

Подписано электронной подписью:
Вержицкий Данил Григорьевич
Должность: Директор КГПИ ФГБОУ ВО «КемГУ»
Дата и время: 2024-04-24 00:00:00
471086fad29a3b30e244c728abc3661ab35c9d50210dcf0e75e03a5b6fdf6436

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«Кемеровский государственный университет»
Кузбасский гуманитарно-педагогический институт

Факультет информатики, математики и экономики

«УТВЕРЖДАЮ»

Декан ФИМЭ

А.В. Фомина /

«10» февраля 2023 г.

Рабочая программа дисциплины

Б1.О.12.04 Машинное обучение

Направление подготовки

44.03.05 Педагогическое образование (с двумя профилями подготовки)

Направленность (профиль) подготовки

Информатика и Системы искусственного интеллекта

Программа бакалавриата

Квалификация выпускника

бакалавр

Форма обучения

Очная

Год набора 2021

Новокузнецк 2023

ОГЛАВЛЕНИЕ

Оглавление	2
1 Цель дисциплины	3
1.1 Формируемые компетенции	3
1.2 Индикаторы достижения компетенций	3
1.3 Знания, умения, навыки (ЗУВ) по дисциплине	4
2 Объём и трудоёмкость дисциплины по видам учебных занятий. Формы промежуточной аттестации.	5
3 Учебно-тематический план и содержание дисциплины	5
3.1 Учебно-тематический план.....	5
3.2 Содержание занятий по видам учебной работы	6
4 Порядок оценивания успеваемости и сформированности компетенций обучающегося в текущей и промежуточной аттестации.....	9
5 Учебно-методическое обеспечение дисциплины	9
5.1 Учебная литература	9
5.2 Материально-техническое и программное обеспечение дисциплины.....	10
5.3 Современные профессиональные базы данных и информационные справочные системы.....	11
6 Иные сведения и (или) материалы	12
6.1 Примерные вопросы и задания / задачи для промежуточной аттестации	12

1 ЦЕЛЬ ДИСЦИПЛИНЫ

В результате освоения данной дисциплины у обучающегося должны быть сформированы компетенции основной профессиональной образовательной программы бакалавриата (далее — ОПОП): ОПК-8.

Содержание компетенций как планируемых результатов обучения по дисциплине см. таблицы 1 и 2.

1.1 Формируемые компетенции

Таблица 1 — Формируемые дисциплиной компетенции

Наименование вида компетенции (универсальная, общепрофессиональная, профессиональная)	Наименование категории (группы) компетенций	Код и название компетенции
общепрофессиональная		ОПК-8 Способен осуществлять педагогическую деятельность на основе специальных научных знаний

1.2 Индикаторы достижения компетенций

Таблица 2 — Индикаторы достижения компетенций, формируемые дисциплиной

Код и название компетенции	Индикаторы достижения компетенции по ОПОП	Дисциплины и практики, формирующие компетенцию ОПОП
ОПК-8 Способен осуществлять педагогическую деятельность на основе специальных научных знаний	ОПК.8.1. Применяет специальные научные знания предметной области в педагогической деятельности по профилю подготовки ОПК.8.2. Владеет методами научного исследования в предметной области	Б1.О.03.01 Общая психология Б1.О.04 Возрастная анатомия и физиология Б1.О.06 Специальная психология и коррекционная педагогика Б1.О.08 Методы исследования в деятельности педагога Б1.О.11.01 Линейная алгебра Б1.О.11.02 Компьютерная графика и анимация Б1.О.11.03 Программирование Б1.О.11.04 Теоретические основы информатики Б1.О.11.05 Теория вероятностей и математическая статистика Б1.О.11.06 Компьютерные сети и интернет технологии Б1.О.11.07 Компьютерное моделирование Б1.О.11.08 Математическая логика Б1.О.11.09 Проектирование информационных систем Б1.О.12 Предметная подготовка по профилю "Системы искусственного интеллекта" Б1.О.12.01 Программное обеспечение Б1.О.12.02 Основы учебной робототехники Б1.О.12.03 Алгоритмы и

