

Министерство просвещения Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Российский государственный профессионально-педагогический университет»
Институт инженерно-педагогического образования
Кафедра информационных систем и технологий

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Б1.В.ДВ.01.03.0 «МАТЕМАТИЧЕСКИЕ ОСНОВЫ МАШИННОГО
ОБУЧЕНИЯ»**

Направление подготовки 09.03.03 Прикладная информатика

Профиль программы «Прикладная информатика (по элективным модулям)»

Автор(ы): канд. пед. наук, доцент, И.А. Сулова
заведующий кафедрой

Одобрена на заседании кафедры информационных систем и технологий. Протокол от «20» января 2022 г. №5.

Рекомендована к использованию в образовательной деятельности научно-методической комиссией института ИПО РГППУ. Протокол от «26» января 2022 г. №6.

Екатеринбург
2022

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель освоения дисциплины «Математические основы машинного обучения»: формирование у будущих специалистов теоретических знаний и практических умений по многомерным статистическим методам и методам машинного обучения.

Задачи:

- освоение основ интеллектуального анализа данных, включая преобразование и очистку данных, работу с пропущенными значениями, основные способы визуализации данных (гистограммами, диаграммами плотности, диаграммами рассеяния, ящиками с усами и т.п.), корреляционный анализ;
- освоение различных методов отбора признаков;
- решение различных задач снижения размерности данных, кластеризации, классификации, регрессии.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «Математические основы машинного обучения» относится к формируемой участниками образовательных отношений части учебного плана.

Для изучения учебной дисциплины необходимы знания, умения и владения, формируемые следующими дисциплинами:

1. Прикладная математика и математическая логика.
2. Введение в машинное обучение.
3. Математическая статистика.
4. Большие данные и аналитика данных. Анализ и разработка алгоритмов.

Перечень учебных дисциплин, для которых необходимы знания, умения и владения, формируемые данной учебной дисциплиной:

1. Нейронные сети.
2. Прикладные модели машинного обучения.
3. Поиск, обработка и распознавание графической, аудио и видеoinформации.

3. РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина направлена на формирование следующих компетенций:

- ПКС-13 Способен к обеспечению процесса организации оценки соответствия требованиям существующих и (или) аналогичных программных систем;



- ПКС-14 Способен к представлению заинтересованным лицам концепции программной системы, технического задания и изменений в них.

В результате освоения дисциплины (модуля) обучающийся должен:

Знать:

31. Статистическую теорию машинного обучения, методы классификации и регрессии с опорными векторами, теорию обобщения Вапника-Червоненкиса и алгоритмы построения разделяющих гиперплоскостей;

32. Методы теории адаптивного прогнозирования в режиме онлайн в теоретико-игровой и сравнительной постановке: игры с рандомизированными предсказаниями, предсказания с использованием экспертных стратегий;

33. Основные понятия теории игр.

Уметь:

У1. Применять основные математические методы и алгоритмы теории машинного обучения;

У2. Применять методы теории адаптивного прогнозирования в режиме онлайн.

Владеть:

В1. Навыком освоения большого объема информации;

В2. Навыками постановки научно-исследовательских задач.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

4.1 Объем дисциплины и виды контактной и самостоятельной работы

Общая трудоёмкость дисциплины составляет 8 зач. ед. (288 час.), семестры изучения – 5, 6, распределение по видам работ представлено в табл. № 1.

Таблица 1. Распределение трудоёмкости дисциплины по видам работ

Вид работы	Форма обучения
	очная
	Семестр изучения
	5, 6 сем.
	Кол-во часов
Общая трудоёмкость дисциплины по учебному плану	288
Контактная работа, в том числе:	100
Лекции	32
Лабораторные работы	68
Самостоятельная работа студента	188
Промежуточная аттестация, в том числе:	



Зачет	5 сем.
Экзамен	6 сем.
Курсовая работа	6 сем.

