



Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Смирнов Сергей Николаевич
Должность: врио ректора
Дата подписания: 01.10.2024 10:55:26
Уникальный программный ключ:
69e375c64f7e975d4e8830e7b4fcc2ad1bf53108

Министерство науки и высшего образования Российской Федерации
ФГБОУ ВО «Тверской государственный университет»

Утверждаю:
Руководитель ООП

А.В. Солнышкин
«21» мая 2024 г.



Рабочая программа дисциплины (с аннотацией)
Программные средства обработки и анализа данных

Направление подготовки

03.04.02 Физика

профиль

Физика конденсированного состояния вещества

Для студентов

2 курса, очной формы обучения

Составитель: к.ф.-м.н. Романов А.А.

Тверь, 2024

Аннотация

1. Цель и задачи дисциплины

Целью освоения дисциплины является приобретение знаний об принципах объектно-ориентированного программирования и получение практических навыков программирования на языке высокого уровня, используемых при обработке и анализе данных.

Задачами освоения дисциплины являются:

1. формирование общепрофессиональных компетенций;
2. изучение и освоение основных методов, способов и средств обработки данных;
3. изучение и освоение объектно-ориентированного языка программирования на языке высокого уровня;
4. изучение и освоение инструментальных средств программирования и прикладных программ для решения инженерно-технических задач.
5. развитие логического и алгоритмического мышления студентов.

2. Место дисциплины в структуре ООП

Дисциплина «Программные средства обработки и анализа данных» относится к Блоку 1. Дисциплины обязательной части учебного плана.

Освоение дисциплины будет основой для выполнения научно-исследовательской работы, прохождения преддипломной практики и выполнения выпускной квалификационной работы.

Дисциплина базируется на знаниях курсов модуля «Информационные технологии» ООП бакалавриата.

3. Объем дисциплины: 3 зачетных единицы, 108 академических часов, **в том числе:**

контактная аудиторная работа: лабораторные работы 26 часов;

самостоятельная работа: 82 часа.

4. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы

<i>Планируемые результаты освоения образовательной программы (формируемые компетенции)</i>	<i>Планируемые результаты обучения по дисциплине</i>
ОПК-3. Способен применять знания в области информационных технологий, использовать современные компьютерные сети, программные продукты и ресурсы информационно-телекоммуникационной сети "Интернет" (далее - сеть "Интернет") для решения задач профессиональной деятельности, в том числе находящихся за пределами профильной подготовки;	ОПК-3.1. Использует информационные технологии для решения задач профессиональной деятельности; ОПК-3.2. Применяет специализированные программные продукты для обработки и анализа данных ОПК-3.3. Использует ресурсы информационно-телекоммуникационной сети "Интернет" для решения задач профессиональной деятельности, в том числе находящихся за пределами профильной подготовки.
ПК-2. Проводит работу по обработке и анализу научно-технической информации и результатов исследований.	ПК-2.3. Проводит анализ научных данных, результатов экспериментов и наблюдений;
ПК-3. Способен выполнять проектирование и разработку продукции в части, касающейся разработки объемных нанометаллов, сплавов и композитов на их основе, а также выбора расходных и вспомогательных материалов	ПК-3.3. Анализирует результаты испытаний образцов материалов.

5. Форма промежуточной аттестации и семестр прохождения

Зачет в 3 семестре.

6. Язык преподавания: русский.

II. Содержание дисциплины, структурированное по темам с указанием отведенного на них количества академических часов и видов учебных занятий.

