

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Сибирский государственный индустриальный университет»
Кафедра прикладных информационных технологий и программирования

УТВЕРЖДАЮ
Директор института
информационных технологий и
автоматизированных систем
_____ Л.Д. Павлова
подпись
« ____ » _____ 20__ г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Теория искусственных нейронных сетей и машинное обучение

09.03.03 «Прикладная информатика»
(направленность (профиль): «Прикладная информатика»)

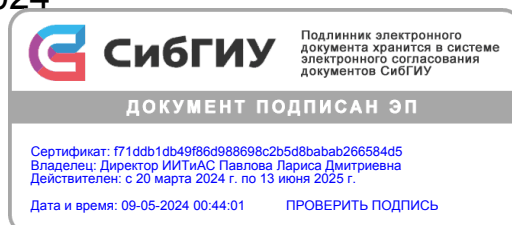
Квалификация выпускника
Бакалавр

Форма обучения
Очная форма

Срок обучения: 4 года

Год начала подготовки 2024

Новокузнецк
2024



1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

- формирование у обучающихся знаний и навыков применения методов машинного обучения для задач управления; моделирования и прогнозирования ML-методами, а также оценки построенной модели.

Задачами учебной дисциплины являются:

- изучение математических основ построения и архитектуры искусственных нейронных сетей;
- изучение различных алгоритмов обучения и настройки ИНС;
- изучение методов сбора и подготовки реальных данных для машинного обучения;
- изучение этапов проектирования признаков для построения моделей машинного обучения;
- изучение процесса построения моделей классификации и регрессии методами ML;
- изучение подходов к оценке построенных моделей.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам обязательной части **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.03.03 «Прикладная информатика».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

- Базы данных;
- Постреляционные базы данных;
- Теория вероятностей и математическая статистика;
- Численные методы решения инженерных задач;
- Анализ дискретных структур;
- Python для анализа данных;
- Обработка и анализ данных.

Учебная дисциплина дополняет знания и умения, получаемые по одновременно изучаемым и последующим учебным дисциплинам:

- Оптимизация в технике и технологиях;
- Распознавание образов.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

- **Профессиональные компетенции**

Наименование категории (группы) ПК	Код и наименование ПК	Код и наименование индикатора достижения ПК	Планируемые результаты обучения
	ПК-2: Способен участвовать в проведении аналитического исследования с применением технологии больших данных	ПК-2.1 Участвует в выборе методов и инструментальных средств анализа данных для проведения аналитических работ, в том числе с применением технологий больших данных	<p>– знать: архитектуры нейронных сетей, подходы к обучению ML-моделей, алгоритмы построения классификаторов, регрессоров, кластеров данных и ассоциативных правил, основные методы выделения признаков, типовые методы моделирования и прогнозирования, методы оптимизации параметров модели.</p> <p>– уметь: выбирать подход к обучению и алгоритм для ML-моделей, выполнять построение классификаторов, регрессоров, кластеров данных и ассоциативных правил, оптимизировать параметры модели машинного обучения при помощи сеточного поиска.</p>
ПК-2.2 Участвует в проведении аналитического исследования, в том числе с применением технологий больших данных		<p>– знать: методы сбора данных для машинного обучения, подходы к формированию наборов данных, методы разведочного анализа данных, методы и алгоритмы машинного обучения, основные этапы базового рабочего процесса ML.</p> <p>– уметь: формировать признаки для решаемых задач; выполнять разведочный анализ данных, выполнять</p>	

			построение модели машинного обучения с оценкой прогностической точности.
		ПК-2.3 Участвует в подготовке отчета по результатам аналитических работ	– знать: способы визуализации больших объемов данных, типовые диаграммы для представления данных. – уметь: строить графики и диаграммы, визуализирующие связь между входными признаками и целевой переменной.

4 Объем и содержание учебной дисциплины

Учебные занятия по учебной дисциплине проводятся в форме контактной работы и в форме самостоятельной работы обучающихся.

