ответил ни на один вопрос на зачете.

7.2.7 Паспорт оценочных материалов

№ п/п	Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
1	Задачи Data Mining.	ПК-4, ПК-6	Тест, защита лабораторных работ
2	Данные. Методы и стадии Data Mining	ПК-4, ПК-6	Тест, защита лабораторных работ
3	Классификация и кластеризация	ПК-4, ПК-6	Тест, защита лабораторных работ
4	Прогнозирование и визуализация	ПК-4, ПК-6	Тест, защита лабораторных работ
5	Методы классификации и прогнозирования. Ансамблирование.	ПК-4, ПК-6	Тест, защита лабораторных работ
6	Методы кластерного анализа. Отбор значимых признаков.	ПК-4, ПК-6	Тест, защита лабораторных работ
7	Методы поиска ассоциативных правил	ПК-4, ПК-6	Тест, защита лабораторных работ
8	Способы визуального представления данных. Методы визуализации.	ПК-4, ПК-6	Тест, защита лабораторных работ
9	Рынок инструментов Data Mining. Существующие программные продукты.	ПК-4, ПК-6	Тест, защита лабораторных работ

7.3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Тестирование осуществляется, либо при помощи компьютерной системы тестирования, либо с использованием выданных тест-заданий на бумажном носителе. Время тестирования 30 мин. Затем осуществляется проверка теста экзаменатором и выставляется оценка согласно методике выставления оценки при проведении промежуточной аттестации.

Решение стандартных задач осуществляется, либо при помощи компьютерной системы тестирования, либо с использованием выданных задач на бумажном носителе. Время решения задач 30 мин. Затем осуществляется проверка решения задач экзаменатором и выставляется оценка, согласно методике выставления оценки при проведении промежуточной аттестации.

Решение прикладных задач осуществляется, либо при помощи компьютерной системы тестирования, либо с использованием выданных задач на бумажном носителе. Время решения задач 30 мин. Затем осуществляется провер-

ка решения задач экзаменатором и выставляется оценка, согласно методике выставления оценки при проведении промежуточной аттестации.

Защита курсового проекта осуществляется согласно требованиям, предъявляемым к работе, описанным в методических материалах. Примерное время защиты на одного студента составляет 20 мин.

8 УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

8.1 Перечень учебной литературы, необходимой для освоения дисциплины

1 Нестеров С.А. Интеллектуальный анализ данных средствами MS SQL Server 2008 / Нестеров С.А.. — Москва: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 303 c. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: https://www.iprbookshop.ru/62813.html

2. Пальмов С.В. Интеллектуальный анализ данных: учебное пособие / Пальмов С.В.. — Самара : Поволжский государственный университет телекоммуникаций и информатики, 2017. — 127 c. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: https://www.iprbookshop.ru/75376.html

3. Афанасьева Т.В. Введение в проектирование систем интеллектуального анализа данных: учебное пособие / Афанасьева Т.В., Афанасьев А.Н.. — Ульяновск : Ульяновский государственный технический университет, 2017. — 64 c. — ISBN 978-5-9795-1686-8. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: https://www.iprbookshop.ru/106086.html

4. Замятин А.В. Введение в интеллектуальный анализ данных: учебное пособие / Замятин А.В.. — Томск : Издательский Дом Томского государственного университета, 2016. — 119 c. — ISBN 978-5-94621-531-2. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: https://www.iprbookshop.ru/109021.html

5. Мельниченко А.С. Математическая статистика и анализ данных: учебное пособие / Мельниченко А.С.. — Москва : Издательский Дом МИСиС, 2018. — 45 c. — ISBN 978-5-906953-62-9. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: https://www.iprbookshop.ru/78563.html

6. Организация самостоятельной работы обучающихся: методические указания для студентов, осваивающих основные образовательные программы высшего образования – бакалавриата, специалитета, магистратуры: методические указания / сост. В.Н. Почечихина, И.Н. Крючкова, Е.И. Головина, В.Р. Демидов; ФГБОУ ВО «Воронежский государственный технический университет ». – Воронеж, 2020. – 14 с.

7. Методические рекомендации по выполнению курсовых проектов (работ) по программам высшего образования – программам бакалавриата,

специалитета, магистратуры / ФГБОУ ВО «Воронежский государственный технический университет»; сост. В.Н. Почечихина, И.Н. Крючкова, Е.И. Головина. Воронеж: Изд-во ВГТУ, 2020. 10 с.

8.2 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень лицензионного программного обеспечения, ресурсов информационно-телекоммуникационной сети «Интернет», современных профессиональных баз данных и информационных справочных систем

Лицензионное ПО:

- Windows Professional 7 Single Upgrade MVL A Each Academic
- Microsoft Office Word 2007
- Microsoft Office Power Point 2007

Свободно распространяемое ПО:

- Microsoft Visual Studio Community Edition
- Microsoft SQL Server Express
- Microsoft SQL Server Managment Studio
- СУБД MS SQL Server 2012

Отечественное ПО:

- Яндекс.Браузер
- Архиватор 7z
- Astra Linux

Ресурс информационно-телекоммуникационной сети «Интернет»:

- Образовательный портал ВГТУ
- http://www.edu.ru/
- https://metanit.com/

Информационно-справочные системы:

- http://window.edu.ru
- https://wiki.cchgeu.ru/

Современные профессиональные базы данных:

- https://proglib.io
- https://msdn.microsoft.com/ru-ru/
- https://docs.microsoft.com/

Информационные технологии, используемые при осуществлении образовательного процесса по дисциплине:

- лекции с применением мультимедийных средств;
- обучение прикладным информационным технологиям, ориентированным на специальность, в рамках лабораторных работ с применением лицензионного программного обеспечения.

9 МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

Для проведения лекционных занятий необходима аудитория, оснащенная оборудованием для лекционных демонстраций и проекционной аппаратурой.

Для проведения лабораторных работ необходима лаборатория с ПК, оснащенными программами для проведения лабораторного практикума и обеспечивающими возможность доступа к локальной сети кафедры и Интернет, из следующего перечня:

- 311 (Лаборатория разработки программных систем)
- 320 (Лаборатория общего назначения)
- 322 (Лаборатория распределённых вычислений)
- 324 (Специализированная лаборатория сетевых систем управления (научно-образовательный центр «АТОС»))
- 325 (Лаборатория автоматизации проектирования вычислительных комплексов и сетей)

Лаборатории расположены по адресу: 394066, г. Воронеж, Московский проспект, 179 (учебный корпус №3).

10 МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

По дисциплине «Интеллектуальный анализ данных» читаются лекции, проводятся лабораторные занятия, выполняется курсовой проект.

Основой изучения дисциплины являются лекции, на которых излагаются наиболее существенные и трудные вопросы, а также вопросы, не нашедшие отражения в учебной литературе.

Лабораторные работы выполняются на лабораторном оборудовании в соответствии с методиками, приведенными в указаниях к выполнению работ.

Большое значение по закреплению и совершенствованию знаний имеет самостоятельная работа студентов. Информацию обо всех видах самостоятельной работы студенты получают на занятиях.

Методика выполнения курсового проекта изложена в учебно-методическом пособии. Выполнять этапы курсового проекта должны своевременно и в установленные сроки.

Контроль усвоения материала дисциплины производится проверкой курсового проекта, защитой курсового проекта, защитой лабораторных работ и на зачете при ответе на вопросы.

Вид учебных занятий	Деятельность студента (*особенности деятельности студента инвалида и лица с ОВЗ, при наличии таких обучающихся*)
Лекция	Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения, выводы, формулировки, обобщения; помечать важные мысли, выделять ключевые слова, термины. Проверка терминов, понятий с помощью энциклопедий, словарей, справочников с выписыванием толкований в тетрадь. Обозначение вопросов, терминов, материала, которые вызывают трудности, поиск ответов в рекомендуемой литературе. Если самостоятельно не удается разо-

	браться в материале, необходимо сформулировать вопрос и задать преподавателю на лекции или на практическом занятии.
Лабораторные занятия	Лабораторные работы позволяют научиться применять теоретические знания, полученные на лекции при решении конкретных задач. Чтобы наиболее рационально и полно использовать все возможности лабораторных занятий для подготовки к ним необходимо: разобрать лекцию по соответствующей теме, ознакомится с соответствующим разделом учебного пособия, проработать дополнительную литературу и источники, изучить методическое обеспечение лабораторной работы.
Самостоятельная работа	Самостоятельная работа студентов способствует глубокому усвоения учебного материала и развитию навыков самообразования. Самостоятельная работа предполагает следующие составляющие: - работа с текстами: учебниками, справочниками, дополнительной литературой, а также проработка конспектов лекций; - работа над темами для самостоятельного изучения; - участие в работе студенческих научных конференций, олимпиад; - выполнение курсового проекта; - оформление расчетно-пояснительной записки; - подготовка к лабораторным занятиям; - подготовка к промежуточной аттестации.
Подготовка к дифференцированному зачету	При подготовке к зачету необходимо ориентироваться на конспекты лекций, рекомендуемую литературу и решение индивидуальных заданий на лабораторных занятиях.

ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ

№ п/п	Перечень вносимых изменений	Дата внесения изменений	Подпись заведующего кафедрой, ответственной за реализацию ОПОП