Код и название компетенции	Индикаторы достижения компетенции по ОПОП	Дисциплины и практики, формирующие компетенцию ОПОП
		структуры данных Б1.О.12.04 Машинное обучение Б1.О.12.05 Основы искусственного интеллекта Б1.О.12.06 Электроника и автоматика Б1.О.12.07 Информационная безопасность Б1.О.12.08 Дистанционные системы обучения Б1.О.12.09 Моделирование интеллектуальных систем Б2.О.01(У) Ознакомительная практика Б2.О.03(П) Психолого-педагогическая практика Б2.О.06(П) Педагогическая практика. Основная школа Б2.О.07(П) Педагогическая практика. Старшая школа ФТД.02 Видеомонтаж

1.3 Знания, умения, навыки (ЗУВ) по дисциплине

Таблица 3 — Знания, умения, навыки, формируемые дисциплиной

Код и название компетенции	Индикаторы достижения компетенции, закрепленные за дисциплиной	Знания, умения, навыки (ЗУВ), формируемые дисциплиной
ОПК-8 Способен осуществлять педагогическую деятельность на основе специальных научных знаний	ОПК.8.1. Применяет специальные научные знания предметной области в педагогической деятельности по профилю подготовки ОПК.8.2. Владеет методами научного исследования в предметной области	Знать: - научное содержание и современное состояние предметной области «Машинное обучение» - методы проведения научного исследования в предметной области «Машинное обучение»; Уметь: - использовать научные знания предметной области «Машинное обучение» в педагогической деятельности по профилю подготовки; - применять научные знания предметной области «Машинное обучение» при разработке образовательных программ, рабочих программ учебных предметов, курсов внеурочной деятельности; Владеть: - методами научного исследования в области интеллектуального анализа данных; - способами получения информации о современном состоянии научных исследований в предметной области «Машинное обучение»

2 ОБЪЁМ И ТРУДОЁМКОСТЬ ДИСЦИПЛИНЫ ПО ВИДАМ УЧЕБНЫХ ЗАНЯТИЙ. ФОРМЫ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ.

Таблица 4 — Объем и трудоемкость дисциплины по видам учебных занятий

Общая трудоемкость и виды учебной работы по дисциплине, проводимые в разных формах	Объём часов по формам обучения		
	ОФО	ОЗФО	ЗФО
1 Общая трудоемкость дисциплины	360		
2 Контактная работа обучающихся с преподавателем (по видам учебных занятий) (всего)	106		
Аудиторная работа (всего):	106		
в том числе:			
лекции	52		
практические занятия, семинары			
практикумы			
лабораторные работы	64		
в интерактивной форме			
в электронной форме			
Внеаудиторная работа (всего):			
в том числе, индивидуальная работа обучающихся с преподавателем			
подготовка курсовой работы /контактная работа			
групповая, индивидуальная консультация и иные виды учебной деятельности, предусматривающие групповую или индивидуальную работу обучающихся с преподавателем)			
творческая работа (эссе)			
3 Самостоятельная работа обучающихся (всего)	172		
4 Промежуточная аттестация обучающегося			
— экзамен (7 семестр)	36		
— экзамен (8 семестр)	36		

3 УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3.1 Учебно-тематический план

Таблица 5 — Учебно-тематический план очной формы обучения

№ недели п/п	Разделы и темы дисциплины по занятиям	Общая трудоемкость (всего час.)	Трудоемкость занятий (час.)						Формы текущего контроля и промежуточной аттестации успеваемости
			ОФО			ЗФО			
			Аудиторн. занятия	СР С	СР С	Аудиторн. занятия		СР С	
лекц.	практ.	лекц.	практ.						
Семестр 7									
1	<i>Основы машинного обучения</i>	62	10	12	40				Устный опрос, индивидуальное задание
2	<i>Инструментальные средства машинного обучения</i>	82	10	20	52				Устный опрос,

								индивидуальное задание
	Промежуточная аттестация - экзамен	36						Экзамен
	ИТОГО по семестру ...	180	20	32	92			
	Семестр 8							
1	Моделирование и прогнозирование	24	6	8	10			Устный опрос, индивидуальное задание
2	Оценка и оптимизация модели	36	8	4	24			Устный опрос, индивидуальное задание
3	Основы проектирования признаков	26	6	4	16			Устный опрос, индивидуальное задание
4	Искусственные нейронные сети в задачах машинного обучения	58	12	16	30			Устный опрос, индивидуальное задание
	Промежуточная аттестация экзамен	36						экзамен
	ИТОГО по семестру ...	180	32	32	80			
	Всего:	360	52	64	172			