Контактная работа включает в себя занятия лекционного типа (лекции), занятия семинарского типа (семинары, практические занятия, практикумы, лабораторные работы, коллоквиумы), промежуточную аттестацию обучающихся и иные формы взаимодействия обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации ООП на иных условиях, в том числе при проведении промежуточной аттестации обучающихся. Контактная работа может проводиться с применением электронного обучения, дистанционных образовательных технологий.

Объем учебной дисциплины

Семестр / курс		ИТОГО	6 семестр	7 семестр
Форма промежуточной аттестации			экзамен	экзамен
Трудоёмкость	<i>академ. час.</i>	324	144	180
	<i>зачетных единиц</i>	9	4	5
Лекции, <i>академ. час.</i>		32	16	16
в форме практической подготовки		0	0	0
Лабораторные работы, <i>академ. час.</i>		16	0	16
в форме практической подготовки		0	0	0
Практические занятия, <i>академ. час.</i>		48	16	32
в форме практической подготовки		0	0	0
Курсовая работа / проект, <i>академ. час.</i>		0	0	0
в форме практической подготовки		0	0	0
Консультации, <i>академ. час.</i>		0	0	0
в форме практической подготовки		0	0	0

Самостоятельная работа, <i>академ. час.</i>	138	76	62
в форме практической подготовки	0	0	0
Контроль, <i>академ. час.</i>	90	36	54
в форме практической подготовки	0	0	0

Содержание учебной дисциплины

Раздел 1 Теоретические основы построения ИНС;

Тема 1.1 Основные понятия (Основные термины и понятия. Понятия искусственного нейрона и нейронной сети. Задачи построения нейронной сети. Этапы построения нейронной сети.);

Тема 1.2 Архитектура классических и современных ИНС (Типы нейронных сетей. Архитектура многослойного персептрона, сети Кохонена, Маркова, Хопфилда, сверточной сети, автоэнкодера, сети глубокого обучения (DCN, DN, DCIGN));

Тема 1.3 Способы обучения ИНС (Подходы к обучению. Обучение нейронной сети без учителя. Методы обучения с учителем. Обучение нейронной сети с подкреплением);

Раздел 2 Сбор и подготовка реальных данных для машинного обучения;

Тема 2.1 Сбор данных (Определение набора входных признаков. Наблюдаемое значение целевой переменной. Достаточный объем обучающих данных. Репрезентативность обучающей выборки);

Тема 2.2 Введение в структуры данных pandas и базовая функциональность (Объект Series. Объект DataFrame. Индексные объекты. Переиндексация. Доступ по индексу, выборка и фильтрация. Арифметические операции и выравнивание данных. Применение функций и отображение. Сортировка и ранжирование);

Тема 2.3 Редукция и вычисление описательных статистик (Корреляция и ковариация. Уникальные значения, счетчики значений и членство. Фильтрация отсутствующих данных);

Тема 2.4 Подготовка данных к моделированию (Категориальные признаки. Отсутствующие данные. Нормализация данных. Переформатирование данных: очистка, преобразование, слияние, изменение формы.);

Тема 2.5 Визуализация данных (API библиотеки matplotlib. Функции построения графиков в pandas. Нанесение данных на карту. Инструментальная экосистема визуализации для Python. Мозаичные диаграммы. Диаграммы размаха. Графики плотности. Диаграммы рассеяния);

Раздел 3 Моделирование и прогнозирование;

Тема 3.1 Основы моделирования с машинным обучением (Поиск связи между входными данными и целевой переменной. Типы методов моделирования. Обучение с учителем и без);

Тема 3.2 Классификация (Построение классификатора и получение предсказаний. Классификация сложных нелинейных данных. Классификация в случае множества классов);

Тема 3.3 Регрессия (Построение регрессора и генерация прогнозов. Регрессия для сложных нелинейных данных);

Тема 3.4 Кластеризация и задача ассоциативных правил (Задача кластеризации. Обнаружение аномалий. Решение задачи ассоциации. Автоэнкодеры.);

Раздел 4 Оценка и оптимизация модели;

