

«УТВЕРЖДАЮ»
Директор института цифровых технологий
и математики _____ С.А. Рощупкин/


РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

2.1.1.1 Машинное обучение и анализ больших данных

Шифр и наименование группы научных специальностей

5.12. Когнитивные науки

Шифр и наименование научной специальности

5.12.4. Когнитивное моделирование

Форма обучения: очная

Институт: математики, естествознания и техники

Кафедра: математического моделирования, компьютерных технологий и информационной безопасности

Трудоёмкость в ЗЕТ – 1

Трудоёмкость в часах – 36

Разработчик: к.т.н., Петров А.А.

Общие положения

Рабочая программа дисциплины «Физика полупроводников» разработана в соответствии с федеральными государственными требованиями, утвержденными приказом Министерства образования и науки высшего образования Российской Федерации от 20 октября 2021 г. № 951.

1. ЦЕЛИ И ЗАДАЧИ УЧЕБНОЙ ДИСЦИПЛИНЫ

Цель дисциплины:

Предоставление обучаемым знаний по современному состоянию исследований и разработок в области машинного обучения и анализа больших данных; проведение анализа существующих проблем «data mining», способов их решения и перспективных направлений развития; выделение основных тенденций в области эффективного использования ресурсов в IT-отрасли, направленных на защиту окружающей среды.

Задачи изучения дисциплины:

- дать обучающимся представление о современных проблемах искусственного интеллекта в разрезе когнитивного моделирования;
- изучение средств и интегрированных сред разработки приложений;
- освоение обучающимися концептуальных моделей предметной области когнитивного моделирования;
- освоение обучающимися методов обработки данных;
- дать представление о способах извлечения знаний из данных и языках представления знаний;
- ознакомление обучающихся с состоянием и перспективами развития технического обеспечения машинного обучения и анализа больших данных; с положениями синергетики об основах эволюции, законах и процессах самоорганизации систем;
- ознакомление обучающихся с положениями синергетики об основах эволюции, законами и процессами самоорганизации систем.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ПРОГРАММЫ АСПИРАНТУРЫ

Дисциплина относится к образовательному компоненту программы подготовки научных и научно-педагогических кадров в аспирантуре.

3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

По окончании изучения дисциплины аспиранты должны:

знать:

- методы исследования, основные области и задачи применения машинного обучения и анализа больших данных;
- современные информационные и телекоммуникационные технологии,

программные и аппаратные средства в науке о данных;

- актуальные тенденции, задачи и проблематику научных исследований в машинного обучения и анализа больших данных;

уметь:

- применять на практике методы машинного обучения и анализа больших данных;

- использовать информационные технологии и современные программно-аппаратные средства при решении задач машинного обучения;

- осуществлять внедрение результатов собственной научно-исследовательской деятельности в практику машинного обучения и анализа больших данных;

владеть:

- навыками применения технологий машинного обучения и анализа больших данных для решения профессиональных задач;

- навыками применения современных информационными технологиями в искусственном интеллекте;

- навыками внедрения результатов исследований в области машинного обучения и анализа больших данных в практику информационных технологий;

4. СТРУКТУРА, ОБЪЕМ И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1. Объем дисциплины и виды учебной работы

Трудоёмкость в ЗЕТ – 1

Трудоёмкость в часах – 36

Лекций – 18 ч.

Самостоятельная работа – 18 ч.

4.2. Разделы дисциплины и виды занятий

№	Наименование раздела дисциплины	Виды учебной нагрузки и их трудоемкость, академ. часы			
		Всего часов	Лекции	аудиторные занятия	
				практические (лабораторные)	Самостоятельная работа
1	Раздел 1. Вопросы машинного обучения	20	10		10
2	Тема 1. Введение в машинное обучение. Основные задачи машинного обучения.	8	4		4
3	Тема 2. Машинное обучение с учителем. Регрессия.	8	4		4

	Искусственные нейронные сети.				
4	Тема 3. Машинное обучение без учителя. Кластеризация. Поиск ассоциаций в данных.	4	2		2
5	<i>Раздел 2. Вопросы анализа больших данных</i>	16	8		8
6	Тема 4. Понятие «Big Data». Связь анализа данных с машинным обучением.	4	2		2
7	Тема 5. Корреляционный и регрессионный анализ. Снижение пространства размерности.	8	4		4
8	Тема 6. Классификация многомерных наблюдений. Анализ временных рядов.	4	2		2
ИТОГО:		36	18		18

5. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

5.1. Текущий контроль по дисциплине осуществляется в форме тестового задания и реферата

Типовой вариант тестового задания

1. Является ли распознавание образов направлением искусственного интеллекта?
 - + Да
 - Нет
2. Является ли генерация и распознавание речи направлением искусственного интеллекта?
 - + Да
 - Нет
3. Относятся ли экспертные системы к интеллектуальным информационным системам?
 - + Да
 - Нет
4. Относятся ли самообучающиеся системы к интеллектуальным информационным системам?
 - + Да
 - Нет
5. Относятся ли адаптивные системы к интеллектуальным информационным системам?
 - + Да
 - Нет
6. Какая разновидность экспертных систем основана на интеграции различных источников данных?
 - Классифицирующие
 - Трансформирующие
 - + Мультиагентные
7. Для решения каких задач предназначены экспертные системы?
 - + Неформализованных
 - Формализованных
 - + Стохастических

