

Минобрнауки России
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)

УТВЕРЖДАЮ
Заведующий кафедрой
Сирота Александр Анатольевич
Кафедра технологий обработки и защиты информации
31.08.2020

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Б1.Б.26 Технология обработки информации

1. Код и наименование направления подготовки/специальности:

10.05.01 КОМПЬЮТЕРНАЯ БЕЗОПАСНОСТЬ

2. Профиль подготовки/специализация:

Анализ безопасности компьютерных систем, Математические методы защиты информации

3. Квалификация (степень) выпускника:

Специалитет

4. Форма обучения:

Очная

5. Кафедра, отвечающая за реализацию дисциплины:

Кафедра технологий обработки и защиты информации

6. Составители программы:

Сирота Александр Анатольевич, д.т.н., профессор

7. Рекомендована:

31.08.2020, №7

8. Учебный год:

2023-2024

Семестр(ы):

9

9. Цели и задачи учебной дисциплины:

Изучение теоретических основ и овладение практическими навыками применения методов и средств обработки информации в интересах сопровождения и проектирования информационных, информационно-измерительных и управляющих систем различного назначения; получение профессиональных компетенций в области современных технологий обработки информации.

Основные задачи дисциплины:

обучение студентов базовым понятиям современных технологий обработки информации; обучение студентов современным методам и алгоритмам анализа данных с использованием машинного обучения в рамках статистического и детерминистского подходов;

овладение практическими навыками разработки алгоритмов обработки информации с использованием современных программных средств и технологий;

✗ раскрытие принципов построения и эксплуатации информационных, информационно-измерительных и управляющих систем с точки зрения решения базовых задач обработки информации.

10. Место учебной дисциплины в структуре ООП:

Входит в блок обязательной вариативной части дисциплины Б1.В.

Входные знания в области математического анализа, теории множеств, матричной алгебры, теории вероятностей и математической статистики. Дисциплина является предшествующей для дисциплины «Моделирование систем».

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями выпускников):

Код	Название	Знать	Уметь	Владеть
ОПК-3	способность понимать значение информации в развитии современного общества, применять достижения информационных технологий для поиска и обработки информации по профилю деятельности в глобальных компьютерных сетях, библиотечных фондах и иных источниках информации	Базовые понятия методов обработки информации, современные методы и алгоритмы анализа данных и машинного обучения в рамках статистического и детерминистского подходов	Использовать стандартное и оригинальное программное обеспечение для проведения исследований и разработок в области алгоритмов анализа данных и машинного обучения, формировать рекомендации по принципам построения и параметрам алгоритмов в области профессиональной деятельности.	Практическими навыками разработки и применения алгоритмов и технологий обработки информации в части анализа данных и машинного обучения.

Код	Название	Знать	Уметь	Владеть
ОПК-7	способность учитывать современные тенденции развития информатики и вычислительной техники, компьютерных технологий в своей профессиональной деятельности, работать с программными средствами общего и специального назначения	Средства и технологии организации и проведения статистического компьютерного моделирования алгоритмов обработки информации, возможности современных программных сред для реализации исследований и разработок в области алгоритмов анализа данных и машинного обучения.	Проводить синтез и анализ алгоритмов обработки информации для решения конкретных практических задач, использовать математические методы в интересах аналитической и численной оценки основных показателей эффективности алгоритмов анализа данных и машинного обучения	Навыками разработки и моделирования алгоритмов обработки информации в современных инструментальных средах, навыками проведения компьютерного эксперимента по оценке эффективности алгоритмов анализа данных и машинного обучения, навыками тестирования компьютерных моделей алгоритмов обработки информации в среде Matlab.