Тема 4.1 Оценка прогностической точности на новых данных (Проблема переобучения и чрезмерно оптимистической оценки модели. Скользящий контроль. Перекрестная проверка);

Тема 4.2 Оценка моделей классификации (Точность для отдельных классов и таблица сопряженности. Компромиссы при оценке точности и ROC-кривые. Многоклассовая классификация);

Тема 4.3 Оценка моделей регрессии (Простые показатели эффективности регрессионных моделей. Исследование остатков);

Тема 4.4 Оценка моделей кластеризации (Проверка однородности и полноты кластеров. Метрики качества кластеризации. Метрики эталонных кластеров);

Тема 4.5 Оптимизация модели путем подбора параметров (Параметры настройки ML-алгоритмов. Сеточный поиск);

Раздел 5 Основы проектирования признаков;

Тема 5.1 Основные этапы проектирования признаков (Проектирование признаков и знание предметной области. Этапы проектирования признаков. Извлечение признаков из обычного текста);

Тема 5.2 Агрегирование данных и групповые операции (Механизм GroupBy. Агрегирование данных. Групповые операции и преобразования. Сводные таблицы и кросс-табуляция);

Тема 5.3 Выбор признаков (Прямой отбор и обратное исключение. Отбор признаков для исследования данных. Практический пример отбора признаков).

5 Перечень тем лекций

№ раздела / темы дисциплины	Темы лекций	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 1.	Теоретические основы построения ИНС		
Тема 1.1.	Основные понятия	2	
Тема 1.2.	Архитектура классических и современных ИНС	2	
Тема 1.3.	Способы обучения ИНС	4	
Раздел 2.	Сбор и подготовка реальных данных для		

	машинного обучения		
Тема 2.1.	Сбор данных	4	
Тема 2.4.	Подготовка данных к моделированию	2	
Тема 2.5.	Визуализация данных	2	
Раздел 3.	Моделирование и прогнозирование		
Тема 3.1.	Основы моделирования с машинным обучением	2	
Тема 3.2.	Классификация	2	
Тема 3.3.	Регрессия	2	
Тема 3.4.	Кластеризация и задача ассоциативных правил	1	
Раздел 4.	Оценка и оптимизация модели		
Тема 4.1.	Оценка прогностической точности на новых данных	2	
Тема 4.2.	Оценка моделей классификации	1	
Тема 4.3.	Оценка моделей регрессии	1	
Тема 4.4.	Оценка моделей кластеризации	1	
Тема 4.5.	Оптимизация модели путем подбора параметров	2	
Раздел 5.	Основы проектирования признаков		
Тема 5.1.	Основные этапы проектирования признаков	1	
Тема 5.3.	Выбор признаков	1	
Итого:		32	0

6 Перечень тем практических занятий (семинаров)

№ раздела / темы дисциплины	Темы практических занятий (семинаров)	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Тема 2.1.	Сбор и подготовка данных	4	
Тема 2.2.	Работа с объектами Series и DataFrame	2	
Тема 2.3.	Редукция и вычисление описательных статистик в Jupyter Notebook	2	
Тема 2.4.	Работа с различными типами данных в датасете	4	

Тема 2.5.	Визуализация данных	4	
Тема 3.1.	Поиск связи между входными данными и целевой переменной	4	
Тема 4.1.	Оценка прогностической точности на новых данных	4	
Тема 4.2.	Оценка моделей классификации	4	
Тема 4.3.	Оценка моделей регрессии	4	
Тема 4.4.	Оценка моделей кластеризации	4	
Тема 4.5.	Оптимизация модели путем подбора параметров	4	
Тема 5.1.	Этапы проектирования признаков	2	
Тема 5.2.	Агрегирование данных в Jupyter Notebook	4	
Тема 5.3.	Прямой отбор и обратное исключение признаков для набора данных	2	
Итого:		48	0

