

ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»

ФИЗИКО-ТЕХНИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра компьютерных технологий



УТВЕРЖДАЮ:

Проректор по научно-методической
и учебной работе

Е.И. Скафа

22 апреля 2020 г.

Рабочая программа учебной дисциплины

«ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ»

Направления подготовки:	09.04.01 Информатика и вычислительная техника
Магистерская программа:	Информатика и вычислительная техника
Образовательная программа:	академическая магистратура
Квалификация:	магистр
Форма обучения:	<u>очная</u> , очно-заочная, <u>заочная</u>

Донецк 2020

УТВЕРЖДАЮ:

Декан физико-технического факультета

С.А. Фоменко

«17» апреля 2020 г.

МП №2



Программа учебной дисциплины «Интеллектуальный анализ данных» составлена на основании Федерального государственного образовательного стандарта высшего образования (ФГОС ВО) направления подготовки 09.04.01 Информатика и вычислительная техника, утвержденного приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г. № 918;

Порядка организации учебного процесса в образовательных организациях высшего профессионального образования Донецкой Народной Республики, утвержденного приказом Министерства образования и науки ДНР № 1171 от «10» ноября 2017 г.; учебного плана и основной образовательной программы Информатика и вычислительная техника, направления подготовки 09.04.01 Информатика и вычислительная техника, разработанных в ГОУ ВПО «Донецкий национальный университет».

Разработчик:

к.т.н, доцент кафедры компьютерных технологий

Бондаренко В.И.

Программа учебной дисциплины утверждена на заседании кафедры компьютерных технологий

Протокол № 12 от «2» апреля 2020 г.

Зав. кафедрой компьютерных технологий

Ермоленко Т.В.

Программа учебной дисциплины одобрена учебно-методической комиссией физико-технического факультета

Протокол № 5 от «15» апреля 2020 г.

Председатель учебно-методической комиссии физико-технического факультета

Котенко В.Н.

1. ОБЛАСТЬ ПРИМЕНЕНИЯ И МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ

Учебная дисциплина «Интеллектуальный анализ данных» относится к вариативной части профессионального блока и состоит из двух содержательных модулей.

Основывается на базе дисциплин: «Математическая логика», «Базы данных», «Технологии извлечения знаний», «Методы оптимизации и идентификации систем», «Программирование».

2. СТРУКТУРА ДИСЦИПЛИНЫ

<i>Характеристика учебной дисциплины</i>		
Направление подготовки	09.04.01 Информатика и вычислительная техника	
Профиль	Информатика и вычислительная техника	
Образовательная программа	Академическая магистратура	
Квалификация	Магистр	
Количество содержательных модулей	2	
Дисциплина базовой / вариативной части образовательной программы	Профессиональный блок. Вариативная часть	
Формы контроля (МК, экзамен, зачет)	Модульный контроль, экзамен	
Показатели	очная форма обучения	заочная форма обучения
Количество зачётных единиц (кредитов)	3,5	3,5
Год подготовки	2	2
Семестр	3	3
Количество часов	126	126
- лекционных	14	2
- практических, семинарских		
- лабораторных	28	6
- самостоятельной работы	84	118
в т. ч. индивидуальное задание		
Недельное количество часов, т. ч.	7,88	
аудиторных	3	

3. ОПИСАНИЕ ДИСЦИПЛИНЫ

Цели и задачи.

Целью изучения дисциплины «Интеллектуальный анализ данных» является формирование у студентов представления о типах задач, возникающих в области интеллектуального анализа данных (Data Mining); изучение основных подходов и алгоритмов решения задач анализа данных и особенностей их применения к решению реальных задач; получение студентами навыка по выявлению, формализации и успешному решению практических задач анализа данных; получение практического навыка в работе с существующими программными пакетами по анализу данных.

Требования к результатам освоения дисциплины.

Процесс изучения дисциплины «Интеллектуальный анализ данных» направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО РФ направления подготовки 09.04.01 Информатика и вычислительная техника и основной образовательной программы высшего профессионального образования направления

подготовки 09.04.01 Информатика и вычислительная техника (магистерская программа: Информатика и вычислительная техника):

а) универсальных (УК):

- способностью осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий (УК-1);
- способностью организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели (УК-3);
- способностью применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия (УК-4);
- способностью анализировать и учитывать разнообразие культур в процессе межкультурного взаимодействия (УК-5);
- способностью определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки (УК-6).