12. Объем дисциплины в зачетных единицах/час:

5/180

Форма промежуточной аттестации:

Экзамен

13. Виды учебной работы

Вид учебной работы	Семестр 9	Всего
Аудиторные занятия	68	68
Лекционные занятия	34	34
Практические занятия		0
Лабораторные занятия	34	34
Самостоятельная работа	76	76
Курсовая работа		0
Промежуточная аттестация	36	36
Часы на контроль	36	36
Всего	180	180

13.1. Содержание дисциплины

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины
п/п	Наименование раздела дисциплины	Содержание раздела дисциплины
1	Введение. Базовые понятия технологий обработки информации	Лекция по разделу 1.Обработка информации, анализ данных, машинное обучение. Основные понятия и определения.
2	Статистическая теория распознавания образов	Лекции по разделу 2.Случайные величины и случайные векторы. 3.Байесовская теория принятия решения применительно к задаче распознавания образов. 4.Распознавание образов, описываемых гауссовскими случайными векторами. 5.Распознавание образов, описываемых произвольными законами распределения. 6.Распознавание образов в условиях параметрической неопределенности на основе обучения с учителем. 7. Распознавание образов в условиях непараметрической неопределенности на основе обучения с учителем. Лабораторные работы по разделу 1. Моделирование случайных величин и случайных векторов. 2. Распознавание образов, описываемых гауссовскими случайными векторами с одинаковыми матрицами ковариаций. 3. Распознавание образов, описываемых гауссовскими случайными векторами с различными матрицами ковариаций. 4. Распознавание образов, описываемых бинарными признаками. 5. Исследование непараметрических алгоритмов оценивания плотности распределения случайной величины. 6. Распознавание образов на основе непараметрических алгоритмов оценивания плотности распределения случайной величины.

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины
3	Распознавание образов в рамках детерминистского (геометрического) подхода	<p>Лекции по разделу</p> <p>8. Распознавание образов с использованием функций расстояния.</p> <p>9. Нелинейные преобразования и спрямляющие пространства.</p> <p>10. Метод опорных векторов и его модификации</p> <p>11. Композиционные алгоритмы распознавания образов. Бэггинг и бустинг.</p> <p>Лабораторные занятия по разделу</p> <p>7. Распознавание образов на основе метода машин опорных векторов.</p> <p>8. Распознавание образов с использованием алгоритма «Случайный лес».</p>
4	Основы кластерного анализа в рамках статистического и детерминистского подходов	<p>Лекции по разделу</p> <p>11. Кластеризация при известном числе классов. Алгоритм K- внутригрупповых средних и алгоритм иерархической кластеризации.</p> <p>12. Критерии оценки числа классов и использование алгоритмов кластеризации в условиях неизвестного числа классов.</p> <p>Лабораторные занятия по разделу</p> <p>9. Исследование методов кластеризации данных на примере алгоритмов иерархической группировки и k-средних.</p>
5	Теоретические основы регрессионного анализа	<p>Лекции по разделу</p> <p>14. Основы регрессионного анализа данных. Постановка задачи методы и подходы</p> <p>15. Построение линейной и нелинейной регрессии в рамках детерминистского подхода</p>

13.2. Темы (разделы) дисциплины и виды занятий

№ п/п	Наименование темы (раздела) дисциплины	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа	Всего
1	Введение. Базовые понятия технологий обработки информации	2	0	0	4	6
2	Статистическая теория распознавания образов	12		16	20	48

№ п/п	Наименование темы (раздела) дисциплины	Лекционные занятия	Практические занятия	Лабораторные занятия	Самостоятельная работа	Всего
3	Распознавание образов в рамках детерминистского (геометрического) подхода	10		12	20	42
4	Основы кластерного анализа в рамках статистического и детерминистского подходов	6		6	20	32
5	Теоретические основы регрессионного анализа	4			12	16
		34	0	34	76	144

14. Методические указания для обучающихся по освоению дисциплины

1. При изучении дисциплины рекомендуется использовать следующие средства: рекомендуемую основную и дополнительную литературу;

- методические указания и пособия;
- контрольные задания для закрепления теоретического материала;
- электронные версии учебников и методических указаний для выполнения лабораторно - практических работ (при необходимости материалы рассылаются по электронной почте).

2. Для максимального усвоения дисциплины рекомендуется проведение письменного опроса (тестирование, решение задач) обучающихся по материалам лекций и практических работ. Подборка вопросов для тестирования осуществляется на основе изученного теоретического материала. Такой подход позволяет повысить мотивацию при конспектировании лекционного материала.

3. При проведении лабораторных занятий обеспечивается максимальная степень соответствия с материалом лекционных занятий и осуществляется экспериментальная проверка методов, алгоритмов и технологий обработки информации, излагаемых в рамках лекций.