7 Перечень тем лабораторных работ

№ раздела / темы дисциплины	Темы лабораторных работ	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 3.	Решение задачи классификации при помощи библиотеки Scikit-learn	4	
Раздел 3.	Решение задачи регрессии Scikit-learn	4	
Раздел 3.	Решение задачи кластеризации Scikit-learn	4	
Раздел 3.	Решение задачи поиска ассоциативных правил Scikit-learn	4	
Итого:		16	0

8 Перечень тем курсовых работ (проектов)

№ раздела / темы дисциплины	Темы курсовых работ (проектов)	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
	<i>Отсутствуют</i>		

Итого:	0	0
---------------	----------	----------

9 Виды самостоятельной работы

№ раздела / темы дисциплины	Виды самостоятельной работы	Трудоемкость, <i>академ. час</i>	
		всего	в форме практической подготовки
Раздел 1.	1. Изучение лекционного материала; 2. Прохождение тестирования.	32	
Раздел 2.	1. Изучение лекционного материала; 2. Оформление отчета о практической работе; 3. Подготовка к практическому занятию; 4. Прохождение тестирования.	44	
Раздел 3.	1. Изучение лекционного материала; 2. Оформление отчета о практической работе; 3. Оформление отчета по лабораторной работе; 4. Подготовка к лабораторной работе; 5. Подготовка к практическому занятию; 6. Прохождение тестирования.	32	
Раздел 4.	1. Изучение лекционного материала; 2. Оформление отчета о практической работе; 3. Подготовка к практическому занятию; 4. Прохождение тестирования.	16	
Раздел 5.	1. Изучение лекционного материала; 2. Оформление отчета о практической работе; 3. Подготовка к практическому занятию; 4. Прохождение тестирования.	14	
<i>Контроль</i>	<i>Подготовка к экзамену (6 семестр)</i>	36	
<i>Контроль</i>	<i>Подготовка к экзамену (7</i>	54	

	семестр)		
Итого:		228	0

10 Учебно-методическое и информационное обеспечение учебной дисциплины

а) литература:

1 Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных : учебное пособие / П. Флах. – Москва : ДМК-пресс, 2015. – 400 с. – ISBN 978-5-97060-273-7. – URL: <https://www.studentlibrary.ru/book/ISBN9785970602737.html> (дата обращения: 06.05.2024);

2 Рашка, С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения : практическое пособие / С. Рашка. – Москва : ДМК-пресс, 2017. – 418 с. – ISBN 978-5-97060-409-0. – URL: <https://www.studentlibrary.ru/book/ISBN9785970604090.html> (дата обращения: 06.05.2024);

3 Глубокое обучение без математики. Т. 2: Практика : учебное пособие / Э. Гласснер. – Москва : ДМК-пресс, 2020. – ISBN 978-5-97060-767-1. – URL: <https://www.studentlibrary.ru/book/ISBN9785970607671.html> (дата обращения: 06.05.2024);

4 Глубокое обучение без математики. Т. 1: Основы : учебное пособие / Э. Гласснер. – Москва : ДМК-пресс, 2019. – ISBN 978-5-97060-701-5. – URL: <https://www.studentlibrary.ru/book/ISBN9785970607015.html> (дата обращения: 06.05.2024);

5 Маккинли, У. Python и анализ данных / У. Маккинли. – Москва : ДМК Пресс, 2015. – 482 с. – ISBN 978-5-97060-315-4. – URL: <https://www.studentlibrary.ru/ru/book/ISBN9785970603154.html> (дата обращения: 06.05.2024).

б) ресурсы информационно-телекоммуникационной сети «Интернет»:

1 Консультант студента : электронно-библиотечная система / ООО «КОНСУЛЬТАНТ СТУДЕНТА». – Москва, [200 –]. – URL: <http://www.studentlibrary.ru>. – Режим доступа: для авторизир. пользователей;

2 ЛАНЬ : электронно-библиотечная система : [коллекция «Инженерно-технические науки»] / ООО «Издательство ЛАНЬ». – Санкт-Петербург, [200 –]. – URL: <http://e.lanbook.com>. – Режим доступа: для авторизир. пользователей;