1. Для студентов очной формы обучения

Учебная программа - наименование разделов и тем	Всего (час.)	Контактная работа (час.)	Самостоятельная работа, в том числе Контроль (час.)
		Лабораторные работы	
1. Регрессионный анализ. Линейная, полиномиальная и логарифмическая регрессия.	8	2	6
2 Классификация: Логистическая регрессия и SVM	8	2	6
3 Функции потерь и оптимизация	8	2	6
4. Оценка точности модели, переобучение, регуляризация	8	2	6
5. Проблема качества данных	8	2	6
6 Работа с пропусками	8	2	6
7 Работа с переменными	8	2	6
8. Деревья решений	8	2	6
9. Поиск выбросов и генерация новых признаков	8	2	6
10. Feature Selection	8	2	6
11. Машинное обучение. Введение в НС и библиотеку Keras	8	2	6
12. Углубление в НС и библиотеку Keras	8	2	6
13. Введение в сверточные НС	6	1	5
14. Введение в рекуррентные НС	6	1	5
ИТОГО	108	26	82

III. Образовательные технологии

Учебная программа-наименование разделов и тем	Вид занятия	Образовательные технологии
1. Регрессионный	<i>Лабораторные работы</i>	<i>Активное слушание.</i>

анализ. Линейная, полиномиальная и логарифмическая регрессия.		Решение индивидуальных задач Самостоятельное изучение теоретического материала
2 Классификация: Логистическая регрессия и SVM	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
3 Функции потерь и оптимизация	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
4. Оценка точности модели, переобучение, регуляризация	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
5. Проблема качества данных	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
6 Работа с пропусками	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
7 Работа с переменными	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
8. Деревья решений	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
9. Поиск выбросов и генерация новых признаков	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
10. Feature Selection	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала

11. Машинное обучение. Введение в НС и библиотеку Keras	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
12. Углубление в НС и библиотеку Keras	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
13. Введение в сверточные НС	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала
14. Введение в рекуррентные НС	Лабораторные работы	Активное слушание. Решение индивидуальных задач Самостоятельное изучение теоретического материала

IV. Оценочные материалы для проведения текущей и промежуточной аттестации

Форма проведения промежуточного контроля: студенты, освоившие программу курса «Программные средства обработки и анализа данных» могут получить зачет по итогам семестровой аттестации согласно «Положению о промежуточной аттестации (экзаменах и зачетах) обучающихся по программам высшего образования ТвГУ» (протокол №11 от 28 апреля 2021 г.).

ОПК-3. Способен применять знания в области информационных технологий, использовать современные компьютерные сети, программные продукты и ресурсы информационно-телекоммуникационной сети "Интернет" (далее - сеть "Интернет") для решения задач профессиональной деятельности, в том числе находящихся за пределами профильной подготовки:

ОПК-3.1. Использует информационные технологии для решения задач профессиональной деятельности;

ОПК-3.2. Применяет специализированные программные продукты для обработки и анализа данных

ОПК-3.3. Использует ресурсы информационно-телекоммуникационной сети "Интернет" для решения задач профессиональной деятельности, в том числе находящихся за пределами профильной подготовки.

Задание:

1. Проведение экспериментов с программной реализацией линейной регрессии. Исследование влияния параметров алгоритма на значение целевой функции на обучающей и тестовой выборке
2. Программная реализация рекуррентных НС.

Форма: устная

Критерии оценивания:

- Студент дает правильный развернутый и исчерпывающий ответ на вопрос. (5 баллов)
- Студент дает правильный ответ, который содержит отдельные неточности в определениях (3 балла)
- Ответ частично правильный (1 балл)

ПК-2. Проводит работу по обработке и анализу научно-технической информации и результатов исследований:

ПК-2.3. Проводит анализ научных данных, результатов экспериментов и наблюдений;

Задание:

1. Запишите выражение, нахождения функции потерь для логарифмической прогрессии.
2. Используя критерий Джини, найдите наилучшее разбиение.

Форма: письменный

Критерии оценивания:

- Студент умеет эффективно использовать теоретические знания для решения поставленной задачи (5 баллов)
- Получен правильный результат, однако ответ содержит отдельные неточности в определениях (3 балла)
- Результат частично правильный (1 балл)
- Неправильный ответ (0 баллов)

ПК-3. Способен выполнять проектирование и разработку продукции в части, касающейся разработки объемных нанометаллов, сплавов и композитов на их основе, а также выбора расходных и вспомогательных материалов:

ПК-3.3. Анализирует результаты испытаний образцов материалов.

Задание:

1. Приведите пример использования нейронных сетей.
2. Какой механизм применяется для стохастического градиентного спуска?