4. При переходе на дистанционный режим обучения для создания электронных курсов, чтения лекций онлайн и проведения лабораторно- практических занятий используется информационные ресурсы Образовательного портала "Электронный университет ВГУ (<https://edu.vsu.ru>), базирующегося на системе дистанционного обучения Moodle, развернутой в университете.

5. При использовании дистанционных образовательных технологий и электронного обучения обучающиеся должны выполнять все указания преподавателей, вовремя подключаться к онлайн - занятиям, ответственно подходить к заданиям для самостоятельной работы.

15. Перечень основной и дополнительной литературы, ресурсов интернет, необходимых для освоения дисциплины

а) основная литература:

№ п/п	Источник
1	Сирота, Александр Анатольевич. Методы и алгоритмы анализа данных и их моделирование в MATLAB : [учебное пособие] / А.А. Сирота .— Санкт-Петербург : БХВ-Петербург, 2016 .— 381 с. : ил. — Библиогр.: с. 371-374 .— Предм. указ.: с. 377-381 .— ISBN 978-5-9775-3778-0.
2	Хултен Д. Разработка интеллектуальных систем : введение в технологию машинного обучения: практическое пособие/Д. Хултен. - Москва: ДМК Пресс, 2019.-284с. : ил. Библиогр. в конце разд. — ISBN 978-5-97060-760-2. ЭБС ВГУ «Университетская библиотека online»

б) дополнительная литература:

№ п/п	Источник
1	Алгазинов, Эдуард Константинович. Анализ и компьютерное моделирование информационных процессов и систем : [учебное пособие для студ. вузов, обуч. по специальности 080801 "Приклад. информатика" и др. междисциплинар. специальностям] / Э.К. Алгазинов, А.А. Сирота ; под общ. ред. А.А. Сироты .— М. : Диалог-МИФИ, 2009 .— 416 с. : ил .— Библиогр. в конце разд. — ISBN 978-5-86404-233-5
2	Чубукова, И.А. Data Mining / И.А. Чубукова .- 2-е изд., испр. - Москва : Интернет-Университет Информационных Технологий, 2008 .- 383 с. - (Основы информационных технологий) .- ISBN 978-5-94774-819-2 .- <URL: http://biblioclub.ru/index.php?page=book&id=233055 >.
3	Андерсон Т. Введение в многомерный статистический анализ / Т. Андерсон. – М.: Физматлит, 1963. – 500 с.
4	Фукунага К. Введение в статистическую теорию распознавания образов / К. Фукунага. – М.: Наука, 1979. – 368 с.

в) информационные электронно-образовательные ресурсы:

№ п/п	Ресурс
1	Электронный каталог Научной библиотеки Воронежского государственного университета. – (http // www.lib.vsu.ru/).
2	Образовательный портал «Электронный университет ВГУ».– (https://edu.vsu.ru/)

№ п/п	Ресурс
3	<p>«Университетская библиотека online» - Контракт № 3010-07/33-19 от 11.11.2019</p> <p>«Консультант студента» - Контракт № 3010-07/34-19 от 11.11.2019</p> <p>ЭБС «Лань» - Договор 3010-04/05-20 от 26.02.2020</p> <p>«РУКОНТ» (ИТС Контекстум) - Договор ДС-208 от 01.02.2018</p> <p>ЭБС «Юрайт» - Договор № 43/8 от 10.02.2020</p>

16. Перечень учебно-методического обеспечения для самостоятельной работы

№ п/п	Источник
1	Сирота, Александр Анатольевич. Методы и алгоритмы анализа данных и их моделирование в MATLAB : [учебное пособие] / А.А. Сирота .— Санкт-Петербург : БХВ-Петербург, 2016 .— 381 с. : ил. — Библиогр.: с. 371-374 .— Предм. указ.: с. 377-381 .— ISBN 978-5-9775-3778-0.
2	Практикум по курсу "Моделирование систем" [Электронный ресурс] : учебно-методическое пособие для вузов : [для студ. 4-5 курсов фак. компьютер. наук днев. и вечер. формы обучения; для направлений: 230200 - Информ. системы, 230400 - Информ. системы и технологии; специальности, 230201 - Информ. системы и технологии]. Ч. 1 / Воронеж. гос. ун-т ; сост.: А.А. Сирота, Е.Ю. Митрофанова , М.А. Дрюченко .— Электрон. текстовые дан. — Воронеж : Издательско-полиграфический центр Воронежского государственного университета, 2013

