МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«Национальный исследовательский ядерный университет «МИФИ»

Обнинский институт атомной энергетики –

филиал федерального государственного автономного образовательного учреждения высшего образования

«Национальный исследовательский ядерный университет «МИФИ»

**(ИАТЭ НИЯУ МИФИ)**

,

Одобрено УМС ИАТЭ НИЯУ МИФИ,

Протокол №2-8/2021 От 30.08.2021

**РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ**

|  |
| --- |
| Распознавание образов |
| *Шифр, название дисциплины* |
|  |
| для студентов направления подготовки |
|  |
| 01.03.02 «Прикладная математика и информатика» |
| *Шифр, название специальности/направления подготовки* |
|  |
|  |
| профиля |
| Прикладная информатика |
| *Шифр, название специализации/профиля* |
|  |
|  |
| Форма обучения: очная |

**г. Обнинск 2021 г.**

Программа составлена в соответствии с требованиями образовательного стандарта высшего образования национального исследовательского ядерного университета «МИФИ» по направлению подготовки 01.03.02 – Прикладная математика и информатика.

Программу составила:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А.Н. Чепурко, доцент, к.ф.-м.н, доцент

Рецензент:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Г.Е. Деев, доцент, к.ф.-м.н, доцент

Программа рассмотрена на заседании отделения интеллектуальных кибернетических систем (О) (протокол № 5/7 от «30» июля 2021 г.)

Руководитель образовательной программы

01.03.02 – «Прикладная математика и информатика»

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ С.В. Ермаков

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_2021 г.

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате освоения ООП бакалавриата обучающийся должен овладеть следующими результатами обучения по дисциплине:

|  |  |  |
| --- | --- | --- |
| Коды компетенций | **Результаты освоения ООП**  **Содержание компетенций\*** | **Перечень планируемых результатов обучения по дисциплине\*\*** |
| ПК-8 | Способность использовать современную вычислительную технику, многопроцессорные ЭВМ, сетевые технологии при решении производственных задач в области математического моделирования физических процессов. | Знать: основные методы классификации и кластеризации.  Уметь: реализовывать на практике наиболее распространённые алгоритмы классификации.  Владеть: навыками решения задач распознавания образов. |

2. Место дисциплины в структуре ООП бакалавриата

Дисциплина реализуется в рамках профессионального модуля.

Для освоения дисциплины необходимы компетенции, сформированные в рамках изучения следующих дисциплин: «Аналитическая геометрия и линейная алгебра», «Программирование», «Теория вероятностей и математическая статистика».

Дисциплина изучается на 3 курсе в 6 семестре.

3. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам занятий) и на самостоятельную работу обучающихся

Общая трудоемкость (объем) дисциплины составляет 3 зачетных единиц (з.е.), 108 академических часов.

3.1. Объём дисциплины по видам учебных занятий (в часах)

|  |  |  |
| --- | --- | --- |
| Объем дисциплины | Всего часов | |
| Очная форма обучения | Заочная форма обучения |
| Общая трудоемкость дисциплины | 108 |  |
| Контактная работа обучающихся с преподавателем  (по видам учебных занятий) (всего) | 32 |  |
| Аудиторная работа (всего): | 32 |  |
| *в том числе:* |  |  |
| лекции | 16 |  |
| семинары, практические занятия | 16 |  |
| лабораторные работы | - |  |
| Внеаудиторная работа (всего): | - |  |
| *в том числе*, индивидуальная работа  обучающихся с преподавателем: | - |  |
| курсовое проектирование | - |  |
| групповая, индивидуальная консультация  и иные виды учебной  деятельности, предусматривающие  групповую или индивидуальную  работу обучающихся с преподавателем | - |  |
| творческая работа (эссе) | - |  |
| Самостоятельная работа обучающихся (всего) | 76 |  |
| Вид промежуточной аттестации обучающегося (зачет) | - |  |

4. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины и трудоемкость по видам учебных занятий