3.2 Содержание занятий по видам учебной работы

Таблица 6 – Содержание дисциплины

№ п/п	Наименование раздела, темы дисциплины	Содержание занятия
	Семестр 7	
	<i>Содержание лекционного курса</i>	
1	<i>Основы машинного обучения</i>	
1.1	Принятие решений на основе данных	Традиционные подходы. Подходы с машинным обучением. Преимущества машинного обучения. Проблемы машинного обучения
1.2	Рабочий процесс машинного обучения	Сбор и подготовка данных. Обучение на модели данных. Оценка производительности модели. Оптимизация производительности модели.
1.3	Способы повышения эффективности обучения.	Предварительная обработка данных и проектирование признаков Непрерывное совершенствование моделей. Масштабирование моделей.
1.4	Сбор данных	Определение набора входных признаков. Значение целевой переменной. Достаточный объем обучающих данных. Репрезентативность обучающей выборки.
1.5	Подготовка данных к моделированию	Категориальные признаки. Отсутствующие данные. Основы проектирования признаков. Нормализация данных.
2	<i>Инструментальные средства машинного обучения</i>	
2.1	Командная оболочка IPython	Назначение оболочки IPython. Взаимодействие оболочки IPython и командного процессора. Организация отладки в оболочке IPython. Профилирование и мониторинг скорости выполнения кода.
2.2	Библиотека NumPy в задачах машинного обучения	Основы библиотеки NumPy. Базовые типы данных библиотеки NumPy.
2.3	Массивы в библиотеке NumPy.	Особенности массивов в NumPy. Универсальные функции. Агрегирование. Транслирование.
2.4	Основы пакета Pandas	Назначение пакета Pandas. Индексация и выборка данных. Обработка отсутствующих данных. Объединение наборов данных.
2.5	Визуализация данных	Мозаичные диаграммы. Диаграммы размаха. Графики плотности. Диаграммы рассеяния.

№ п/п	Наименование раздела, темы дисциплины	Содержание занятия
<i>Содержание лабораторных занятий</i>		
1	<i>Основы машинного обучения</i>	
1.1	Организация сбора и подготовки данных с машинном обучении	Способы представления данных. Форматы входных данных. Инструменты преобразования данных в требуемый формат.
1.2	Получение исходных данных из источников данных	Файловые операции в Python. Библиотеки для работы с форматами данных. Кодировки. Использование готовых наборов данных.
1.3	Осуществление первичной обработка данных.	Изменение представления данных. Осуществление нормализации данных.
2	<i>Инструментальные средства машинного обучения</i>	
2.1	Знакомство с командной оболочкой IPython	Запуск IPython. Получение справки. Использование горячих клавиш. Использование команд IPython. История ввода и вывода. Подавление вывода.
2.2	Выполнение операций с массивами в NumPy.	Создание массивов из списков. Создание пустых массивов. Атрибуты массивов. Доступ к отдельным элементам. Срезы массивов. Изменение формы массивов. Слияние и разбиение массивов.
2.3	Выполнение вычислений с использованием массивов в NumPy	Суммирование значений. Поиск минимума и максимума. Работа с булевыми массивами. Сортировка массивов. Выполнение частичной сортировки.
2.4	Работа с библиотекой Pandas	Создание и использование объектов Series, DataFrame и Index. Выборка данных из объекта Series. Выборка данных из объекта DataFrame. Операции над пустыми значениями. Создание иерархической индексации.
2.5	Построение графиков и диаграмм с помощью пакета matplotlib	Отображение графиков на экране и сохранение их в файл. Настройка цветов и стилей линий. Настройка осей координат. Построение линейных графиков. Построение диаграмм рассеяния. Визуализация погрешностей. Построение гистограмм. Построение графиков плотностей и контурных графиков.
	Промежуточная аттестация - экзамен	
	Семестр 8	
<i>Содержание лекционного курса</i>		
1.	Моделирование и прогнозирование	
1.1	Основы моделирования с машинным обучением	Поиск связи между входными данными и целевой переменной. Типы методов моделирования. Обучение с учителем и обучение без учителя.
1.2	Классификация	Построение классификатора и получение предсказаний. Классификация сложных нелинейных данных. Классификация в случае множества классов.
1.3	Регрессия.	Построение регрессора и генерация прогнозов. Регрессия для сложных нелинейных данных.
2.	Оценка и оптимизация модели	
2.1	Оценка прогностической точности на новых данных.	Переобучение и чрезмерно оптимистическая оценка модели. Скользящий контроль.
2.2	Оценка моделей классификации	Точность для отдельных классов и таблиц сопряженности. Компромиссы при оценке точности и ROC-кривые. Многоклассовая классификация.
2.3	Оценка моделей регрессии	Простые показатели эффективности регрессионных моделей. Исследование остатков.
2.4	Оптимизация моделей путем подбора параметров.	Параметры настройки ML-алгоритмов. Сеточный поиск.
3.	Основы проектирования признаков	
3.1	Назначение проектирования признаков	Суть проектирования признаков. Причины проектирования признаков. Проектирование признаков и знание предметной области.
3.2	Основные этапы проектирования признаков	Первичный анализ признаков. Обработка даты и времени. Извлечение признаков из текста.
3.3	Выбор признаков	Прямой отбор и обратное исключение. Отбор признаков для