б) общепрофессиональных (ОПК):

- способностью самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте (ОПК-1);
- способностью разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач (ОПК-2);
- способностью анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями (ОПК-3);
- способностью применять на практике новые научные принципы и методы исследований (ОПК-4);
- способностью разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем (ОПК-5);
- способностью адаптировать зарубежные комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий (ОПК-7);
- способностью осуществлять эффективное управление разработкой программных средств и проектов (ОПК-8).

в) профессиональных (ПК):

производственно-технологическая деятельность:

- способностью управлять развитием баз данных (ПК-1);
- способностью осуществлять управление сервисами информационных технологий (ПК-2);
- способностью осуществлять технологическую поддержку подготовки технических публикаций (ПК-3);
- способностью осуществлять управление развитием инфокоммуникационной системы организации (ПК-6);
- способностью осуществлять научно-методическое и учебно-методическое обеспечение реализации программ профессионального обучения, среднего профессионального образования и дополнительного профессионального образования (ПК-9);

проектная деятельность:

- способностью проектировать сложные пользовательские интерфейсы (ПК-10);

организационно-управленческая деятельность:

- способностью управлять работами по сопровождению и проектами по созданию

(модификации) информационных систем, автоматизирующих задачи организационного управления и бизнес-процессы (ПК-13);

– способностью осуществлять руководство разработкой комплексных проектов на всех стадиях и этапах выполнения работ (ПК-20);

научно-исследовательская деятельность:

– способностью осуществлять экспертный анализ эргономических характеристик программных продуктов и/или аппаратных средств (ПК-21).

В результате изучения учебной дисциплины студент должен:

Знать:

- приёмы структурирования входящей информации;
- основные подходы к анализу данных и способы их модификации;
- математические модели, лежащие в основе различных подходов к решению задач анализа данных;
- структуру и содержимое программных пакетов для анализа данных (например sklearn в python, R);
- основные подходы для решения задач интеллектуального анализа данных.

Уметь:

- анализировать предметную область, делать обобщения и синтезировать знания о ней;
- настраивать алгоритмы под особенности прикладных задач;
- обосновывать оптимальность алгоритмов в рамках моделей;
- уметь работать с пакетами анализа и решать конкретные прикладные задачи;
- выбрать модель, наиболее адекватную решаемой задаче и обосновать ее эффективность.

Владеть:

- программными инструментами, позволяющими гибкую подстройку алгоритма;
- основными методами обработки информации.

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ И ФОРМЫ ОРГАНИЗАЦИИ УЧЕБНОГО ПРОЦЕССА

Порядковый номер и тема	Краткое содержание темы
	<i>Содержательный модуль 1.</i>
Тема 1. Обзор состояния проблемы искусственного интеллекта.	История развития. Основные проблемы, сформулированные для систем искусственного интеллекта. Определения искусственного интеллекта. Задачи интеллектуального анализа данных
Тема 2. Методы предварительной обработки данных.	Задача анализа данных, как задача поиска закономерностей в протоколах наблюдений. Определение и свойства эмпирических гипотез. Эмпирические закономерности их подтвержденность и потенциальная опровержимость. Алгоритм усиления гипотез. Теория измерений. Типы измерительных шкал. Допустимые преобразования для различных типов шкал. Нормировка шкал, вычисление расстояний. Разнотипные признаки. Вычисление расстояний в пространстве разнотипных признаков. Вычисление расстояний между признаками в зависимости от их типа.

