

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 «ПРИКЛАДНЫЕ МОДЕЛИ 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 «Прикладные модели машинного обучения»: изучение основ теории обучения машин, включая дискриминантный, кластерный и регрессионный анализ, овладение приемами практического решения задач интеллектуального анализа данных.

Задачи:

- сформировать теоретические знания по основам машинного обучения для построения формальных математических моделей и интерпретации результатов моделирования;
- выработать умения по практическому применению методов машинного обучения для построения формальных математических моделей и интерпретации результатов моделирования при решении прикладных задач в различных прикладных областях;
- выработать умения использования различных программных инструментов анализа баз данных и систем машинного обучения.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «Прикладные модели машинного обучения» относится к формируемой участниками образовательных отношений части учебного плана.

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

1. Математические основы машинного обучения.
2. Основы научно-исследовательской работы.
3. Экспертные системы.

3. РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина направлена на формирование следующих компетенций:

- ПКС-11 Способен выполнять анализ проблемной ситуации заинтересованных лиц;
- ПКС-13 Способен к обеспечению процесса организации оценки соответствия требованиям существующих и (или) аналогичных программных систем.

В результате освоения дисциплины (модуля) обучающийся должен:

Знать:

31. Теоретические основы многомерных статистических методов;



32. Теоретические основы моделей машинного обучения для решения различных задач анализа данных и построения предсказательных моделей.

Уметь:

У1. Выбирать признаки объектов для решения различных задач анализа данных и построения предсказательных моделей;

У2. Подбирать семейства моделей машинного обучения для конкретных задач, производить их обучение, гиперпараметрическую оптимизацию, контроль качества обучения.

Владеть:

В1. Навыками написания приложений, использующих различные модели машинного обучения для решения прикладных задач анализа данных.

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 – 7, 8, распределение по видам работ представлено в табл. № 1.

Таблица 1. Распределение трудоёмкости дисциплины по видам работ

Вид работы	Форма обучения
	очная
	Семестр изучения
	7, 8 сем.
	Кол-во часов
Общая трудоёмкость дисциплины по учебному плану	288
Контактная работа, в том числе:	92
Лекции	24
Лабораторные работы	68
Самостоятельная работа студента	196
Промежуточная аттестация, в том числе:	
Зачет с оценкой	8 сем.
Экзамен	7 сем.
Курсовая работа	8 сем.

**Распределение трудоёмкости 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. Введение в машинное обучение и анализ многомерных данных	7, 8	76	8	-	22	46
2. Методы обучения с учителем	7, 8	100	8	-	24	68
3. Методы обучения без учителя	7, 8	112	8	-	22	82

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

4.3 Содержание разделов (тем) дисциплин

Раздел 1. Введение в машинное обучение и анализ многомерных данных

Понятие машинного обучения. Соотнесение областей искусственного интеллекта, машинного обучения, нейронных сетей и глубокого обучения. Задача обучения по прецедентам. Признаковое описание объектов. Задание ответов. Типы задач машинного обучения. Предсказательная модель. Примеры задач классификации и регрессии. Этапы обучения и применения моделей машинного обучения. Функционалы качества. Сведение задачи обучения к задаче оптимизации. Объекты, признаки и особенности задач медицинской диагностики, кредитного скоринга, предсказания оттока клиентов, категоризации текстовых документов, прогнозирования стоимости недвижимости, прогнозирования объемов продаж, предсказания прибыли ресторана, ранжирования поисковой выдачи, ранжирования в рекомендательных системах, предсказания перехода по контекстной рекламе.

Исследовательский анализ данных с помощью библиотеки Pandas. Тип DataFrame. Загрузка данных, получение информации о первых или последних пяти записях, колонках, типах. Описательная статистика. Операции доступа к отдельным колонкам, строчкам, элементам DataFrame, их группам. Получение уникальных значений в конкретных колонках, подсчет частот уникальных значений в конкретных колонках. Сортировка и фильтрация данных, применение функций к отдельным ячейкам, колонкам и строчкам. Маппирование, изменение типа колонок данных. Добавление, удаление, замена колонок данных. Группировка данных. Получение SummaryTables.

Раздел 2. Методы обучения с учителем



Понятие дерева решений. Их использование для решений задач классификации и регрессии. Понятие прироста информации. Метрики качества разбиения в задачах классификации и регрессии. Процедуры построения деревьев классификации и регрессии. Понятие деревьев решений. Использование стандартных классов `DecisionTreeClassifier` и `DecisionTreeRegressor` из библиотеки `scikit-learn`. Визуализация построенных деревьев решений. Проблема переобучения деревьев решений, подбор гиперпараметров. Достоинства и недостатки деревьев решений.

Раздел 3. Методы обучения без учителя

Понятие обучения без учителя. Метод главных компонент. Методы кластеризации: `k-means`, `affinity propagation`, спектральная кластеризация, агломеративная кластеризация. Метрики качества кластеризации. Основные возможности библиотеки `scikit-learn` для реализации данных методов.

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. 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>



2. Сузи Р.А. Язык программирования Python [Электронный ресурс] / Р.А. Сузи. — Электрон. текстовые данные. — М. : Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 350 с. — 5-9556-0058-2. — Режим доступа: <http://www.iprbookshop.ru/52211.html>.— ЭБС «IPRbooks»

3. 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>.

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.

