

Компонент ОПОП 09.04.01 Информатика и вычислительная техника
Направленность (профиль) Компьютерный анализ и интерпретация данных.
Data Science
Б1.О.15
шифр дисциплины

РАБОЧАЯ ПРОГРАММА

Дисциплины Машинное обучение

Разработчик (и):
Шиманский С.А.
доцент

Утверждено на заседании кафедры
Цифровых технологий, математики
и экономики
наименование кафедры

протокол № 8 от 23.05.2023

И.о. заведующего кафедрой
ЦТМиЭ

_____ Мотина Т.Н.
подпись _____ ФИО

Мурманск
2023

Пояснительная записка

Объем дисциплины 4 з. е.

1. Результаты обучения по дисциплине, соотнесенные с индикаторами достижения компетенций, установленными образовательной программой

Компетенции	Индикаторы достижения компетенций	Результаты обучения по дисциплине
ОПК-2 Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	ИД-1 _{ОПК-2} Понимает принципы и применяет методы разработки алгоритмов и программных средств ИД-2 _{ОПК-2} Использует современные интеллектуальные технологии при разработке алгоритмов и программных средств ИД-2 _{ОПК-3} Решает профессиональные задачи путем разработки оригинальных алгоритмов и программных средств	Знать: приемы сбора и обработки информации для задач профессиональной деятельности в области машинного обучения; математические и вычислительные (компьютерные) основы МО; принципы работы современных информационных технологий и программных средств, применяемых в задачах МО. Уметь: осуществлять поиск, сбор, обработку, критический анализ и синтез информации, полученной из разных источников при решении задач в области МО; решать стандартные профессиональные задачи в области МО и ИНС с применением естественнонаучных и инженерных знаний, методов математического анализа и моделирования, информационных технологий.
ОПК-6 Способен разрабатывать компоненты программно-аппаратных комплексов обработки информации и автоматизированного проектирования	ИД-3 _{ОПК-6} Использует инструментальные средства разработки компонент программно-аппаратных комплексов	Владеть: навыком поиска, сбора, обработки, критического анализа и синтеза информации, методикой системного подхода для решения задач в области МО; навыками применения современных информационных технологий и программных средств при решении задач профессиональной деятельности в области МО.

2. Содержание дисциплины

Тема 1. *Введение в предмет.* Понятие машинного обучения (МО). История. Основные термины и определения. Решаемые задачи и их классификация. Практические примеры. Связь функционального анализа и машинного обучения.

Тема 2. *Оптимизационные модели МО.* Градиентные методы. Метрические пространства и МО. Оптимизация функционалов. Постановка оптимизационных задач. Нормирование метрики и метод имитации отжига.

Тема 3. *Задача автоматической классификации.* Метрические алгоритмы МО. Обучение без учителя, нормированные метрические пространства и обучение сетей Кохонена. Линейные топологические пространства. Иерархическая классификация. Самоорганизующиеся карты. Банаховы пространства.

Тема 4. *Основные нелинейные модели МО.* Линейные метрические пространства и метод роения частиц. Гильбертовы пространства и генетические алгоритмы. Пространства Лебега и генетические алгоритмы.

Вероятностное метрическое пространство и МО. Прогнозирование временных рядов. Байесовский классификатор. Метод опорных векторов. Метод главных компонент.

Многомерная регрессия, обобщенная регрессионная модель. Кластеризация.

Топологические пространства и обучение с учителем.

Линейные операторы, функционалы и обучение с подкреплением. Задача понижения размерности. Жадные алгоритмы и машинное обучение. SARSA. Q-learning.

Тема 5. *Композиция моделей МО*. Композиция моделей (ансамблевые методы – стекинг, бэггинг, бустинг). Ранжирование и рекомендательные системы. Современные тренды в области МО: AutoML. Некорректные задачи.

3. Перечень учебно-методического обеспечения дисциплины

Методические материалы для обучающихся по освоению дисциплины представлены на официальном сайте МАУ в разделе «Информация по образовательным программам, в том числе адаптированным».

4. Фонд оценочных средств по дисциплине

Является отдельным компонентом образовательной программы, разработан в форме отдельного документа, включает в себя:

- перечень компетенций с указанием этапов их формирования в процессе освоения дисциплины;
- задания текущего контроля;
- задания промежуточной аттестации;
- задания внутренней оценки качества образования.

5. Перечень основной и дополнительной учебной литературы (печатные издания, электронные учебные издания и (или) ресурсы электронно-библиотечных систем)

Основная литература:

1. Матренин, П. В. Методы стохастической оптимизации : учебное пособие : [16+] / П. В. Матренин, М. Г. Гриф, В. Г. Секаев ; Новосибирский государственный технический университет. – Новосибирск : Новосибирский государственный технический университет, 2016. – 67 с. : ил., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=576420>. – Библиогр. в кн. – ISBN 978-5-7782-2861-0.
2. Гультаева, Т. А. Методы статистического обучения в задачах регрессии и классификации / Т. А. Гультаева, А. А. Попов, А. С. Саутин ; Новосибирский государственный технический университет. – Новосибирск : Новосибирский государственный технический университет, 2015. – 323 с. : ил., табл., граф. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=576264>. – Библиогр.: с. 268-292. – ISBN 978-5-7782-2817-7.