17. Информационные технологии, используемые для реализации учебной дисциплины, включая программное обеспечение и информационно-справочные системы (при необходимости):

Для реализации учебного процесса используются:

1. ПО Microsoft в рамках подписки "Imagine/Azure Dev Tools for Teaching", договор №3010-16/96-18 от 29 декабря 2018г.
2. ПО MATLAB Classroom ver. 7.0, 10 конкурентных бессрочных лицензий на каждый, компоненты: Matlab, Simulink, Stateflow, 1 тулбокс, N 21127/VRN3 от 30.09.2011 (за счет проекта ЕК TEMPUS/ERAMIS).
3. ПО Матлаб в рамках подписки "Университетская лицензия на программный комплекс для ЭВМ - MathWorks, Headcount – 25 ": лицензия до 31.01.2022, сублицензионный контракт 3010-07/01-19 от 09.01.19.
4. При проведении занятий в дистанционном режиме обучения используются технические и информационные ресурсы Образовательного портала "Электронный университет ВГУ (<https://edu.vsu.ru>), базирующегося на системе дистанционного обучения Moodle, развернутой в университете, а также другие доступные ресурсы сети Интернет.

18. Материально-техническое обеспечение дисциплины:

1. Мультимедийная лекционная аудитория (корп.1а, ауд. № 479), ПК-Intel-i3, рабочее место преподавателя: проектор, видеоконмутатор, микрофон, аудиосистема, специализированная мебель: доски меловые 2 шт., столы 60 шт., лавки 30 шт., стулья 64 шт.; доступ к фондам учебно - методической документации и электронным библиотечным системам, выход в Интернет.

2. Компьютерный класс (один из №1-4 корп. 1а, ауд. № 382-385), ПК-Intel-i3 16 шт., специализированная мебель: доска маркерная 1 шт., столы 16 шт., стулья 33 шт.; доступ к фондам учебно-методической документации и электронным изданиям, доступ к электронным библиотечным системам, выход в Интернет.

19. Фонд оценочных средств:

19.1 Перечень компетенций с указанием этапов формирования и планируемых результатов обучения

Код и содержание компетенции (или ее части)	Планируемые результаты обучения (показатели достижения заданного уровня освоения компетенции посредством формирования знаний, умений, навыков)	Этапы формирования компетенции (разделы (темы) дисциплины или модуля и их наименование)	ФОС* (средства оценивания)
ОПК-4 Способность понимать значение информации в развитии современного общества, применять информационные технологии для поиска и обработки информации	Знает базовые понятия методов обработки информации, современные методы и алгоритмы анализа данных и машинного обучения в рамках статистического и детерминистского подходов. Умеет использовать стандартное и оригинальное программное обеспечение для проведения исследований и разработок в области алгоритмов анализа данных и машинного обучения, формировать рекомендации по принципам построения и параметрам алгоритмов в области профессиональной деятельности. Владеет практическими навыками разработки и применения алгоритмов и технологий обработки информации в части анализа данных и машинного обучения.	Разделы 1-5 Введение. Базовые понятия технологий обработки информации. Статистическая теория распознавания образов. Распознавание образов в рамках детерминистского (геометрического) подхода. Основы кластерного анализа в рамках статистического и детерминистского подходов. Теоретические основы регрессионного анализа.	Контрольная работа по соответствующим разделам и темам. Задания для выполнения лабораторных работ 1-9.