+ Детерминированных

9. База знаний включает в себя:

+ факты

+ понятия

+ правила

- механизм принятия решений

- подсказки-стимулы

- инструменты фиксации идей

- инструменты генерирования идей

- инструменты комбинирования идей

- пользовательский интерфейс

- нет правильного ответа

10. В отличие от базы данных, база знаний включает в себя:

+ правила принятия решений

- подсказки-стимулы

- инструменты фиксации идей

- инструменты генерирования идей

- инструменты комбинирования идей

- пользовательский интерфейс

- нет правильного ответа

11. Функция принадлежности к нечеткому множеству может принимать:

- любое положительное значение,

+ значение между нулем и единицей,

- любое отрицательное значение.

12. Основными критериями стратегии поиска являются:

- полнота,

- сложность и оптимальность,

+ все эти критерии.

13. Поиск в глубину – это:

- слепой поиск,

- направленный поиск,

+ итеративный поиск.

14. Двухнаправленный поиск – это:

- поиск в глубину,

- поиск в ширину,

+ прямой поиск от корневой вершины и обратный от целевой вершины

Примерная тематика рефератов

1. Способы представления и управления знаниями.

2. Генеральная и выборочные совокупности.

3. Модели и языки метаданных и онтологий.

4. Эволюционные вычисления.

5. Генетические методы и алгоритмы.

6. Синергетика – новое научное междисциплинарное направление.

7. Нейроинформатика и нейросистемы.
8. Геоинформатика и геоинформационные системы.
9. Современное состояние и перспективы развития телекоммуникационных систем и технологий.
10. Современное состояние и тенденции развития методов кодирования и сжатия данных.
11. Тенденции развития методов проектирования автоматизированных систем.
12. Тенденции развития интегрированных автоматизированных систем.
13. Облачные нейросетевые вычисления.
14. Современное состояние и перспективы развития графических процессоров для анализа данных и машинного обучения.
15. Современное состояние и тенденции развития квантовых процессоров.

5.2. Промежуточная аттестация по дисциплине проводится в 3 семестре в форме зачета

Вопросы к зачету

Задача машинного обучения

1. Формальная постановка задачи машинного обучения
2. Предобработка данных
3. Рекомендации к выбору признаков данных
4. Методы оценки качества машинного обучения
5. Проблема переобучения
6. Постановка задачи восстановления линейной регрессии
7. Метод градиентного спуска для обучения модели линейной регрессии
8. Переход от линейной регрессии к нелинейной регрессии
9. Линейные классификаторы
10. Связь линейных классификаторов с нейросетевыми моделями
11. Алгоритм обучения логистической регрессии
12. Байесовский подход к классификации
13. Алгоритм Парзена — Розенблатта
14. Использование энтропии в деревьях принятия решений
15. Построение дерева принятия решений
16. Классификация новых данных при помощи дерева принятия решений
17. Искусственные нейронные сети простого типа
18. Обучение искусственной нейронной сети простого типа (персептрона)
19. Многослойные искусственные нейронные сети
20. Функции активации для многослойных персептронов
21. Расчет ответа для многослойного персептрона на основе алгоритма прямого распространения
22. Обучение многослойного персептрона на основе алгоритма обратного распространения
23. Постановка задачи кластеризации
24. Оценка качества кластеризации

25. Алгоритм кластеризации k-средними
26. Постановка задачи иерархической кластеризации
27. Алгоритм иерархической кластеризации
28. Описание модели карты Кохонена
29. Алгоритм обучения карты Кохонена
30. Постановка задачи поиска ассоциаций в данных
31. Алгоритм Apriori
32. Понятие корреляционного анализа
33. Понятие регрессионного анализа
34. Классификация многомерных наблюдений
35. Предварительный анализ данных
36. Предмет и объект описательной статистики
37. Анализ временных рядов

Оценочные средства для проведения текущей и промежуточной аттестации по дисциплине

Критерии оценивания для зачета

Оценка «зачтено». Систематическое посещение занятий в течение учебного года - аспирант посетил более 75% аудиторных занятий. В процессе обучения показал заинтересованность в предмете.

Оценка «не зачтено». Пропущено значительное количество занятий без уважительной причины - аспирант посетил менее 75% аудиторных занятий. В процессе обучения не проявил интереса к предмету.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

6.1. Основная литература

1. *Платонов, А. В.* Машинное обучение : учебное пособие для вузов / А. В. Платонов. — Москва : Издательство Юрайт, 2022. — 85 с. — (Высшее образование). — ISBN 978-5-534-15561-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/508804> (дата обращения: 01.09.2022).

1. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2022. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/489100> (дата обращения: 01.09.2022).

6.2. Дополнительная литература

1. *Миркин, Б. Г.* Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2022. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/469306> (дата обращения: 01.09.2022).

6.3. Электронные образовательные ресурсы

1.	urait.ru	Юрайт Образовательная платформа	Регистрация через любой университетский компьютер. В дальнейшем предоставляется неограниченный индивидуальный доступ из любой точки, в которой имеется доступ к сети Интернет
2.	www.elibrary.ru	Российский информационный портал в области науки, технологии, медицины и образования	Свободный доступ

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы.