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

09.03.01 «Информатика и вычислительная техника»
(направленность (профиль): «Информатика и вычислительная техника»)

Квалификация выпускника
Бакалавр

Форма обучения
Очная форма

Срок обучения: 3 года 5 месяцев

Год начала подготовки 2023

Новокузнецк
2023

СибГИУ
Подлинник электронного документа хранится в системе электронного согласования документов СибГИУ
ДОКУМЕНТ ПОДПИСАН ЭП
Сертификат: 861a8d56fc2cfb6de53bd665403c592fd02b1419
Владелец: Ректор Юрьев Алексей Борисович
Действителен: с 12 февраля 2024 г. по 7 мая 2025 г.
Дата и время: 01-03-2024 00:31:33 ПРОВЕРИТЬ ПОДПИСЬ

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) 09.03.01 «Информатика и вычислительная техника».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

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

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

– Оптимизация в технике и технологиях;
– Проектная деятельность 3.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Общепрофессиональные компетенции

Наименование категории (группы) ОПК	Код и наименование ОПК	Код и наименование индикатора достижения ОПК	Планируемые результаты обучения
	ОПК-1: Способен применять естественнонаучные и общеинженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности	ОПК-1.3 Применяет методы математического анализа и моделирования в теоретических и экспериментальных исследованиях профессиональной деятельности	– знать: основные методы выделения признаков, типовые методы моделирования и прогнозирования, методы оптимизации параметров модели. – уметь: подбирать наиболее оптимальный набор входных признаков для модели, оптимизировать параметры модели машинного обучения при помощи сеточного поиска. – владеть: решать задачи по подбору входных признаков для модели, решать задачи по оптимизации модели машинного обучения.
	ОПК-2: Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности	ОПК-2.3 Применяет современные информационные технологии и программные средства при решении прикладных задач	– знать: методы сбора данных для машинного обучения, подходы к формированию наборов данных, методы разведочного анализа данных. – уметь: формировать признаки для решаемых задач; выполнять разведочный анализ данных. – владеть: решать задачи по сбору и разведочному анализу данных для моделей машинного обучения в соответствии с

			предметной областью.
	ОПК-3: Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности	ОПК-3.3 Решает поставленные задачи на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий	– знать: архитектуры нейронных сетей, подходы к обучению ML-моделей, алгоритмы построения классификаторов, регрессоров, кластеров данных и ассоциативных правил. – уметь: выбирать подход к обучению и алгоритм для ML-моделей, выполнять построение классификаторов, регрессоров, кластеров данных и ассоциативных правил. – владеть: решать задачи по подбору методов моделирования и прогнозирования, решать задачи по классификации, регрессии, кластеризации, ассоциативным правилам.
	ОПК-9: Способен осваивать методики использования программных средств для решения практических задач.	ОПК-9.2 Осваивает методики использования современных программных средств для решения научно-исследовательских и прикладных задач	– знать: способы визуализации больших объемов данных, типовые диаграммы для представления данных. – уметь: строить графики и диаграммы, визуализирующие связь между входными признаками и целевой переменной. – владеть: решать задачи по визуализации

			больших объемов данных при помощи библиотек и прикладных пакетов программ для анализа данных.
		ОПК-9.3 Решает типовые задачи в области профессиональной деятельности на основе известных методик, типовых алгоритмов и программных решений	– знать: методы и алгоритмы машинного обучения, основные этапы базового рабочего процесса ML. – уметь: выполнять построение модели машинного обучения с оценкой прогностической точности. – владеть: решать задачи по построению моделей машинного обучения для задач в предметной области.

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***	***4 семестр***	***5 семестр***
Форма промежуточной аттестации			*экзамен*	*экзамен*
Трудоёмкость	*академ. час.*	***324***	*144*	*180*
	зачетных единиц	***9***	*4*	*5*
Лекции, *академ. час.*		***32***	*16*	*16*
в форме практической подготовки		***0***	*0*	*0*
Лабораторные работы, *академ. час.*		***16***	*0*	*16*
в форме практической подготовки		***0***	*0*	*0*

Практические занятия, *академ. час.*	***48***	*16*	*32*
в форме практической подготовки	***0***	*0*	*0*
Курсовая работа / проект, *академ. час.*	***0***	*0*	*0*
в форме практической подготовки	***0***	*0*	*0*
Консультации, *академ. час.*	***0***	*0*	*0*
в форме практической подготовки	***0***	*0*	*0*
Самостоятельная работа, *академ. час.*	***156***	*76*	*80*
в форме практической подготовки	***0***	*0*	*0*
Контроль, *академ. час.*	***72***	*36*	*36*
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, *академ.час***	
		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, *академ.час***	
		всего	**в форме практической подготовки**
Тема 2.1.	Сбор и подготовка данных	4	
Тема 2.2.	Работа с объектами	2	

	Series и DataFrame		
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, *академ.час***	
		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, *академ.час*	
		всего	в форме практической подготовки
	Отсутствуют		
Итого:		**0**	**0**

9 Виды самостоятельной работы

№ раздела / темы дисциплины	Виды самостоятельной работы	Трудоемкость, *академ.час*	
		всего	в форме практической подготовки
Раздел 1.	1. Прохождение тестирования.	32	
Раздел 2.	1. Оформление отчета о практической работе; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	44	
Раздел 3.	1. Оформление отчета о практической работе; 2. Оформление отчета по лабораторной работе; 3. Подготовка к лабораторной работе; 4. Подготовка к практическому занятию; 5. Прохождение тестирования.	40	
Раздел 4.	1. Оформление отчета о практической работе; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	22	
Раздел 5.	1. Оформление отчета о практической работе; 2. Подготовка к практическому занятию; 3. Прохождение тестирования.	18	
Контроль	*Подготовка к экзамену (4 семестр)*	36	
Контроль	*Подготовка к экзамену (5 семестр)*	36	
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: 25.05.2023);

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: 25.05.2023);

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: 25.05.2023);

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: 25.05.2023);

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: 25.05.2023).

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 Национальная электронная библиотека (НЭБ) : информационная система / ФГБУ «РГБ». – Москва, [2015 –]. – URL: http://rusneb.ru. – Режим доступа: по подписке;

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

6 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;

7 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;

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

9 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;
– Kaspersky Endpoint Security;
– Notepad++;
– PyCharm.

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) 09.03.01 «Информатика и вычислительная техника».

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)
09.03.01 «Информатика и вычислительная техника»
(направленность (профиль): «Информатика и вычислительная техника»)
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) 09.03.01 «Информатика и вычислительная техника».

Учебная дисциплина базируется на предварительном усвоении обучающимися учебных дисциплин:

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

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

– Оптимизация в технике и технологиях;

– Проектная деятельность 3.

3 Планируемые результаты обучения по учебной дисциплине

Процесс изучения учебной дисциплины направлен на формирование следующих компетенций:

– Общепрофессиональные компетенции

Наименование категории (группы) ОПК	Код и наименование ОПК	Код и наименование индикатора достижения ОПК	Планируемые результаты обучения
	ОПК-1: Способен применять естественнонаучные и общеинженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности	ОПК-1.3 Применяет методы математического анализа и моделирования в теоретических и экспериментальных исследованиях профессиональной деятельности	– знать: основные методы выделения признаков, типовые методы моделирования и прогнозирования, методы оптимизации параметров модели. – уметь: подбирать наиболее оптимальный набор входных признаков для модели, оптимизировать параметры модели машинного обучения при помощи сеточного поиска. – владеть: решать задачи по подбору входных признаков для модели, решать задачи по оптимизации модели машинного обучения.
	ОПК-2: Способен понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и	ОПК-2.3 Применяет современные информационные технологии и программные средства при решении прикладных задач	– знать: методы сбора данных для машинного обучения, подходы к формированию наборов данных, методы разведочного анализа данных. – уметь: формировать признаки для

	использовать их при решении задач профессиональной деятельности		решаемых задач; выполнять разведочный анализ данных. – владеть: решать задачи по сбору и разведочному анализу данных для моделей машинного обучения в соответствии с предметной областью.
	ОПК-3: Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности	ОПК-3.3 Решает поставленные задачи на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий	– знать: архитектуры нейронных сетей, подходы к обучению ML-моделей, алгоритмы построения классификаторов, регрессоров, кластеров данных и ассоциативных правил. – уметь: выбирать подход к обучению и алгоритм для ML-моделей, выполнять построение классификаторов, регрессоров, кластеров данных и ассоциативных правил. – владеть: решать задачи по подбору методов моделирования и прогнозирования, решать задачи по классификации, регрессии, кластеризации, ассоциативным правилам.
	ОПК-9: Способен осваивать методики использования программных средств для решения практических задач.	ОПК-9.2 Осваивает методики использования современных программных средств для решения научно-	– знать: способы визуализации больших объемов данных, типовые диаграммы для представления данных.

		исследовательских и прикладных задач	– уметь: строить графики и диаграммы, визуализирующие связь между входными признаками и целевой переменной. – владеть: решать задачи по визуализации больших объемов данных при помощи библиотек и прикладных пакетов программ для анализа данных.
		ОПК-9.3 Решает типовые задачи в области профессиональной деятельности на основе известных методик, типовых алгоритмов и программных решений	– знать: методы и алгоритмы машинного обучения, основные этапы базового рабочего процесса ML. – уметь: выполнять построение модели машинного обучения с оценкой прогностической точности. – владеть: решать задачи по построению моделей машинного обучения для задач в предметной области.

4 Объем учебной дисциплины

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

Консультации, *академ. час.*	***0***	*0*	*0*
в форме практической подготовки	***0***	*0*	*0*
Самостоятельная работа, *академ. час.*	***156***	*76*	*80*
в форме практической подготовки	***0***	*0*	*0*
Контроль, *академ. час.*	***72***	*36*	*36*
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).