Код и содержание компетенции (или ее части)	Планируемые результаты обучения (показатели достижения заданного уровня освоения компетенции посредством формирования знаний, умений, навыков)	Этапы формирования компетенции (разделы (темы) дисциплины или модуля и их наименование)	ФОС* (средства оценивания)
<p>ОПК-7 Способность учитывать современные тенденции развития информатики и вычислительной техники, компьютерных технологий в своей профессиональной деятельности, работать с программными средствами общего и специального назначения</p>	<p>Знает средства и технологии организации и проведения статистического компьютерного моделирования алгоритмов обработки информации, возможности современных программных сред для реализации исследований и разработок в области алгоритмов анализа данных и машинного обучения. Умеет проводить синтез и анализ алгоритмов обработки информации для решения конкретных практических задач, использовать математические методы в интересах аналитической и численной оценки основных показателей эффективности алгоритмов анализа данных и машинного обучения. Владеет навыками разработки и моделирования алгоритмов обработки информации в современных инструментальных средах, навыками проведения компьютерного эксперимента по оценке эффективности алгоритмов анализа данных и машинного обучения, навыками тестирования компьютерных моделей алгоритмов обработки информации в среде Matlab.</p>	<p>Разделы 2-5 Статистическая теория распознавания образов. Распознавание образов в рамках детерминистского (геометрического) подхода. Основы кластерного анализа в рамках статистического и детерминистского подходов. Теоретические основы регрессионного анализа.</p>	<p>Контрольная работа по соответствующим разделам и темам. Задания для выполнения лабораторных работ 1-9.</p>

Код и содержание компетенции (или ее части)	Планируемые результаты обучения (показатели достижения заданного уровня освоения компетенции посредством формирования знаний, умений, навыков)	Этапы формирования компетенции (разделы (темы) дисциплины или модуля и их наименование)	ФОС* (средства оценивания)
Промежуточная аттестация			Комплект КИМ

19.2 Описание критериев и шкалы оценивания компетенций (результатов обучения) при промежуточной аттестации

Для оценивания результатов обучения на экзамене используются следующие содержательные показатели (формулируется с учетом конкретных требований дисциплины):

1. знание теоретических основ учебного материала, основных определений, понятий и используемой терминологии;
2. умение проводить обоснование и представление основных теоретических и практических результатов (теорем, алгоритмов, методик) с использованием математических выкладок, блок-схем, структурных схем и стандартных описаний к ним;
3. умение связывать теорию с практикой, иллюстрировать ответ примерами, в том числе, собственными, умение выявлять и анализировать основные закономерности, полученные, в том числе, в ходе выполнения лабораторно-практических заданий;
4. умение обосновывать свои суждения и профессиональную позицию по излагаемому вопросу;
5. владение навыками программирования и экспериментирования с компьютерными моделями алгоритмов обработки информации в среде Matlab в рамках выполняемых лабораторных заданий;
6. владение навыками проведения компьютерного эксперимента, тестирования компьютерных моделей алгоритмов обработки информации.

Различные комбинации перечисленных показателей определяют критерии оценивания результатов обучения (сформированности компетенций) на государственном экзамене:

- высокий (углубленный) уровень сформированности компетенций;
- повышенный (продвинутый) уровень сформированности компетенций;
- пороговый (базовый) уровень сформированности компетенций.

Для оценивания результатов обучения на государственном экзамене используется 4-балльная шкала: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно». Для оценивания результатов обучения на зачете используется – зачтено, не зачтено по результатам тестирования.

Соотношение показателей, критериев и шкалы оценивания результатов обучения на государственном экзамене представлено в следующей таблице.

Критерии оценивания компетенций и шкала оценок на экзамене

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
Обучающийся демонстрирует полное соответствие знаний, умений, навыков по приведенным критериям свободно оперирует понятийным аппаратом и приобретенными знаниями, умениями, применяет их при решении практических задач.	Повышенный уровень	Отлично

<p>Ответ на контрольно-измерительный материал не полностью соответствует одному из перечисленных выше показателей, но обучающийся дает правильные ответы на дополнительные вопросы. При этом обучающийся демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателям, но допускает незначительные ошибки, неточности, испытывает затруднения при решении практических задач.</p>	<p>Базовый уровень</p>	<p>Хорошо</p>
<p>Обучающийся демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателям, допускает значительные ошибки при решении практических задач. При этом ответ на контрольно-измерительный материал не соответствует любым двум из перечисленных показателей, обучающийся дает неполные ответы на дополнительные вопросы.</p>	<p>Пороговый уровень</p>	<p>Удовлетворительно</p>
<p>Ответ на контрольно-измерительный материал не соответствует любым трем из перечисленных показателей. Обучающийся демонстрирует отрывочные, фрагментарные знания, допускает грубые ошибки</p>	<p>–</p>	<p>Неудовлетворительно</p>