3 НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА eLIBRARY.RU : база данных / ООО «НЭБ». – Москва, [200 –]. – URL: <http://elibrary.ru>. – Режим доступа: по подписке;

4 Образовательная платформа ЮРАЙТ / ООО «Электронное издательство ЮРАЙТ». – Москва, [200 –]. – URL: <https://urait.ru>. – Режим доступа: для авторизир. пользователей;

5 Университетская библиотека онлайн : электронно-библиотечная система / ООО «Директ-Медиа». – Москва, [200 –]. – URL: <https://biblioclub.ru>. – Режим доступа: для авторизир. пользователей. – URL: <http://www.biblioclub.ru>;

6 Электронная библиотека // Научно-техническая библиотека СибГИУ : сайт. – Новокузнецк, [200 –]. – URL: <http://library.sibsiu.ru/LibrELibraryFullText.asp>. – Режим доступа: для авторизир. пользователей. – URL: <https://library.sibsiu.ru/LibrELibraryFullText.asp>;

7 Электронные периодические издания ИВИС : универсальная база данных / ООО «ИВИС». – Москва, [200 –]. – URL: <http://eivis.ru>. – Режим доступа: по подписке;

8 Электронный каталог : сайт / Научно-техническая библиотека СибГИУ. – Новокузнецк, [199 –]. – URL: <http://libr.sibsiu.ru>. – URL: <https://libr.sibsiu.ru>.

в) лицензионное и свободно распространяемое программное обеспечение:

- 7-Zip;
- Adobe Acrobat Reader;
- Astra Linux Special Edition;
- JetBrains PyCharm Community Edition;
- Kaspersky Endpoint Security;
- Microsoft Office;
- Microsoft Windows;
- Notepad++;
- OnlyOffice;
- P7-Офис.

г) базы данных и информационно-справочные системы:

1 ГАРАНТ : справочно-правовая система / ООО «Правовой центр «Гарант». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.;

2 КонсультантПлюс : справочно-правовая система / ООО «Информационный центр АНВИК». – Новокузнецк, [199 –]. – Режим доступа: компьютерная сеть библиотеки Сиб. гос. индустр. ун-та.;

3 Техэксперт : информационно-справочная система / ООО «Группа компаний «Кодекс». – Кемерово, [200 –]. – Режим доступа: компьютерная сеть Сиб. гос. индустр. ун-та.

11 Материально-техническое обеспечение учебной дисциплины

Материально-техническое обеспечение учебной дисциплины включает учебные аудитории, оснащенные оборудованием, компьютерной техникой, и техническими средствами обучения, в том числе:

- учебную аудиторию для проведения занятий лекционного типа, оборудованную учебной доской, экраном и мультимедийным проектором;
- учебную аудиторию для проведения занятий семинарского типа (практических занятий), оснащенную персональными компьютерами и выходом в сеть «Интернет»;
- учебную аудиторию для проведения занятий семинарского типа (лабораторных работ), оснащенную персональными компьютерами и выходом в сеть «Интернет»;
- учебную аудиторию (помещения) для групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации;
- помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, научно-техническую библиотеку СибГИУ.

Рабочая программа дисциплины составлена в соответствии с требованиями ФГОС ВО по направлению подготовки (специальности) 09.03.03 «Прикладная информатика».

Составитель(и):

доцент Бабичева Надежда Борисовна (кафедра прикладных информационных технологий и программирования).

Рабочая программа дисциплины рассмотрена и утверждена на заседании кафедры.

Приложение

Аннотация

рабочей программы дисциплины «Теория искусственных нейронных сетей и машинное обучение»

по направлению подготовки (специальности)

09.03.03 «Прикладная информатика»

(направленность (профиль): «Прикладная информатика»)

форма обучения – Очная форма

1 Цели и задачи освоения учебной дисциплины

Целями учебной дисциплины являются:

- формирование у обучающихся знаний и навыков применения методов машинного обучения для задач управления; моделирования и прогнозирования ML-методами, а также оценки построенной модели.