(в академических часах)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № п/п | Наименование раздела /темы дисциплины | Общая трудоём- кость всего  (в часах) | Виды учебных занятий, включая самостоятельную работу обучающихся и трудоемкость  *(в часах)* | | | | Формы текущего контроля успевае-мости |
| Аудиторные учебные занятия | | | СРО |
| Лек | Сем/Пр | Лаб |
| 1. | **Введение** | 14 | 2 | 2 | - | 10 |  |
| 1.1. | Основные понятия распознавания образов | 6 | 1 | 1 | - | 4 | ИДЗ № 1 |
| 1.2. | Решающие функции. | 8 | 1 | 1 | - | 6 | ИДЗ № 1 |
| 2. | **Выявление кластеров** | 42 | 6 | 6 | - | 30 |  |
| 2.1. | Задача кластеризации. | 14 | 2 | 2 | - | 10 | ИДЗ № 2, ИДЗ № 3 |
| 2.2. | Алгоритмы кластеризации. | 28 | 4 | 4 | - | 20 | ИДЗ № 2, ИДЗ № 3 |
| 3. | Классификация образов | 52 | 8 | 8 | - | 36 |  |
| 3.1. | На основе функции расстояния | 10 | 2 | 2 | - | 6 | ИДЗ № 1 |
| 3.2. | На основе функции правдоподобия | 21 | 3 | 3 | - | 15 | ИДЗ № 4 |
| 3.3. | Обучаемые классификаторы. | 21 | 3 | 3 | - | 15 | ИДЗ № 5 |

4.2. Содержание дисциплины, структурированное по разделам (темам)

Лекционный курс

|  |  |  |
| --- | --- | --- |
| № | Наименование раздела /темы дисциплины | Содержание |
| 1. | **Введение** | |
| 1.1. | Основные понятия распознавания образов | Основные задачи распознавания образов. Примеры автоматических систем распознавания образов. |
| 1.2. | Решающие функции. | Линейные решающие функции. Обобщённые решающие функции. Пространство образов и пространство весов. |
| 2. | **Выявление кластеров** | |
| 2.1. | Задача кластеризации. | Постановка задачи кластеризации. Меры сходства. Оценка результатов процесса кластеризации. |
| 2.2. | Алгоритмы кластеризации. | Простейший алгоритм, использующий пороговую величину. Алгоритм максиминного расстояния. Алгоритм К внутригрупповых средних. Алгоритм ИСОМАД. |
| 3. | Классификация образов | |
| 3.1. | На основе функции расстояния | Классификация образов по критерию минимума расстояния. Случай единственности эталона. Случай множественности эталонов. |
| 3.2. | На основе функции правдоподобия | Классификация образов как задача теории статистических решений. Байесовский классификатор в случае образов, характеризующихся нормальным распределением. Вероятности ошибок. |
| 3.3. | Обучаемые классификаторы. | Детерминистский подход в построении обучаемых классификаторов. Алгоритмы перцептрона, НСКО как разновидности градиентного метода. Доказательство сходимости алгоритмов перцептрона и НСКО.  Статистический подход в построении обучаемых классификаторов. Алгоритм Роббинса-Монро. Алгоритм корректирующих приращений. Алгоритм НСКО. |

Практические/семинарские занятия

|  |  |  |
| --- | --- | --- |
| № | Наименование раздела /темы дисциплины | Содержание |
| 1. | **Введение** | |
| 1.1. | Основные понятия распознавания образов | Задачи классификации и кластеризации. |
| 1.2. | Решающие функции. | Линейные решающие функции. Разделяющие границы. Области непринятия решений. |
| 2. | **Выявление кластеров** | |
| 2.1. | Задача кластеризации. | Постановка задачи кластеризации. |
| 2.2. | Алгоритмы кластеризации. | Использование пороговой величины Т. Алгоритм максиминного расстояния. Алгоритм К внутригрупповых средних. |
| 3. | **Численные методы оптимизации** | |
| 3.1. | На основе функции расстояния | Классификация образов по критерию минимума расстояния в случае единственности эталона. |
| 3.2. | На основе функции правдоподобия | Байесовский классификатор. |
| 3.3. | Обучаемые классификаторы. | Алгоритм подкрепления-наказания.  Алгоритм НСКО. |

Лабораторные занятия

Не предусмотрены.