№ п/п	Наименование раздела, темы дисциплины	Содержание занятия
		исследования данных.
4.	Искусственные нейронные сети в задачах машинного обучения	
4.1	Понятие искусственных нейронных сетей	Искусственный нейрон. Представление данных для нейронных сетей. Нейронные сети прямого распространения. Функции активации.
4.2	Обратное распространение	Выходы нейронов и ошибка сети. Использование ошибки для изменения весов. Скорость обучения. Пакеты данных.
4.3	Оптимизаторы.	Настройка скорости обучения. Стратегии обновления. Варианты градиентного спуска.
4.4	Сверточные нейронные сети.	Понятие свертки. Фильтры. Иерархия фильтров. Многомерная свертка. Транспонированная свертка
4.5	Рекуррентные нейронные сети.	Понятие состояния. Использование состояния. Структура ячейки рекуррентной нейронной сети. Организация входов рекуррентной нейронной сети. Обучение рекуррентной нейронной сети. LSTM и GRU.
4.6	Порождающие состязательные нейронные сети.	Структура порождающих состязательных нейронных сетей. Дискриминатор. Генератор. Обучение сети. Проблемы применения порождающих состязательных нейронных сетей
<i>Содержание лабораторных занятий</i>		
1.	Моделирование и прогнозирование	
1.1	Решение задач классификации	Формулировка задачи и подготовка данных Визуализация данных. Базовая линейная модель. Нелинейный классификатор. Добавление категориальных признаков. Добавление временных признаков.
1.2	Решение задач регрессии	Формулировка задачи и подготовка данных Визуализация данных. Линейная регрессия. Логистическая регрессия. Выбор структуры модели.
2.	Оценка и оптимизация модели	
2.1	Скользкий контроль модели	Использование метода отложенных данных. Использование контроля по k-блокам.
3.	Основы проектирования признаков	
3.1	Использование текстовых признаков и признаков, извлекаемых из изображений.	Модель «Мешок слов». Тематическое моделирование. Расширение содержимого. Простые признаки. Извлечение объектов и форм.
4.	Искусственные нейронные сети в задачах машинного обучения	
4.1	Использование библиотеки Keras для построения нейронных сетей	Назначение библиотеки Keras. Структура Библиотеки Keras. Создание модели. Создание слоя нейронной сети. Компиляция нейронной сети. Обучение нейронной сети. Оценка качества обучения.
4.2	Создание многослойного перцептрона	Чтение исходных данных. Нормализация входных данных. Создание классов выходных данных. Выбор функции активации. Обучение перцептрона. Использование перцептрона для кластеризации данных.
4.3	Создание сверточной сети.	Задача распознавания рукописного ввода. Набор данных MNIST. Ввод и подготовка исходных данных. Выбор архитектуры сверточной нейронной сети. Обучение сети. Исследование количества скрытых параметров сети на точность классификации.
4.4	Создание рекуррентной нейронной сети	Создание рекуррентной сети для генерации текста.

4 ПОРЯДОК ОЦЕНИВАНИЯ УСПЕВАЕМОСТИ И СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ОБУЧАЮЩЕГОСЯ В ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Для положительной оценки по результатам освоения дисциплины обучающемуся необходимо выполнить все установленные виды учебной работы. Оценка результатов работы обучающегося в баллах (по видам) приведена в таблице 7.