Задачами учебной дисциплины являются:

- изучение математических основ построения и архитектуры искусственных нейронных сетей;
- изучение различных алгоритмов обучения и настройки ИНС;
- изучение методов сбора и подготовки реальных данных для машинного обучения;
- изучение этапов проектирования признаков для построения моделей машинного обучения;
- изучение процесса построения моделей классификации и регрессии методами ML;
- изучение подходов к оценке построенных моделей.

2 Место учебной дисциплины в структуре ООП по направлению подготовки (специальности)

Учебная дисциплина относится к учебным дисциплинам обязательной части **Блока 1 «Дисциплины (модули)»** ООП по направлению подготовки (специальности) 09.03.03 «Прикладная информатика».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

- Базы данных;
- Постреляционные базы данных;
- Теория вероятностей и математическая статистика;
- Численные методы решения инженерных задач;
- Анализ дискретных структур;
- Python для анализа данных;
- Обработка и анализ данных.

Учебная дисциплина дополняет знания и умения, получаемые по одновременно изучаемым и последующим учебным дисциплинам:

- Оптимизация в технике и технологиях;
- Распознавание образов.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Профессиональные компетенции

Наименование категории (группы) ПК	Код и наименование ПК	Код и наименование индикатора достижения ПК	Планируемые результаты обучения
	ПК-2: Способен участвовать в проведении аналитического исследования с применением технологии больших данных	ПК-2.1 Участвует в выборе методов и инструментальных средств анализа данных для проведения аналитических работ, в том числе с применением технологий больших данных	<p>– знать: архитектуры нейронных сетей, подходы к обучению ML-моделей, алгоритмы построения классификаторов, регрессоров, кластеров данных и ассоциативных правил, основные методы выделения признаков, типовые методы моделирования и прогнозирования, методы оптимизации параметров модели.</p> <p>– уметь: выбирать подход к обучению и алгоритм для ML-моделей, выполнять построение классификаторов, регрессоров, кластеров данных и ассоциативных правил, оптимизировать параметры модели машинного обучения при помощи сеточного поиска.</p>
ПК-2.2 Участвует в проведении аналитического исследования, в том числе с применением технологий больших данных		<p>– знать: методы сбора данных для машинного обучения, подходы к формированию наборов данных, методы разведочного анализа данных, методы и алгоритмы</p>	

			<p>машинного обучения, основные этапы базового рабочего процесса ML.</p> <p>– уметь: формировать признаки для решаемых задач; выполнять разведочный анализ данных, выполнять построение модели машинного обучения с оценкой прогностической точности.</p>
		ПК-2.3 Участвует в подготовке отчета по результатам аналитических работ	<p>– знать: способы визуализации больших объемов данных, типовые диаграммы для представления данных.</p> <p>– уметь: строить графики и диаграммы, визуализирующие связь между входными признаками и целевой переменной.</p>

4 Объем учебной дисциплины

Семестр / курс		ИТОГО	6 семестр	7 семестр
Форма промежуточной аттестации			<i>экзамен</i>	<i>экзамен</i>
Трудоёмкость	<i>академ. час.</i>	324	144	180
	<i>зачетных единиц</i>	9	4	5
Лекции, <i>академ. час.</i>		32	16	16
в форме практической подготовки		0	0	0
Лабораторные работы, <i>академ. час.</i>		16	0	16
в форме практической подготовки		0	0	0
Практические занятия, <i>академ. час.</i>		48	16	32
в форме практической подготовки		0	0	0
Курсовая работа / проект, <i>академ. час.</i>		0	0	0
в форме практической подготовки		0	0	0
Консультации, <i>академ. час.</i>		0	0	0
в форме практической подготовки		0	0	0
Самостоятельная работа, <i>академ. час.</i>		138	76	62
в форме практической подготовки		0	0	0
Контроль, <i>академ. час.</i>		90	36	54
в форме практической подготовки		0	0	0

5 Краткое содержание учебной дисциплины

В структуре учебной дисциплины выделяются следующие основные разделы (темы):