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

В качестве учебно-методических материалов используется рекомендованная литература.

6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

6.1. Паспорт фонда оценочных средств по дисциплине

|  |  |  |  |
| --- | --- | --- | --- |
| **№ п/п** | **Контролируемые разделы (темы) дисциплины (результаты по разделам)** | **Код контролируемой компетенции (или её части) / и ее формулировка** | **Наименование оценочного средства** |
| 1.1. | Основные понятия распознавания образов | ПК-8: Способность использовать современную вычислительную технику, многопроцессорные ЭВМ, сетевые технологии при решении производственных задач в области математического моделирования физических процессов. | ИДЗ № 1 |
| 1.2. | Решающие функции. | ПК-8 | ИДЗ № 1 |
| 2.1. | Задача кластеризации. | ПК-8 | ИДЗ № 2, ИДЗ № 3 |
| 2.2. | Алгоритмы кластеризации. | ПК-8 | ИДЗ № 2, ИДЗ № 3 |
| 3.1. | На основе функции расстояния | ПК-8 | ИДЗ № 1 |
| 3.2. | На основе функции правдоподобия | ПК-8 | ИДЗ № 4 |
| 3.3. | Обучаемые классификаторы. | ПК-8 | ИДЗ № 5 |

6.2. Типовые контрольные задания или иные материалы

*6.2.1. Зачет*

В зачетном задании два теоретических вопроса.

Теоретические вопросы для зачета:

1. Основные задачи распознавания образов.
2. Линейные решающие функции. Обобщённые решающие функции.
3. Пространство образов и пространство весов.
4. Постановка задачи кластеризации. Меры сходства.
5. Оценка результатов процесса кластеризации.
6. Алгоритм кластеризации, использующий пороговую величину.
7. Алгоритм максиминного расстояния.
8. Алгоритм К внутригрупповых средних.
9. Алгоритм ИСОМАД.
10. Классификация образов по критерию минимума расстояния в случае единственности эталона.
11. Классификация образов по критерию минимума расстояния в случае множественности эталонов
12. Классификация образов как задача теории статистических решений.
13. Байесовский классификатор в случае образов, характеризующихся нормальным распределением.
14. Вероятности ошибок в Байесовском классификаторе.
15. Детерминистский подход в построении обучаемых классификаторов. Алгоритм перцептрона как разновидность градиентного метода.
16. Доказательство сходимости алгоритма перцептрона.
17. Детерминистский подход в построении обучаемых классификаторов. Алгоритм НСКО как разновидность градиентного метода.
18. Доказательство сходимости алгоритма НСКО.
19. Алгоритм Роббинса-Монро.
20. Статистический подход в построении обучаемых классификаторов. Алгоритм корректирующих приращений.
21. Статистический подход в построении обучаемых классификаторов. Алгоритм НСКО.

Критерий оценки – правильность и полнота ответа на вопросы. Оценка выставляется по шкале от 0 до 40 баллов: теоретические вопросы –30 баллов, 10 баллов – дополнительные вопросы. Зачет считается сданным при оценке не ниже 25 баллов.

*6.2.2. Индивидуальное домашнее задание (ИДЗ)*

Индивидуальное домашнее задание состоит из пяти заданий по следующим темам: классификация образов по критерию минимума расстояния, алгоритм максиминного расстояния, алгоритм К внутригрупповых средних, Байесовский классификатор, алгоритм перецептрона.

а) типовое индивидуальное домашнее задание:

№ 1. Даны два кластера: **X**1={(0,0),(0,1),(1,0),(1,1)}; **X**2={(5,0),(5,1),(6,0),(6,1),(6,3),(6,2)}. Найти решающие функции по критерию минимума расстояния. Найти уравнение разделяющей границы.