Тема 3. Классификация задач анализа данных. Базовые гипотезы.	Классификация задач анализа данных, как задач заполнения пробелов в таблице объект-свойство в зависимости от типов шкал; местоположения и количества пробелов. Гипотеза компактности, как одна из основных базовых гипотез в анализе данных.
Тема 4. Задачи классификации, кластеризации и частичного обучения.	Задача классификации и основные подходы к ее решению. Задача кластеризации и основные подходы к ее решению. Задача частичного обучения (semi-supervised learning) как промежуточный вариант между задачей классификации и кластеризации, эффективность постановки в условиях непредставительности обучающей выборки.
Содержательный модуль 2.	
Тема 5. Задача выбора и генерации информативной системы признаков.	Основные подходы к решению задачи выбора информативной системы признаков (feature selection). Критерии качества. Алгоритмы перебора признаковов подсистем (градиентные и стохастические). Основные подходы к решению задачи генерации информативной системы признаков (feature extraction). Метод главных компонент. Нелинейные преобразования признаков.
Тема 6. Задача прогнозирования и заполнения пробелов.	Задача прогнозирования (регрессионного анализа), связь геометрической и статистической постановок. Построение линейной регрессии. Влияние вида функции потерь при восстановлении регрессионной зависимости. Заполнение пробелов в таблицах данных и поиск ошибок в данных. Алгоритм ZET. Группа алгоритмов WANGA для разнотипных переменных.
Тема 7. Задачи комбинированного типа.	Переход от основных задач анализа данных к задачам комбинированных. Особенности выбора признаков для задач классификации. Принципы натуральной классификации. Задача SDX-наиболее общая задача анализа данных
Тема 8. Функция конкурентного сходства и основанные на ней алгоритмы.	Конкурентное сходство как базовая гипотеза для решения различных задач анализа данных. Определения функции конкурентного сходства. Измерение компактности с помощью функции конкурентного сходства. Алгоритмы, основанные на функции конкурентного сходства
Тема 9. Анализ временных рядов	Основные подходы, применяемые при анализе временных рядов. Выделение трендов и сезонных составляющих. Разложение в ряд Фурье для периодических процессов. Марковские модели.

Курс дисциплины «Интеллектуальный анализ данных» предусматривает следующие **формы организации учебного процесса:**

1. лекции;
2. лабораторные занятия;
3. самостоятельная работа студента.

Ссылки на электронные материалы по всем формам организации учебного процесса размещены на сайте <http://donnu.ru/phys/kt/bondarenko>.

По источнику передачи и восприятия учебной информации используются словесные (лекция, беседа), наглядные (иллюстрация, демонстрация), практические (исследования, упражнения, лабораторные работы) методы.

По характеру познавательной деятельности студентов используются объяснительно-иллюстративные и репродуктивные методы, проблемное преподавание, частично-поисковый и исследовательский методы.

В зависимости от основной дидактической цели и задач используются методы устного изложения знаний, закрепление учебного материала, самостоятельной работы студентов по осмыслению и усвоению нового материала, работы по применению знаний на практике и выработке умений и навыков, проверки и оценки знаний, умений и навыков.

Используются следующие методы контроля:

1. устный контроль (экспресс-опрос на лекциях);
2. проверка конспектов;
3. защита лабораторных работ;
4. проверка самостоятельных работ;
5. модульная контрольная работа (дидактическое тестирование);
6. экзамен.

Тематический план

	Содержательный модуль 1											
Названия содержательных модулей и тем	Количество часов											
	Очная форма обучения						Заочная форма обучения					
	всего	В Т.Ч.					всего	В Т.Ч.				
		лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа		лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа
Тема 1. Обзор состояния проблемы искусственного интеллекта	9	1		2	6		9	0,13		0,5	8,37	0,1
Тема 2. Методы предварительной обработки данных.	16	2		4	10		16	0,26		0,5	15,24	0,2
Тема 3. Классификация задач анализа данных. Базовые гипотезы.	11	1		2	8		11	0,13		0,5	10,37	0,2
Тема 4. Задачи классификации, кластеризации и частичного обучения.	18	2		4	12		18	0,26		0,5	17,24	0,2
Итого по 1-му содержательному модулю	54	6		12	36		54	0,78		2	51,22	0,8

	Содержательный модуль 2											
Названия содержательных модулей и тем	Количество часов											
	Очная форма обучения						Заочная форма обучения					
	всего	В Т.Ч.					всего	В Т.Ч.				
		лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа		лекции	практические	лабораторные	самостоятельная работа	индивидуальная работа
Тема 5. Задача выбора и генерации информативной системы признаков.	13	1		2	10		13	0,05		0,5	12,45	
Тема 6. Задача прогнозирования и заполнения пробелов.	18	2		4	12		18	0,26		1	16,74	
Тема 7. Задачи комбинированного типа.	14	2		4	8		14	0,26		1	12,74	
Тема 8. Функция конкурентного сходства и основанные на ней алгоритмы.	14	2		4	8		14	0,26		1	12,74	
Тема 9. Анализ временных рядов	13	1		2	10		13	0,39		0,5	12,11	
Итого по 2-му содержательному модулю	72	8		16	48		72	1,22		4	66,781	
Всего часов	126	14		28	84		126	2		6	118	

5. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ПРОВЕДЕНИЯ ЛЕКЦИОННЫХ, ПРАКТИЧЕСКИХ И ЛАБОРАТОРНЫХ ЗАНЯТИЙ

Темы лекционных занятий

№ п/п	Название темы	Количество часов
1.	Обзор состояния проблемы искусственного интеллекта	1
2.	Методы предварительной обработки данных.	2

3.	Классификация задач анализа данных. Базовые гипотезы.	1
4.	Задачи классификации, кластеризации и частичного обучения.	2
5.	Задача выбора и генерации информативной системы признаков.	1
6.	Задача прогнозирования и заполнения пробелов.	2
7.	Задачи комбинированного типа.	2
8.	Функция конкурентного сходства и основанные на ней алгоритмы.	2
9.	Анализ временных рядов	1
	ВСЕГО	14

Темы лабораторных занятий

<i>№ п/п</i>	<i>Название темы</i>	<i>Количество часов</i>
1.	Решение задач из теории измерений с помощью средств предобработки данных в языках R или Python.	5
2.	Решение задач классификации, кластеризации и частичного обучения с помощью библиотеки sklearn.	5
3.	Проведение регрессионного анализа с помощью библиотеки sklearn или средств языка R.	5
4.	Решение прикладных задач комбинированного типа с использованием языка python и его библиотек (или языка R). Сравнение качества различных алгоритмов. Изучение методик тестирования.	4
5.	Решение прикладных задач конкурентного сходства с использованием языка python и его библиотек (или языка R). Сравнение качества различных алгоритмов. Выбор подходящего алгоритма.	4
6.	Программирование прикладных задач анализа временных рядов.	5
	ВСЕГО	28

6. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Организация самостоятельной работы студентов

Самостоятельная работа студентов по курсу «Интеллектуальный анализ данных» предусматривает:

- систематическое ведение конспекта лекций и повседневную проработку лекционного материала;
- изучение дополнительной технической литературы и интернет-источников, рекомендуемых этой программой;
- добросовестную подготовку к лабораторным занятиям;
- самостоятельную разработку алгоритмов и текстов программ лабораторных работ;
- изучение дополнительного инструментария;
- своевременное и качественное оформление отчётов по лабораторным работам.

<i>№ п/п</i>	<i>Название темы</i>	<i>Количество часов</i>
1.	Обзор состояния проблемы искусственного интеллекта	6
2.	Методы предварительной обработки данных.	10

3.	Классификация задач анализа данных. Базовые гипотезы.	8
4.	Задачи классификации, кластеризации и частичного обучения.	12
5.	Задача выбора и генерации информативной системы признаков.	10
6.	Задача прогнозирования и заполнения пробелов.	12
7.	Задачи комбинированного типа.	8
8.	Функция конкурентного сходства и основанные на ней алгоритмы.	8
9.	Анализ временных рядов	10
	ВСЕГО	84

7. ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ

К каждой теме предусмотрены индивидуальные задания. Пример индивидуальных вариантов заданий:

1. Имеется таблица с информацией о геологических образцах

№	Месторождение	Масса образца	радиоактивность	наличие примесей
12456	Алуфьевское	123	высокая	да
2341	Мочаева	67	средняя	да
562	Алуфьевское	34	низкая	да
12221	Алуфьевское	567	средняя	нет
231	Мочаева	32	средняя	да
3245	Кристел	13	низкая	да
12	Алуфьевское	124	высокая	нет
15	Мочаева	45	низкая	нет

а) Какой столбец (признак) к какому типу относится?

б) Введите метрику (с учетом типов признаков) для этой задачи и вычислите расстояние между образцами 15 и 12.

в) Вычислите расстояние между признаками «радиоактивность» и «месторождение».