Таблица 7 — Балльно-рейтинговая оценка результатов учебной работы обучающихся по видам (БРС)

Учебная работа (виды)	Сумма баллов	Виды и результаты учебной работы	Оценка в аттестации	Баллы (17 недель)
Текущая учебная работа в семестре (Посещение занятий по расписанию и выполнение заданий)	60	Лекционные занятия (конспект) (16 занятий)	1 балл — посещение 1 лекционного занятия	9 - 17
		Лабораторные работы (отчет о выполнении лабораторной работы) (22 работы).	1 балл — посещение 1 практического занятия и выполнение работы на 51-65% 2 балла — посещение 1 занятия и существенный вклад на занятии в работу всей группы, самостоятельность и выполнение работы на 85,1-100%	22 - 44
Итого по текущей работе в семестре				51 - 100
Промежуточная аттестация (экзамен)	40 (100% /баллов приведенной шкалы)	Теоретический вопрос 1	5 баллов (пороговое значение) 10 баллов (максимальное значение)	5 - 10
		Теоретический вопрос 2	5 баллов (пороговое значение) 10 баллов (максимальное значение)	5 - 10
		Выполнение задания	10 баллов (пороговое значение) 20 баллов (максимальное значение)	10 - 20
Итого по промежуточной аттестации (экзамену)				(51 – 100% по приведенной шкале) 20 – 40 б.
Суммарная оценка по дисциплине: Сумма баллов текущей и промежуточной аттестации				51 – 100 б.

5 УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.

5.1 Учебная литература

Основная учебная литература

1. Воронина, В. В. Теория и практика машинного обучения : учебное пособие / В. В. Воронина. — Ульяновск : УлГТУ, 2017. — 290 с. — ISBN 978-5-9795-1712-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165053>. — Режим доступа: для авториз. пользователей.

2. Шарден, Б. Крупномасштабное машинное обучение вместе с Python : учебное пособие / Б. Шарден, Л. Массарон, А. Боскетти ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2018. — 358 с. — ISBN 978-5-97060-506-6. — Текст : электронный //

Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/105836>. — Режим доступа: для авториз. пользователей.

Дополнительная учебная литература

1. Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук ; перевод с английского А. Б. Огурцова. — Москва : ДМК Пресс, 2018. — 250 с. — ISBN 978-5-97060-508-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/97353>. — Режим доступа: для авториз. пользователей.

2. Рашка, С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения : руководство / С. Рашка ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2017. — 418 с. — ISBN 978-5-97060-409-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/100905>. — Режим доступа: для авториз. пользователей.

3. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. — Москва : ДМК Пресс, 2015. — 400 с. — ISBN 978-5-97060-273-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/69955>. — Режим доступа: для авториз. пользователей.

5.2 Материально-техническое и программное обеспечение дисциплины

Таблица 8 — Информационные технологии и программное обеспечение аудиторных занятий и самостоятельной работы

№п/п	Наименование помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом, в том числе помещения для самостоятельной работы, с указанием перечня основного оборудования, учебно-наглядных пособий и используемого программного обеспечения	Адрес (местоположение) помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом (в случае реализации образовательной программы в сетевой форме дополнительно указывается наименование организации, с которой заключен договор)
1	<p>604 Учебная аудитория для проведения:</p> <ul style="list-style-type: none"> - занятий лекционного типа; - занятий семинарского (практического) типа; - групповых и индивидуальных консультаций; - текущего контроля и промежуточной аттестации. <p>Специализированная (учебная) мебель: доска меловая, столы, стулья.</p> <p>Оборудование для презентации учебного материала: переносное - ноутбук, экран, проектор.</p> <p>Используемое программное обеспечение: MS Windows (Microsoft Imagine Premium 3 year по лицензионному договору № 1212/КМР от 12.12.2018 г. до 12.12.2021 г.), LibreOffice (свободно распространяемое ПО), FoxitReader (свободно распространяемое ПО), Firefox 14 (свободно распространяемое ПО), Яндекс.Браузер (отечественное свободно распространяемое ПО).</p>	654079, Кемеровская область, г. Новокузнецк, Центральный район, просп. Metallургов, дом № 19