**Распределение трудоемкости по видам контактной работы для заочной формы обучения (при наличии) корректируется в соответствии с учебным планом заочной формы обучения.*

4.2 Содержание и тематическое планирование дисциплины

Таблица 2. Тематический план дисциплины

Наименование разделов и тем дисциплины (модуля)	Сем.	Всего, час.	Вид контактной работы, час.			СРС
			Лекции	Практ. занятия	Лаб. работы	
1. Элементы теории классификации и регрессии с опорными векторами	5, 6	66	8	-	12	46
2. Универсальные предсказания	5, 6	70	8	-	16	46
3. Элементы сравнительной теории машинного обучения	5, 6	74	8	-	20	46
4. Элементы теории игр	5, 6	78	8	-	20	50

**Распределение часов по разделам (темам) дисциплины для заочной формы обучения осуществляется научно-педагогическим работником, ведущим дисциплину.*

4.3 Содержание разделов (тем) дисциплин

Раздел 1. Элементы теории классификации и регрессии с опорными векторами

Постановка задачи классификации. Байесовский классификатор. Линейные классификаторы: перцептрон. Алгоритм Розенблатта. Теорема Новикова о сходимости.

Теория обобщения Вапника-Червоненкиса. Верхняя оценка вероятности ошибки классификации через VC-размерность класса функций классификации.

VC-размерность, определение, основное свойство. VC-размерность класса всех линейных (однородных) классификаторов.

Метод опорных векторов. Оптимальная гиперплоскость. Алгоритм построения оптимальной гиперплоскости. Оценка вероятности ошибки обобщения через число опорных векторов.

SVM - метод в пространстве признаков, примеры. Ядра.



Случай неотделимой выборки. Вектор переменных мягкого отступа. Оценка вероятности ошибки обобщения. Оптимизационная задача для классификации с ошибками в квадратичной норме.

Случай неотделимой выборки. Вектор переменных мягкого отступа. Оценка вероятности ошибки обобщения. Оптимизационная задача для классификации с ошибками в линейной норме. Оптимизационная задача для классификации с ошибками в форме задачи линейного программирования.

Задача многомерной регрессии. Простая линейная регрессия. Гребневая регрессия.

Задача многомерной регрессии. Регрессия с опорными векторами. Ошибка обобщения при регрессии. Решение задачи гребневой регрессии с помощью SVM.

Гребневая регрессия в прямой форме и в двойственной форме как частный случай регрессии с опорными векторами в случае квадратичной функции потерь. Нелинейная многомерная гребневая регрессия (с ядром).

Раздел 2. Универсальные предсказания

Задача универсального прогнозирования в режиме он-лайн: статистический подход.

Калибруемость прогнозов. Алгоритм вычисления хорошо калибруемых прогнозов.

Прогнозирование с произвольным ядром.

Раздел 3. Элементы сравнительной теории машинного обучения

Алгоритм оптимального распределения потерь в режиме он-лайн.

Задача нахождения оптимальных решений с учетом экспертных стратегий. Алгоритм экспоненциального взвешивания экспертных решений.

Рандомизированные прогнозы.

Усиление простых классификаторов – Boosting. Алгоритм AdaBoost.

Раздел 4. Элементы теории игр

Антагонистические игры двух игроков. Достаточное условие существования седловой точки.

Достаточное условие существования седловой точки. Смешанные расширения матричных игр. Минимаксная теорема.

Чистые стратегии. Решение матричной игры типа $2 \times M$. Решение игры типа $N \times M$.

Бесконечные игры с рандомизированными предсказаниями. Построение выигрышной стратегии предсказателя с использованием минимаксной теоремы.

Хорошо калибруемые предсказания. Универсальная калибруемость со счетным числом правил выбора.



5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

Для изучения дисциплины используются различные образовательные технологии:

1. Информационно-коммуникационные образовательные технологии, при которых организация образовательного процесса, основывается на применении специализированных программных сред и технических средств работы с информацией. Используются для поддержки самостоятельной работы обучающихся с использованием электронной информационно-образовательной среды (ЭИОС), телекоммуникационных технологий, педагогических программных средств и др.

2. Для организации процесса обучения и самостоятельной работы используются информационно-коммуникационные образовательные технологии, представленные в виде педагогических программных средств и электронной информационно-образовательной среды (ЭИОС). Технологии расширяют возможности образовательной среды, как разнообразными программными средствами, так и методами развития креативности обучаемых. К числу таких программных средств относятся моделирующие программы, поисковые, интеллектуальные обучающие, экспертные системы, программы для проведения деловых игр.