8. КОНТРОЛЬНЫЕ ВОПРОСЫ К ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

1. Система полигон сравнения алгоритмов распознавания.
2. Современные технологии выбора признаков
3. Динамическое распознавание
4. Бикластеризация
5. Метод независимых компонент.
6. Задача частичного обучения. Алгоритм Co-learning.
7. Деревья классификаций. Алгоритмы CART, C4.5.
8. Парзеновские окна.
9. Визуализация данных

9. ОБРАЗЕЦ МОДУЛЬНОГО КОНТРОЛЯ

ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»

Физико-технический факультет

Направление подготовки:

09.04.01 Информатика и вычислительная техника

Магистерская программа:
Программа подготовки:
Семестр
Учебная дисциплина

Информатика и вычислительная техника
академическая магистратура
1
Интеллектуальный анализ данных

МОДУЛЬНАЯ КОНТРОЛЬНАЯ РАБОТА 1 ВАРИАНТ №1

Имеется выборка размера 40x10000. Объекты выборки – пациенты. 20 пациентам поставлен диагноз «рассеянный склероз», оставшиеся 20 – здоровы. Признаки – значения активности 10000 генов. Необходимо найти гены, ответственные за предрасположенность пациентов к рассеянному склерозу. Решить эту задачу средствами стандартного пакета анализа данных, предварительно сформулировав ответы на следующие вопросы.

1. Какая это задача анализа данных? Почему?
2. Какие базовые предположения по Вашему мнению нужно использовать для успешного решения этой задачи? Почему?
3. Какой алгоритм (несколько алгоритмов) Вы бы использовали для решения этой задачи? Почему?

Утверждено на заседании кафедры компьютерных технологий,
протокол № 12 от «2» апреля 2020 г.

Заведующий кафедрой
Преподаватель

Ермоленко Т.В.
Бондаренко В.И.

Критерии оценивания модульного контроля

<i>Номер задания</i>	<i>Количество баллов</i>
Вопрос 1	3
Вопрос 2	3
Вопрос 3	3
Решение задания	11
Всего	20

10. ОБРАЗЕЦ ЭКЗАМЕНАЦИОННОГО БИЛЕТА

ГОУ ВПО «ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ»

Физико-технический факультет

Направление подготовки:
Магистерская программа:
Программа подготовки:
Семестр
Учебная дисциплина

09.04.01 Информатика и вычислительная техника
Информатика и вычислительная техника
академическая магистратура
1
Интеллектуальный анализ данных

Экзаменационный билет 1

1. Меры близости между признаками, измеренными в шкалах наименований и сильных.
2. Метод опорных векторов.

3. Задачи комбинированного типа.

Утверждено на заседании кафедры компьютерных технологий,
протокол № 12 от «2» апреля 2020 г.

Заведующий кафедрой
Экзаменатор

Ермоленко Т.В.
Бондаренко В.И.

Критерии оценивания экзамена

<i>Номер задания</i>	<i>Количество баллов</i>
Задание 1	15
Задание 2	15
Задание 3	10
Всего	40

11. ОБРАЗЕЦ ТЕСТОВОГО ЗАДАНИЯ

Тестовое задание не предусмотрено учебной программой.

12. КРИТЕРИИ ОЦЕНИВАНИЯ

*Распределение баллов, которые могут получить студенты
в процессе изучения дисциплины*

Третий семестр

	Содержательный модуль №1					Содержательный модуль №2					Всего
	Лабораторные работы			Мод. контр. работа	Всего С.М. №1	Лабораторные работы			Экзамен	Всего С.М. №2	
	№1	№2	№3			№4	№5	№6			
Макс. балл	5	5	5	20	35	5	10	10	40	65	100

Согласно модульному принципу организации учебного процесса, содержание дисциплины включает в себя два зачётных модуля. Каждый зачётный модуль состоит из теоретического материала и практических задач, выполнение которых требует овладения теорией в указанном в модуле объёме.

К первому модульному контролю студент должен защитить 3 лабораторные работы. *За первую, вторую и третью* лабораторные работы студент может получить по 5 балла.

На первом модульном контроле студент имеет возможность получить 20 баллов за ответы на тестовые вопросы. К каждому тестовому вопросу предполагается не менее четырех вариантов ответа.

К экзамену студент должен защитить 3 следующие лабораторные работы. За четвёртую работу студент может получить 5 балла, за пятую и шестую работу - по 10 баллов.