19.3 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующие этапы формирования компетенций в процессе освоения образовательной программы

19.3.1 Примерный перечень применяемых оценочных средств в ходе текущей аттестации

№ п/п	Наименование оценочного средства	Представление оценочного средства в фонде	Критерии оценки
1	2	3	4
1	Устный опрос	Вопросы по темам/разделам дисциплины	Правильный ответ – зачтено, неправильный или принципиально неточный ответ - не зачтено
2	Контрольная работа по разделам дисциплины	Теоретические вопросы по темам/разделам дисциплины	Шкала оценивания соответствует приведенной в разделе 19.2
3	Лабораторная работа	Содержит 9 лабораторных заданий, предусматривающие разработку, тестирование и эксплуатацию моделей и алгоритмов анализа данных с использованием различных методов обучения.	При успешно выполнении работы ставится оценка зачтено и осуществляется допуск к экзамену, в противном случае ставится оценка не зачтено и обучающийся не допускается к экзамену.

4	КИМ промежуточной аттестации	Каждый контрольно-измерительный материал для проведения промежуточной аттестации включает 2 задания вопросов для контроля знаний, умений и владений в рамках оценки уровня сформированности компетенции.	Шкалы оценивания приведены в разделе 19.2
---	------------------------------	--	---

19.3.2. Примерный перечень вопросов к экзамену

№	Содержание
1	Общая схема процесса обработки информации. Основные виды и процедуры обработки информации.
2	Классификация базовых подходов к обработке информации. Задачи обработки информации, решаемые в рамках технологии DATA MINING
3	Байесовская теория решений. Решающее правило на основе минимизации условного риска
4	<i>Байесовская теория решений. Решающие правила на основе максимума апостериорной вероятности и функции правдоподобия</i>
5	<i>Понятие разделяющих функций. Обобщенная структура решающего правила</i>
6	<i>Распознавание образов, описываемых гауссовскими векторами с различными математическими ожиданиями</i>
7	<i>Распознавание образов, описываемых гауссовскими векторами с различными ковариационными матрицами</i>
8	<i>Распознавание образов, описываемых произвольными законами распределения</i>
9	<i>Распознавание образов, описываемых бинарными признаками</i>
10	<i>Основы теории оценивания. Параметрическое и непараметрическое оценивание в статистических моделях данных.</i>
11	<i>Линейная регрессия</i>
12	Теорема о нормальной корреляции
13	<i>Фильтр Калмана-Бьюси</i>
14	<i>Основы регрессионного анализа данных.</i>
14	<i>Параметрическое обучение в задачах распознавания. Метод максимума правдоподобия.</i>

15	Параметрическое обучение в задачах распознавания. Метод максимума апостериорной вероятности.
16	<i>Непараметрическое обучение в задачах распознавания. Метод Парзена</i>
17	<i>Непараметрическое обучение в задачах распознавания. Метод К-ближайших соседей</i>
18	Распознавание образов с помощью функций расстояния
19	Метод опорных векторов
20	Композиционные алгоритмы. Бэггинг
21	Композиционные алгоритмы. Бустинг
22	Кластеризация образов. Алгоритм К-средних
23	Кластеризация образов. Иерархическая процедура группирования.
24	Кластерный анализ в условиях неизвестного числа классов
25	Нелинейные преобразования и спрямляющие пространства

19.3.3. Пример задания для выполнения лабораторной работы

Лабораторная работа №5

«Исследование непараметрических алгоритмов оценивания плотности распределения случайной величины»

Цель работы:

Исследовать алгоритмы оценивания плотности распределения случайных величин и случайных векторов на основе методов Парзена и k ближайших соседей.

Форма контроля: *отчёт в электронном виде*

Количество отведённых аудиторных часов: 4

Задание:

Получите у преподавателя вариант задания и напишите код, реализующий соответствующий алгоритм обработки. Для ответа на поставленные вопросы требуется провести численный эксперимент и представить соответствующие графики. Составьте отчёт о проделанной работе, в котором отразите следующие пункты:

1. ФИО исполнителя и номер группы.
2. Название и цель лабораторной работы.
3. Номер своего варианта.
4. Код, написанный исполнителем.
5. Графики, соответствующие варианту задания исполнителя.