№ 2. Дан набор двумерных образов:

x 5,97 5,96 6,21 5,81 5,70 5,01 12,93 11,76 11,78 12,64 12,38 12,62 11,65 6,09 6,61 6,00 4,45 6,60 6,52 4,36 6,36

y 5,90 5,85 6,62 5,47 5,77 5,88 6,61 6,29 5,69 6,79 5,79 6,04 5,31 11,61 11,82 12,23 12,15 13,04 10,82 12,65 11,29

Используя алгоритм максиминного расстояния разбить данное множество на непересекающиеся подмножества.

№ 3. Дан набор двумерных образов:

x 5,97 5,96 6,21 5,81 5,70 5,01 12,93 11,76 11,78 12,64 12,38 12,62 11,65 6,09 6,61 6,00 4,45 6,60 6,52 4,36 6,36

y 5,90 5,85 6,62 5,47 5,77 5,88 6,61 6,29 5,69 6,79 5,79 6,04 5,31 11,61 11,82 12,23 12,15 13,04 10,82 12,65 11,29

Используя алгоритм К внутригрупповых средних разбить данное множество на непересекающиеся подмножества.

№ 4. Даны три кластера:

**X**1={(5,97;5,90),( 5,96; 5,85),( 6,21; 6,62),( 5,81; 5,47),( 5,70; 5,77),( 5,01; 5,88)};

**X**2={(12,93;6,61), ( 11,76;6,29), (11,78; 5,69), (12,64; 6,79), (12,38; 5,79),(12,62; 6,04), (11,65;5,31)}

**X**3={(6,09; 11,61),( 6,61; 11,82),(6,00;12,23),(4,45;12,15),(6,60;13,04),( 6,52; 10,82), (4,36; 12,65), (6,36; 11,29)}

Используя Байесовский классификатор, найти решающие функции и построить разделяющие границы.

№ 5. Даны три кластера:

**X**1={(5,97;5,90),( 5,96; 5,85),( 6,21; 6,62),( 5,81; 5,47),( 5,70; 5,77),( 5,01; 5,88)};

**X**2={(12,93;6,61), ( 11,76;6,29), (11,78; 5,69), (12,64; 6,79), (12,38; 5,79),(12,62; 6,04), (11,65;5,31)}

**X**3={(6,09; 11,61),( 6,61; 11,82),(6,00;12,23),(4,45;12,15),(6,60;13,04),( 6,52; 10,82), (4,36; 12,65), (6,36; 11,29)}.

Используя алгоритм подкрепления-наказания, найти решающие функции и построить разделяющие границы.

б) критерии оценивания компетенций (результатов) – правильность и полнота выполнения всех шагов решения задачи.

в) описание шкалы оценивания:

1, 4 и 5 задания оцениваются по шкале от 0 до 10 баллов. 2 и 3 задания оценивается по шкале от 0 до 15 баллов.

Индивидуальное домашнее задание считается выполненным успешно при отсутствии заданий с нулевой оценкой и при суммарной оценке не ниже 35 баллов.

*6.2.2. Лабораторные работы*

Не предусмотрены.

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

|  |  |  |
| --- | --- | --- |
| Форма аттестации | Наименование оценочного средства | Баллы |
| Зачет (100 баллов) | Индивидуальное домашнее задание (ИДЗ) | 60 |
| Ответы на зачете | 40 |

7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

а) основная учебная литература:

1. Дж. Ту, Р. Гонсалес. Принципы распознавания образов. Изд-во «Мир», Москва, 1978, -412с. (имеется в библиотеке ИАТЭ).

2. К. Фукунага. Введение в статистическую теорию распознавания образов. – М.: Наука, Главная редакция физико-математической литературы, 1979, -368с. (имеется в электронном виде).

б) дополнительная учебная литература:

1. Лепский А.Е., Броневич А.Г. Математические методы распознавания образов. Курс лекций. – Таганрог: Изд-во ТТИ ЮФУ, 2009. – 155с. (имеется в электронном виде).

2. Мазуров В.Д. Математические методы распознавания образов. Уч. Пособ. 2-е изд., доп. и перераб. – Екатеринбург: Изд-во Урал. ун-та, 2010. – 101с. (имеется в электронном виде).