	Интернет с обеспечением доступа в ЭИОС.	
2	<p>602 Компьютерный класс / Лаборатория информационных систем. Учебная аудитория для проведения:</p> <ul style="list-style-type: none"> - занятий семинарского (практического) типа; - учебных и производственных практик; - курсового проектирования (выполнения курсовых работ); - групповых и индивидуальных консультаций; - самостоятельной работы; - текущего контроля и промежуточной аттестации. <p>Специализированная (учебная) мебель: доска меловая, кафедра, столы компьютерные, стулья.</p> <p>Оборудование: стационарное – компьютеры для обучающихся (17 шт.).</p> <p>Используемое программное обеспечение: MS Windows (Microsoft Imagine Premium 3 year по лицензионному договору № 1212/КМР от 12.12.2018 г. до 12.12.2021 г.), LibreOffice (свободно распространяемое ПО), BloodshedDev C++ 4.9.9.2 (свободно распространяемое ПО), FoxitReader (свободно распространяемое ПО), Firefox (свободно распространяемое ПО), Java (бесплатная версия), Microsoft SQL Server 2008 (Microsoft Imagine Premium 3 year по лицензионному договору № 1212/КМР от 12.12.2018 г. до 12.12.2021 г.), SWI-Prolog (свободно распространяемое ПО), UML-диаграммы (бесплатная версия), Denwer (свободно распространяемое ПО), Eclipse (свободно распространяемое ПО), FreePascal (свободно распространяемое ПО), Lazarus (свободно распространяемое ПО), Pascal ABC.NET (свободно распространяемое ПО), Adobe Reader XI (свободно распространяемое ПО), WinDjView (свободно распространяемое ПО), Python (свободно распространяемое ПО), PyCharm Community Edition (бесплатное ПО).</p> <p>Интернет с обеспечением доступа в ЭИОС.</p>	654079, Кемеровская область, г. Новокузнецк, Центральный район, просп. Metallургов, дом № 19

5.3 Современные профессиональные базы данных и информационные справочные системы.

Перечень СПБД и ИСС по дисциплине

1. 100 репозитория по машинному обучению. <http://meta-guide.com/software-meta-guide/100-best-github-machine-learning>
2. Репозиторий по машинному обучению на Github: <https://github.com/josephmisiti/awesome-machine-learning>
3. Список открытых источников данных, на которых можно найти бесплатные наборы данных: <https://github.com/demidovakatyavvedenie-mashinnoe-obuchenie/blob/master/datasets.md>
4. Календарь соревнований по анализу данных: <http://mltrainings.ru/?filter=active>

6 ИНЫЕ СВЕДЕНИЯ И (ИЛИ) МАТЕРИАЛЫ

6.1 Примерные вопросы и задания / задачи для промежуточной аттестации

Таблица 9 - Примерные теоретические вопросы и задачи к экзамену

	Разделы и темы	Примерные теоретические вопросы	Примерные практические задания
7 семестр			
1	Основы машинного обучения		
1.1	Принятие решений на основе данных	1. В чем разница традиционного подхода к составлению алгоритма и машинным обучением? 2. Какими преимуществами и недостатками обладает машинное обучение в сравнении с традиционным подходом к программированию.	
1.2	Рабочий процесс машинного обучения	3. В чем заключаются сбор и подготовка данных для машинного обучения? 4. Как оценивается производительность модели машинного обучения?	1. Сохраните предоставленные данные в формате CSV. 2. Напишите процедуру для чтения данных из файла в формате CSV.
1.3	Способы повышения эффективности обучения.	5. С какой целью выполняется предварительная обработка данных? 6. В чем состоит сущность масштабирования модели?	3. Напишите процедуру, сохраняющую данные с указанной структурой в формате XML. 4. Напишите процедуру для чтения данных с указанной структурой в формате XML.
1.4	Сбор данных	7. Исходя из чего определяется достаточный объем обучающих данных? 8. Что такое «репрезентативность выборки»?	5. Напишите процедуру, получающую данные из набора данных cifar10 6. Напишите процедуру, получающую данные из набора данных по регрессии цен на жилье в Бостоне.
1.5	Подготовка данных к моделированию	9. Что такое «Категориальные признаки»? 10. В чем суть проектирования признаков? 11. Что такое «Нормализация данных»?	7. Напишите процедуру, нормализующую представленные данные. 8. Напишите процедуру, преобразующую тензор второго порядка в тензор первого порядка.
2	Инструментальные средства машинного обучения		
2.1	Командная оболочка IPython	12. Назначение оболочки IPython 13. Преимущества оболочки IPython перед стандартным интерпретатором Python.	9. Измерьте время выполнения указанного фрагмента кода. 10. Получить значение переменной во время выполнения кода.
2.2	Библиотека NumPy в задачах машинного обучения	14. Назначение библиотеки NumPy 15. Основные компоненты библиотеки NumPy.	11. Установите библиотеку NumPy и подключите её к вашей программе. 12. Средствами библиотеки NumPy создайте переменную одинарной точности.
2.3	Массивы в библиотеке NumPy.	16. Способы создания массивов в библиотеке NumPy.	13. Создайте массив и заполните его значениями, загруженными из