Форма: устный

Критерии оценивания:

- Студент дает правильный развернутый и исчерпывающий ответ на вопрос. (5 баллов)
- Студент дает правильный ответ, который содержит отдельные неточности в определениях (3 балла)
- Ответ частично правильный (1 балл)

V. Учебно-методическое и информационное обеспечение дисциплины

1) Рекомендуемая литература

а) Основная литература:

1. Чернышев С.А. Основы программирования на Python : Учебное пособие для вузов / Чернышев Станислав Андреевич; Чернышев С. А. - Электрон. дан. - Москва : Юрайт, 2021. - 286 с. - (Высшее образование). - URL: <https://urait.ru/bcode/477353>

2. Федоров Д.Ю. Программирование на языке высокого уровня Python : Учебное пособие для вузов / Федоров Дмитрий Юрьевич; Федоров Д. Ю. - 3-е изд. - Электрон. дан. - Москва : Юрайт, 2021. - 210 с. - (Высшее образование). - URL: <https://urait.ru/bcode/478098>

3. Григорьев А.А. Методы и алгоритмы обработки данных : Учебное пособие / Григорьев Анатолий Алексеевич, Исаев Евгений Анатольевич; Национальный исследовательский университет "Высшая школа экономики". - 2. - Москва : ООО "Научно-издательский центр ИНФРА-М", 2021. - 383 с. <http://znanium.com/catalog/document?id=361208>

4. Титов, А. Н. Python. Обработка данных : учебно-методическое пособие / А. Н. Титов, Р. Ф. Тазиева. — Казань : КНИТУ, 2022. — 104 с. — ISBN 978-5-7882-3171-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/331019>

5. Сандуляк, Д. А. Анализ и обработка экспериментальных данных : учебно-методическое пособие / Д. А. Сандуляк, М. Н. Полисмакова. — Москва : РТУ МИРЭА, 2023. — 64 с. — ISBN 978-5-7339-1902-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/382541>

6. Методы обработки экспериментальных данных : учебное пособие / С. А. Гордин, А. А. Соснин, И. В. Зайченко, В. Д. Бердоносков. — Комсомольск-на-Амуре : КНАГУ, 2022. — 75 с. — ISBN 978-5-7765-1501-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/310745>

б) Дополнительная литература

1. Информационные аналитические системы [Электронный ресурс] : учебник / Т. В. Алексеева, Ю. В. Амириди, В. В. Дик и др.; под ред. В. В. Дика. - М.: МФПУ Синергия, 2013. - 384 с. Режим доступа: <http://znanium.com/bookread2.php?book=451186>

2. Куприянов М. С., Барсегян, А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP [Электронный ресурс] / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко. - 2-е изд., перераб. и доп. - СПб.: БХВ-Петербург, 2007. - 384 с. Режим доступа:<http://znanium.com/bookread2.php?book=489445>

2) Программное обеспечение

Kaspersky Endpoint Security 10 для Windows

Adobe Acrobat Reader

Google Chrome

OpenOffice

Mozilla Firefox

Mathcad 15 M010

MATLAB R2012b

Foxit Reader

WinDjView

ABBYY Lingvo x5

Notepad++

Cadence SPB/OrCAD

Python

Lego MINDSTORM EV3

Origin 8.1 Sr2

Многофункциональный редактор ONLYOFFICE

ОС Linux Ubuntu

VLC media player

Java SE Development Kit 8 Update 45 (64-bit)

Lazarus 1.4.0

LEGO MINDSTORMS EV3 Microsoft Expression Studio 4

MiKTeX 2.9

MSXML 4.0 SP2 Parser and SDK

R Studio

3) Современные профессиональные базы данных и информационные справочные системы

ЭБС«ZNANIUM.COM» www.znanium.com;

ЭБС «Университетская библиотека онлайн»<https://biblioclub.ru/>;

ЭБС «Лань» <http://e.lanbook.com>

Научная электронная библиотека eLibrary.ru;

Электронная база данных диссертаций РГБ;