Дополнительная литература:

3. Осипов, Г. С. Методы искусственного интеллекта / Г. С. Осипов. – Москва : Физматлит, 2011. – 296 с. : ил., схем., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=457464>. – ISBN 978-5-9221-1323-6.
4. Павлов, С. И. Системы искусственного интеллекта : учебное пособие / С. И. Павлов. – Томск : Томский государственный университет систем управления и радиоэлектроники, 2011. – Часть 2. – 194 с. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=208939>. – ISBN 978-5-4332-0014-2.
5. Учебник по машинному обучению : онлайн-учебник // Школа анализа данных : [сайт] / АНО ДПО «Образовательные технологии Яндекса». – [2007-2022]. – URL: <https://education.yandex.ru/handbook/ml>
6. Математические основы машинного обучения = Machine Learning : курс лекций К. В. Воронцова для студентов МФТИ : 45 видео / [Московский физико-технический

- институт] // YouTube : [видеохостинг]. – URL: <https://www.youtube.com/playlist?list=PLk4h7dmY2eYFmowaPqjFDzSokiiLq5TkT>
7. MachineLearningSeminars : семинары А. В. Грабовского к лекционному курсу К. В. Воронцова // GitHub : [сайт]. – 2023. – URL: <https://github.com/andriygav/MachineLearningSeminars>, <https://andriygav.github.io/MachineLearningSeminars/>
 8. Курс «Машинное обучение» 2019 : 22 видео // Компьютерные науки : [канал видеохостинга YouTube]. – Обновлено 2 мая 2020 г. – URL: <https://www.youtube.com/playlist?list=PLJOzdkh8T5krxc4HsHbB8g8f0hu7973fK>
 9. Python Data Science Handbook : [site] / Jake VanderPlas. – [Washington, 20--?]. – URL: <https://jakevdp.github.io/PythonDataScienceHandbook>
 10. Petersen, K. B. The Matrix Cookbook / K. B. Petersen, M. S. Pedersen. – [S. l.], 2012. – 72 p. – URL: <https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf>

6. Профессиональные базы данных и информационные справочные системы

1. Интернет-университет информационных технологий www.intuit.ru
2. ЭБС «Университетская библиотека онлайн» – <http://biblioclub.ru/>
3. ЭБС «IPRbooks» – <http://iprbookshop.ru/>
4. ЭБС «Консультант студента. Электронная библиотека технического вуза» –
5. <http://www.studentlibrary.ru/>
6. Электронная база данных «EBSCO» – <http://search.ebscohost.com/>
7. Национальная электронная библиотека – <https://rusneb.ru/>
8. Научная электронная библиотека eLIBRARY.RU – <http://elibrary.ru/defaultx.asp>
9. «Словари и энциклопедии на АКАДЕМИКЕ» (открытый доступ) – <http://dic.academic.ru/>

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства

1. Офисный пакет Microsoft Office 2007 Russian Academic OPEN.
2. Офисный пакет Microsoft Office 2010 Russian Academic OPEN.
3. Microsoft Visual Studio 2010 (подписка на образовательные лицензии, сетевые версии), участие в академической программе Microsoft Azure Dev Tools for Teaching.
4. GNU Compiler Collection
5. Свободное программное обеспечение Jupyter Notebook с установленным интерпретатором языка программирования Python 3 и Python-пакетами NumPy, SciPy, matplotlib, sklearn, keras
6. Веб-обозреватель (браузер), обеспечивающий доступ к Google Colaboratory.

8. Обеспечение освоения дисциплины лиц с инвалидностью и ОВЗ

Обучающиеся из числа инвалидов и лиц с ОВЗ обеспечиваются печатными и (или) электронными образовательными ресурсами в формах, адаптированных к ограничениям их здоровья.

9. Материально-техническое обеспечение дисциплины представлено в приложении к ОПОП «Материально-технические условия реализации образовательной программы» и включает:

- учебные аудитории для проведения учебных занятий, предусмотренных программой бакалавриата, оснащенные оборудованием и техническими средствами обучения;
- помещения для самостоятельной работы обучающихся, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде МАУ.

10. Распределение трудоемкости по видам учебной деятельности

Таблица 1 – Распределение трудоемкости

Вид учебной деятельности	Распределение трудоемкости дисциплины по формам обучения			
	Очная			
	Семестр			Всего часов
		3		
Лекции		20		20
Лабораторные работы		24		24
Самостоятельная работа		64		64
Подготовка к промежуточной аттестации		36		36
Всего часов по дисциплине		144		144
/из них в форме практической подготовки				
Формы промежуточной аттестации и текущего контроля				
Экзамен		+		1
Количество РГР		1		1

Перечень лабораторных работ по формам обучения

№ п\п	Темы лабораторных работ
1	2
	Очная форма
1.	Стохастический поиск: метод имитации отжига
2.	Коллективные методы: метод роя частиц, генетический алгоритм
3.	Регрессионные модели: простая линейная регрессия, множественная линейная регрессия, нелинейная регрессия
4.	Классификационные модели: метод ближайших соседей, логистическая регрессия
5.	Наивный байесовский классификатор
6.	Кластеризация: метод К-средних, смешанные гауссовы модели, иерархическая кластеризация, DBSCAN-кластеризация