		17. Операции над массивами в библиотеке NumPy.	изображения в градациях серого. 14. Нормализуйте указанный массив, приведя его значения к диапазону [0; 1]
2.4	Основы пакета Pandas	18. Назначение пакета Pandas. 19. Основные компоненты пакета Pandas.	15. Создайте объект DataFrame. 16. Создайте объект Series.
2.5	Визуализация данных	20. Основные компоненты библиотеки matplotlib. 21. Объектно-ориентированный и процедурный способы работы с библиотекой matplotlib.	17. Постройте график функции $y = x^2$. 18. По приведенным данным постройте диаграмму рассеяния.
8 семестр			
1	Моделирование и прогнозирование		
1.1	Основы моделирования с машинным обучением		
1.2	Классификация	22. Задача классификации. Построение классификатора и получение предсказаний. 23. Классификация в случае множественных классов.	19. По представленным данным постройте для бинарной классификации. 20. По представленным данным постройте для множественной классификации.
1.3	Регрессия.	24. Задача регрессии. Линейная модель. 25. Нелинейные модели регрессии.	21. По представленным данным постройте модель линейной регрессии. 22. По представленным данным постройте модель логистической регрессии.
2	Оценка и оптимизация модели		
2.1	Оценка прогностической точности на новых данных.	26. Понятие «переобучение». Борьба с переобучением. 27. Скользящий контроль.	
2.2	Оценка моделей классификации	28. Способы оценки точности модели классификации. 29. ROC-кривые.	
2.3	Оценка моделей регрессии	30. Показатели эффективности регрессионной модели. 31. Исследование остатков.	23. Оцените эффективность регрессионной модели. 24. Для предложенной модели выполните исследование остатков.
2.4	Оптимизация моделей путем подбора параметров.	32. Параметры настройки алгоритмов машинного обучения. 33. Сеточный поиск.	
3	Основы проектирования признаков		
3.1	Назначение проектирования признаков	34. Суть задачи проектирования признаков. 35. Соотношение задач проектирования признаков и моделирования предметной области.	
3.2	Основные этапы проектирования признаков	36. Состав процедуры первичного анализа признаков. 37. Способы извлечения признаков из неструктурированного текста.	25. Для задачи машинного обучения выполните первичный анализ признаков. 26. Предложите признаки, которые могут быть извлечены из представленного неструктурированного текста.
3.3	Выбор признаков	38. Прямой отбор. 39. Обратное исключение.	
4	Искусственные нейронные сети в задачах машинного обучения		
4.1	Понятие искусственных нейронных сетей	40. Структура искусственного нейрона	

		41. Обобщенная структура нейронной сети.	
4.2	Обратное распространение	42. Использование ошибки для изменения весов. 43. Факторы выбора скорости обучения.	
4.3	Оптимизаторы.	44. Стратегии обновления весов. 45. Варианты реализации градиентного спуска.	27. Создайте нейронную сеть для классификации заемщиков по их надежности. 28. Создайте нейронную сеть для классификации отзывов на кинофильмы по их типу (положительный/отрицательный)
4.4	Сверточные нейронные сети.	46. Сущность свертки. 47. Многомерная свертка.	29. Составьте и обучите сверточную нейронную сеть для распознавания изображений из набор MNIST. 30. Составьте и обучите сверточную нейронную сеть для распознавания изображений из набор FASHION-MNIST.
4.5	Рекуррентные нейронные сети.	48. Структура ячейки нейронной сети. 49. Обучение рекуррентной нейронной сети.	31. Создайте и обучите рекуррентную нейронную сеть для предсказания погоды. 32. Составьте нейронную сеть для предсказания изменения цен на недвижимость в пригородах Бостона.
4.6	Порождающие состязательные нейронные сети.	50. Устройство порождающих состязательных нейронных сетей. 51. Обучение порождающих состязательных нейронных сетей.	