Раздел 1 Теоретические основы построения ИНС;

Тема 1.1 Основные понятия (Основные термины и понятия. Понятия искусственного нейрона и нейронной сети. Задачи построения нейронной сети. Этапы построения нейронной сети.);

Тема 1.2 Архитектура классических и современных ИНС (Типы нейронных сетей. Архитектура многослойного персептрона, сети Кохонена, Маркова, Хопфилда, сверточной сети, автоэнкодера, сети глубокого обучения (DCN, DN, DCIGN));

Тема 1.3 Способы обучения ИНС (Подходы к обучению. Обучение нейронной сети без учителя. Методы обучения с учителем. Обучение нейронной сети с подкреплением);

Раздел 2 Сбор и подготовка реальных данных для машинного обучения;

Тема 2.1 Сбор данных (Определение набора входных признаков. Наблюдаемое значение целевой переменной. Достаточный объем обучающих данных. Репрезентативность обучающей выборки);

Тема 2.2 Введение в структуры данных pandas и базовая функциональность (Объект Series. Объект DataFrame. Индексные объекты. Переиндексация. Доступ по индексу, выборка и фильтрация. Арифметические операции и выравнивание данных. Применение функций и отображение. Сортировка и ранжирование);

Тема 2.3 Редукция и вычисление описательных статистик (Корреляция и ковариация. Уникальные значения, счетчики значений и членство. Фильтрация отсутствующих данных);

Тема 2.4 Подготовка данных к моделированию (Категориальные признаки. Отсутствующие данные. Нормализация данных. Переформатирование данных: очистка, преобразование, слияние, изменение формы.);

Тема 2.5 Визуализация данных (API библиотеки matplotlib. Функции построения графиков в pandas. Нанесение данных на карту. Инструментальная экосистема визуализации для Python. Мозаичные диаграммы. Диаграммы размаха. Графики плотности. Диаграммы рассеяния);

Раздел 3 Моделирование и прогнозирование;

Тема 3.1 Основы моделирования с машинным обучением (Поиск связи между входными данными и целевой переменной. Типы методов моделирования. Обучение с учителем и без);

Тема 3.2 Классификация (Построение классификатора и получение предсказаний. Классификация сложных нелинейных данных. Классификация в случае множества классов);

Тема 3.3 Регрессия (Построение регрессора и генерация прогнозов. Регрессия для сложных нелинейных данных);

Тема 3.4 Кластеризация и задача ассоциативных правил (Задача кластеризации. Обнаружение аномалий. Решение задачи ассоциации. Автоэнкодеры.);

Раздел 4 Оценка и оптимизация модели;

Тема 4.1 Оценка прогностической точности на новых данных (Проблема переобучения и чрезмерно оптимистической оценки модели. Скользящий контроль. Перекрестная проверка);

Тема 4.2 Оценка моделей классификации (Точность для отдельных классов и таблица сопряженности. Компромиссы при оценке точности и ROC-кривые. Многоклассовая классификация);

Тема 4.3 Оценка моделей регрессии (Простые показатели эффективности регрессионных моделей. Исследование остатков);

Тема 4.4 Оценка моделей кластеризации (Проверка однородности и полноты кластеров. Метрики качества кластеризации. Метрики эталонных кластеров);

Тема 4.5 Оптимизация модели путем подбора параметров (Параметры настройки ML-алгоритмов. Сеточный поиск);

Раздел 5 Основы проектирования признаков;

Тема 5.1 Основные этапы проектирования признаков (Проектирование признаков и знание предметной области. Этапы проектирования признаков. Извлечение признаков из обычного текста);

Тема 5.2 Агрегирование данных и групповые операции (Механизм GroupBy. Агрегирование данных. Групповые операции и преобразования. Сводные таблицы и кросс-табуляция);

Тема 5.3 Выбор признаков (Прямой отбор и обратное исключение. Отбор признаков для исследования данных. Практический пример отбора признаков).

6 Составитель(и):

доцент Бабичева Надежда Борисовна (кафедра прикладных информационных технологий и программирования).