3. Проведение лабораторных или практических работ направлено на формирование практических навыков и умений в области решения задач прикладного характера, способствует усилению мотивации к приобретению профессионально значимых навыков за счёт погружения в квазипрофессиональную проектную деятельность, позволяет сконцентрировать внимание обучающегося на совокупности полученных ранее теоретических знаний и отследить их практико-ориентированный характер.

В процессе выполнения лабораторных или практических работ обучающиеся получают первичное знакомство с элементами будущей профессиональной деятельности, формируют представление о принципах практической реализации полученных теоретических сведений.

4. При реализации образовательной программы с применением дистанционных образовательных технологий и электронного обучения:

- состав видов контактной работы по дисциплине (модулю), при необходимости, может быть откорректирован в направлении снижения доли занятий лекционного типа и соответствующего увеличения доли консультаций (групповых или индивидуальных) или иных видов контактной работы;

- информационной основой проведения учебных занятий, а также организации самостоятельной работы обучающихся по дисциплине (модулю) являются представленные в электронном виде методические, оценочные и иные материалы, размещенные в электронной информационно-образовательной среде (ЭИОС) университета, в электронных библиотечных системах и открытых Интернет-ресурсах;



- взаимодействие обучающихся и педагогических работников осуществляется с применением ЭИОС университета и других информационно-коммуникационных технологий (видеоконференцсвязь, облачные технологии и сервисы, др.);

- соотношение контактной и самостоятельной работы по дисциплине (модулю) может быть изменено в сторону увеличения последней, в том числе самостоятельного изучения теоретического материала.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ

6.1 Основная литература

1. Кук, Д. Машинное обучение с использованием библиотеки H2O. Мощные и масштабируемые методы для глубокого обучения и ИИ [Электронный ресурс] 1 / Даррен Кук. - Электрон. текстовые дан. - Москва : ДМК Пресс, 2018. - 248 с. : ил., табл. - Режим доступа: <https://e.lanbook.com/reader/book/97353/#1>.

2. Уэс Маккинли Python и анализ данных : практическое пособие. - Саратов : Профобразование, 2017. - 482 с. - Режим доступа: <http://www.iprbookshop.ru/64058>.

6.2 Дополнительная литература

1. Широков, А. И. Информатика: разработка программ на языке программирования Питон: базовые языковые конструкции : учебник / А. И. Широков, М. О. Пышняк. — Москва : Издательский Дом МИСиС, 2020. — 142 с. — ISBN 978-5-907226-76-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/106713.html>

2. Борзунов С. В. Алгебра и геометрия с примерами на Python : учебное пособие для вузов / Борзунов С. В., Кургалин С. Д. — Санкт-Петербург : Лань, 2020. — 444 с. — Режим доступа: <https://e.lanbook.com/book/149336>.

3. Амоа, К. А. Разработка программных пакетов на языке Python : учебное пособие / К. А. Амоа, Н. А. Рындин, Ю. С. Скворцов. — Воронеж : Воронежский государственный технический университет, ЭБС АСВ, 2020. — 61 с. — ISBN 978-5-7731-0887-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru>

6.3 Программное обеспечение и Интернет-ресурсы

Интернет-ресурсы:

1. Банк программ подготовки рабочих кадров и специалистов, реализуемых в созданных многофункциональных центрах прикладных квалификаций. Режим доступа: <http://mcpk.ntf.ru/>

2. Сайт движения WorldSkillsRussia. Режим доступа: <http://worldskillsrussia.org/>

3. Яндекс Практикум. Режим доступа: <https://praktikum.yandex.ru/>



Программное обеспечение:

1. Операционная система Windows.
2. Офисная система Office Professional Plus.

Информационные системы и платформы:

1. Система дистанционного обучения «Moodle».
2. Информационная система «Таймлайн».
3. Платформа для организации и проведения вебинаров «Mirapolis Virtual Room».

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Перечень материально-технического обеспечения для реализации образовательного процесса по дисциплине:

1. Лаборатория "Робототехнические системы" технологий в области электроники, мехатроники, робототехники, программирования и схемотехники.
2. Компьютерный класс.
3. Учебная аудитория для проведения занятий лекционного типа с мультимедийным оборудованием.
4. Помещения для самостоятельной работы.