8. Перечень ресурсов\* информационно-телекоммуникационной сети «Интернет» (далее - сеть «Интернет»), необходимых для освоения дисциплины

1. Ресурсы электронно-библиотечной системы издательства «Лань» // URL: [www.e.lanbook.com](http://www.e.lanbook.com) (по подписке).

9. Методические указания для обучающихся по освоению дисциплины

|  |  |
| --- | --- |
| Вид учебного занятия | Организация деятельности студента |
| 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. Обозначить вопросы, термины, материал, который вызывает трудности, пометить и попытаться найти ответ в рекомендуемой литературе. Если самостоятельно не удается разобраться в материале, необходимо сформулировать вопрос и задать преподавателю на консультации, на практическом занятии. |
| Практические занятия | Проработка рабочей программы, уделяя особое внимание целям и задачам, структуре и содержанию дисциплины. Работа с конспектом лекций, просмотр рекомендуемой литературы. Изучение выбранной предметной области на примерах решения задач семинарских занятий, индивидуальных домашних заданий. |
| Курсовая работа | Не предусмотрена |
| Индивидуальное домашнее задание | Ознакомиться с основной и дополнительной литературой, включая справочные издания, зарубежные источники, основополагающие термины. Попрактиковаться в решении аналогичных общих домашних задач по всем темам индивидуальных домашних заданий. |
| Лабораторная работа | Не предусмотрена. |
| Подготовка к экзамену | При подготовке к экзамену необходимо ориентироваться на конспекты лекций и рекомендуемую литературу. |

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем (при необходимости)

Издательская система LaTeX для подготовки докладов, презентаций и учебного материала.

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Видеопроектор, компьютер, издательская система LaTeX для подготовки докладов, презентаций и учебного материала.

12. Иные сведения и (или) материалы

12.1. Перечень образовательных технологий, используемых при осуществлении образовательного процесса по дисциплине

Часов в интерактивной форме – 8.

В ходе практических занятий происходит публичное обсуждение каждой решаемой задачи. При этом студенты высказывают свои мнения по выбору наиболее простого способа поиска оптимального решения.

После решения домашних работ на консультациях проводится разбор допущенных студентами ошибок.

12.2. **Формы организации самостоятельной работы обучающихся (темы, выносимые для самостоятельного изучения; вопросы для самоконтроля; типовые задания для самопроверки**

Некоторые темы изучаются студентами самостоятельно. Для изучения используется приведённая в списке основная и дополнительная литература. Контроль освоения материала осуществляется при приёме индивидуального домашнего задания.

| № | Тема и часть, изучаемая (осваиваемая) самостоятельно |
| --- | --- |
| 1.2 | Пространство образов и пространство весов. |
| 2.1 | Оценка результатов процесса кластеризации. |
| 2.2 | Алгоритм ИСОМАД. |
| 3.1 | Классификация образов по критерию минимума расстояния в случае множественности эталонов. |
| 3.2 | Вероятности ошибок в Байесовском классификаторе. |
| 3.3 | Статистический подход в построении обучаемых классификаторов. Алгоритм Роббинса-Монро. Алгоритм корректирующих приращений. Алгоритм НСКО. |

Вопросы и задания для самоконтроля по всем темам:

1. Что такое решающая функция?

2. Что такое область непринятия решения?

3. Как оценить качество кластеризации?

4. Каковы преимущества и недостатки алгоритма максиминного растояния?

5. Каковы преимущества и недостатки алгоритма алгоритма К внутригрупповых средних?

6. Каковы преимущества и недостатки алгоритма ИСОМАД?

7. Что такое Байесовский классификатор?

8. Какова целевая функция в градиентном методе подкрепления-наказания?

9. Какова целевая функция в градиентном методе НСКО?

10. Каковы преимущества и недостатки алгоритма подкрепления-наказания?

11. Каковы преимущества и недостатки алгоритма НСКО?

12. Какую задачу решает алгоритм Роббинса-Монро?

12.3. Краткий терминологический словарь

|  |  |
| --- | --- |
| Решающие функции по критерию минимума расстояния |  |
| Решающее правило | , если |
| Функция критерия в алгоритме перцептрона |  |