Примеры контрольных вопросов:

1. При каком значении параметра оконной функции достигается минимум ошибки оценивания по критерию, указанному в варианте исполнителя? Ответ подтвердить графиками,

представленными в отчёте.

2. Какой вид оконной функции обеспечивает оптимальную оценку плотности распределения в реализованном Вами исследовании? Ответ подтвердить графиками, представленными в отчёте.

Варианты заданий:

1. Вычислить абсолютную ошибку оценивания плотности распределения случайной величины при использовании оценки Парзена. Построить график зависимости ошибки оценивания от величины параметра оконной функции следующего вида:
 - a. гауссовская функция;
 - b. показательная функция;
 - c. оконная прямоугольная функция;
 - d. оконная треугольная функция.
2. Вычислить среднеквадратичную ошибку оценивания плотности распределения случайной величины по методу Парзена для заданного вида окна.
3. Окно какого вида будет обеспечивать оптимальное по критерию среднеквадратичной ошибки оценивание плотности распределения случайной величины в соответствии с методом Парзена? Построить графики зависимостей ошибок от объема обучающей выборки. Сравните следующие виды окон:
 - a. гауссовская функция, показательная функция и оконная прямоугольная функция;
 - b. гауссовская функция, оконная прямоугольная функция и оконная треугольная функция.
4. Вычислить абсолютную ошибку оценивания плотности распределения случайного вектора в двумерном пространстве признаков при использовании оценки Парзена. Построить график зависимости ошибки оценивания от величины параметра оконной функции. Используйте одну из следующих оконных функций:
 - a. гауссовская функция с использованием диагональной матрицы;
 - b. гауссовская функция с использованием матрицы ковариаций;
 - c. показательная функция;
 - d. оконная прямоугольная функция;
 - e. оконная треугольная функция.
5. Окно какого вида будет обеспечивать оптимальное по критерию среднеквадратичной ошибки оценивание плотности распределения двумерного случайного вектора по методу Парзена? Построить графики зависимостей ошибок от объема обучающей выборки. Сравните следующие виды окон:
 - a. гауссовская функция с использованием диагональной матрицы и гауссовская функция с использованием матрицы ковариаций;
 - b. гауссовская функция с использованием диагональной матрицы, показательная функция и оконная треугольная функция;
 - c. гауссовская функция с использованием матрицы ковариаций, оконная прямоугольная функция и оконная треугольная функция.
6. Получить оценки плотности распределения случайной величины на основе метода k ближайших соседей. Определить значение параметра k, при котором достигается минимум среднеквадратичной ошибки оценивания.
7. Реализовать оценку плотности распределения двумерного случайного вектора, плотность которого задается на основе пяти гауссовых функций. Определить оптимальное по критерию среднеквадратичной ошибки оценивания значение параметра k.

19.3.4. Пример контрольно-измерительного материала

УТВЕРЖДАЮ

Направление подготовки / специальность 10.05.01 Компьютерная безопасность

Дисциплина Б1.Б.26 Технологии обработки информации

Форма обучения Очное

Вид контроля Экзамен

Вид аттестации Промежуточная

Контрольно-измерительный материал № 1

1. Общая схема процесса обработки информации. Основные виды и процедуры обработки информации
2. Непараметрическое обучение в задачах распознавания. Метод Парзена.

Преподаватель _____ А.А. Сирота

19.4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Оценка знаний, умений и навыков, характеризующая этапы формирования компетенций в рамках изучения дисциплины осуществляется в ходе текущей и промежуточной аттестаций.

Текущая аттестация проводится в соответствии с Положением о текущей аттестации обучающихся по программам высшего образования Воронежского государственного университета. Текущая аттестация проводится в формах устного опроса (индивидуальный опрос, фронтальная беседа) и письменных работ (контрольные, лабораторные работы). При оценивании могут использоваться количественные или качественные шкалы оценок.

Промежуточная аттестация может включать в себя теоретические вопросы, позволяющие оценить уровень полученных знаний и/или практическое (ие) задание(я), позволяющее (ие) оценить степень сформированности умений и навыков.

При оценивании используется количественная шкала. Критерии оценивания приведены выше в таблице раздела 19.2.