База данных Реферативных журналов ВИНТИ;

Полнотекстовый доступ к журналам AIP (Американский институт физики);

Полнотекстовый доступ к журналам и книгам издательства Springer Verlag;

Полнотекстовый доступ к отдельным журналам и книгам Института инженеров по электротехнике и электронике (Institute of Electrical and Electronics Engineers);

Полнотекстовый доступ к отдельным журналам Optical Society of America;

Полнотекстовый доступ к журналам ACS (Американского химического общества);

Реферативная база Inspec (доступ к рефератам и полным текстам монографий и научных статей в области физики, электротехники, электроники, коммуникаций, компьютерных наук и информационных технологий);

Коллекция электронных книг Оксфордско-Российского фонда;

Корпоративный каталог «КОРБИС (Тверь и партнеры)».

4) Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

Научная библиотека ТвГУ – <http://library.tversu.ru>;

Сервер доступа к модульной объектно-ориентированной динамической учебной среде Moodle – <http://moodle.tversu.ru>;

Сервер обеспечения дистанционного обучения и проведения Web-конференций Mirapolis Virtual Room – <http://mvr.tversu.ru>;

Репозиторий научных публикаций ТвГУ – <http://eprints.tversu.ru>.

Видео-лекции курса "Машинное обучение" от Yandex
<https://yandexdataschool.ru/edu-process/courses/machine-learning>

Научный портал по математическим наукам - <http://www.mathnet.ru>

Портал образовательных ресурсов по ИТ - <http://www.intuit.ru>

Портал статей по применению ИТ и машинному обучению
http://habrahabr.ru/hub/machine_learning/

Профессиональный интернет-ресурс по машинному обучению -
<http://www.machinelearning.ru/>

VI. Методические материалы для обучающихся по освоению дисциплины

- Вопросы для самоподготовки

1. Сущность и этапы анализа данных с помощью методов машинного обучения
2. Классификация методов машинного обучения
3. Основные проблемы машинного обучения
4. Задачи машинного обучения с учителем
5. Задачи машинного обучения без учителя,
6. Задачи машинного обучения с подкреплением.
7. Библиотеки машинного обучения языка Python.
8. Метод регрессии
9. Деревья решений, алгоритм случайного леса.
10. Метод k-ближайших соседей
11. Логистическая регрессия

12. Бинарная линейная классификация

13. Нейронные сети, их обучение методом обратного распространения ошибки.

- Контрольные вопросы и задания

1. Что такое машинное обучение?
2. Чем алгоритмы машинного обучения отличаются от традиционных алгоритмов?
3. Формальная постановка задачи машинного обучения.
4. Объекты и признаки.
5. Основные типы признаков.
6. Что такое обучение с учителем? Приведите примеры задач.
7. Опишите задачу регрессии.
8. Приведите примеры задач регрессии.
9. Опишите задачу классификации.
10. Приведите примеры задач классификации.
11. Как оценить качество модели?
12. Что такое недообучение и переобучение и как их избежать?
13. На какие части и каким образом делится выборка?
14. Что такое обучение без учителя? Приведите примеры задач.
15. Опишите задачу кластеризации.
16. Приведите примеры задач кластеризации.
17. Что такое "близость" в задаче кластеризации?
18. Какие метрики можно использовать в задаче кластеризации?
19. Как оценить качество кластеризации?
20. Опишите структуру и свойства искусственного нейрона.
21. Что называется функцией активации? Приведите примеры.
22. Как организуются слои?
23. Опишите архитектуру многослойного перцептрона.
24. Что значит "обучить" нейронную сеть?

25. Проблемы обучения и тестирования.
26. Как оценить качество обучения нейронной сети?
27. Что такое сверточные нейронные сети?
28. Что такое рекуррентные нейронные сети?
29. Какие задачи можно решать с помощью нейронных сетей?
30. Какие проблемы в области медицины и здравоохранения можно решать с помощью методов машинного обучения?
31. Сформулируйте задачу постановки диагноза как задачу машинного обучения.
32. Опишите процесс обучения нейронной сети.
33. Как нужно подготовить данные для обучения и тестирования?
34. Что такое метод обратного распространения ошибки? Почему он так называется?
35. Что такое параметры и гиперпараметры? На что они влияют?
36. Как подбирать параметры и гиперпараметры?
37. Зачем нужна нормализация?
38. Опишите способы нормализации.
39. Как можно оценить качество работы нейронной сети?
40. Что такое Keras и TensorFlow. Какие возможности они предоставляют?