На экзамене студент имеет возможность получить 40 баллов, ответив на 3 вопроса. Ответы на первые два вопроса оцениваются в 15 баллов, на третий – в 10.

Шкала соответствия баллов национальной шкале

Оценка по шкале ECTS	Оценка по 100-балльной шкале	Оценка по государственной шкале (экзамен, дифференцированный зачет)	Оценка по государственной шкале (зачет)
A	90-100	5 (отлично)	зачтено
B	80-89	4 (хорошо)	зачтено
C	75-79	4 (хорошо)	зачтено
D	70-74	3 (удовлетворительно)	зачтено
E	60-69	3 (удовлетворительно)	зачтено
FX	35-59	2 (неудовлетворительно) с возможностью повторной сдачи	не зачтено
F	0-34	2 (неудовлетворительно) с возможностью повторной сдачи при условии обязательного набора дополнительных баллов	не зачтено

Оценка за овладение курса выставляется по следующим принципам:

– Оценку «отлично» заслуживает студент, который обнаружил глубокие знания при ответах на теоретические вопросы по темам курса, а также выполнил практические задания в полном объёме и набрал более 90 баллов.

– Оценку «хорошо» заслуживает студент, сделавший ошибки в теоретических или практических ответах, которые могут быть интерпретированы как малосущественные для вопросов, которые рассматривались. Студент должен набрать более 75 баллов.

– Оценку «удовлетворительно» заслуживает студент, который выполнил задания неполно и с ошибками, но при этом набрал более 60 баллов.

– Оценку «неудовлетворительно» заслуживает студент, который не выполнил большинства теоретических и практических задач и набрал менее 60 баллов.

13. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОГО ПРОЦЕССА

Лекционные занятия проводятся в аудитории, оснащенной мультимедийной техникой и доской.

Лабораторные занятия проводятся в компьютерном классе, оборудованном компьютерами с лицензионным программным обеспечением, доступом к сети Интернет, столами и доской.

14. РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА

№ п/п	Наименование	Кол-во экземпляров в библиотеке ДонНУ	Наличие электронной версии в ЭБС
----------	--------------	--	--

Основная литература			
1.	Чубукова, И.А. Data Mining / И.А. Чубукова. - 2-е изд., испр. - М.: Интернет-Университет Информационных Технологий, 2008. - 383 с. - (Основы информационных технологий)		
2.	Интеллектуальные системы: учебное пособие / А. Семенов, Н. Соловьев, Е. Чернопрудова, А. Цыганков ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет». - Оренбург : ОГУ, 2013. - 236 с.		
Дополнительная литература			
3.	Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Технологии анализа данных : Data Mining, Visual Mining, Text Mining, OLAP- СПб.: БХВ-Петербург, 2008		

15. ИНФОРМАЦИОННЫЕ РЕСУРСЫ

1. Ссылки на электронные материалы курса. URL: <http://donnu.ru/phys/kt/bondarenko> (дата обращения 10.03.2020 г.)
2. Шитиков В. К., Мастицкий С. Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R. URL: <https://ranalytics.github.io/data-mining/> (дата обращения 19.03.2020 г.)
3. Коллекция материалов по машинному обучению URL: <http://www.machinelearning.ru>
4. Коллекция прикладных задач URL: <http://archive.ics.uci.edu/ml>
5. Документация по пакету sklearn. URL: <http://scikit-learn.org/stable/modules/classes.html>

16. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

1. Python 3 или более старших версий.
2. Пакет sklearn.
3. Пакет R-4 или более старших версий.
4. R-Studio.

Рабочая программа рассмотрена и переутверждена на заседании кафедры компьютерных технологий с изменениями (без изменений) на 2020 год.

Протокол № 12 от «2» апреля 2020 г.

Заведующий кафедрой

Ермоленко Т.В.

Рабочая программа рассмотрена и переутверждена на заседании кафедры компьютерных технологий с изменениями (без изменений) на 2021 год.

Протокол № ____ от «____» _____ 2021 г.

Заведующий кафедрой

Рабочая программа рассмотрена и переутверждена на заседании кафедры компьютерных технологий с изменениями (без изменений) на 2022 год.

Протокол № ____ от «____» _____ 2022 г.

Заведующий кафедрой