VII. Материально-техническое обеспечение

<p>Учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, Лекционная аудитория № 226 (170002 Тверская обл., г. Тверь, Садовый пер., д. 35)</p>	<p>1 Микшерный пульт Yamaha MG-124C 2 Аудиокомплект (мик. пульт, акуст. усилитель, акуст. система, радиосистема) 3 Интерактивная система SMART Board 660i4 4 Мультимедийный проектор Epson EB-4850WU с потолочным креплением 5 Телекоммуникационный шкаф ШТК-М-18.6.6-3ААА с полками 6 Телекоммуникационный шкаф</p>	<p>Kaspersky Endpoint Security 10 для Windows Adobe Acrobat Reader Google Chrome Notepad++ Многофункциональный редактор ONLYOFFICE OpenOffice</p>
--	--	---

	<p>ШТК-М-18.6.6-3AAA с полками 7 Экран настенный Lumien 8 Компьютер iRU Corp 510 15-2400/4096/500/G210-512/DVD-RW/W7S/монитор E-Machines E220HQVB 21,5” 9 Компьютер Ramec Storm Custom W CPU -Intel Core i3-540(3,06ГГц)/2*1024Мб/DVD-RW/400W/клав./мышь/ковр./Win 7 St/Монитор 20" ACER V203HV 10 Комплект учебной мебели на 110 посадочных мест 11 Меловая доска 12. Бактерицидный облучатель-рециркулятор настенный "Мегидез" РБОВ 911-"МСК", 00-000000000002868</p>	
<p>Помещение для самостоятельной работы, учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, практики, Компьютерный класс физико-технического факультета. Компьютерная лаборатория робототехнических систем № 4а (170002 Тверская обл., г. Тверь, Садовый пер., д. 35)</p>	<p>1 Компьютер HP 260 G2 Intel Core i3 6100U, 4 ГБ RAM, 256 Гб SSD, Intel HD Graphics 520, DOS, клавиатура, мышь+ ЖК монитор 21.5" BenQ GL2250 (12 шт) 2 Мультимедийный комплект учебного класса (вариант № 2) Проектор Casio XJ-M140, настенный проекц. экран Lumien 180*180. ноутбук Dell N4050. сумка 15,6",мышь 3 Коммутатор D-Link 10/100/1000mbps 16-portr DGS-1016D 4 Видеокамера IP-FALCON EYE FE-IPC-BL200P., ОнЛайн Трейд ООО 5 Видеокамера IP-FALCON EYE FE-IPC-BL200P., ОнЛайн Трейд ООО 6. Демонстрационное оборудование комплект «LegoMidstormsEV3» 7. Комплект учебной мебели</p>	<p>Kaspersky Endpoint Security 10 для Windows Adobe Acrobat Reader Google Chrome OpenOffice Mozilla Firefox Mathcad 15 M010 MATLAB R2012b Foxit Reader WinDjView ABBYY Lingvo x5 Notepad++ Cadence SPB/OrCAD Python Lego MINDSTORM EV3 Origin 8.1 Sr2 Многофункциональный редактор ONLYOFFICE OC Linux Ubuntu VLC media player Java SE Development Kit 8 Update 45 (64-bit) Lazarus 1.4.0 LEGO MINDSTORMS EV3 Microsoft Expression Studio 4 MiKTeX 2.9 MSXML 4.0 SP2 Parser and SDK R Studio</p>

VIII. Сведения об обновлении рабочей программы дисциплины

№ п.п.	Обновленный раздел рабочей программы дисциплины	Описание внесенных изменений	Реквизиты документа, утвердившего изменения
1.			
2.			