

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»
(ВлГУ)**

Институт информационных технологий и радиоэлектроники



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
«Машинное обучение»

направление подготовки / специальность
09.04.01 «Информатика и вычислительная техника»

направленность (профиль) подготовки
Инженерия искусственного интеллекта

г. Владимир
2022

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Машинное обучение» является освоение студентами основных вопросов теории вероятности, методов оптимизации и стохастических процессов для дальнейшего применения в разработке алгоритмов машинного обучения.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина «Машинное обучение» относится к обязательной части учебного плана.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения ОПОП (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине, в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	
ОПК-1. Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	ОПК-1.1. Знать: математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности. ОПК-1.2. Уметь: решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний. ОПК-1.3. Иметь навыки: теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте.	Знает: математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности. Умеет: решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний. Имеет навыки: теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте.	Тестовые вопросы и задания, задания для контрольной работы, задания для самостоятельной работы, вопросы и задания для зачета с оценкой и экзамена
ПК-3. Способен разрабатывать и применять методы и алгоритмы машинного	ПК-3.1. Знать: ПК-3.1.1. классы методов и алгоритмов машинного обучения ПК-3.1.2. методы и	Знает: классы методов и алгоритмов машинного обучения; методы и критерии оценки качества моделей машинного	Тестовые вопросы и задания, задания для контрольной работы, задания для

обучения для решения задач искусственного интеллекта	<p>критерии оценки качества моделей машинного обучения</p> <p>ПК-3.2. Уметь: ПК-3.2.1. ставить задачи и разрабатывать новые методы и алгоритмы машинного обучения ПК-3.2.2. определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области</p> <p>ПК-3.3. Иметь навыки: ПК-3.3.1. по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области</p>	<p>обучения</p> <p>Умеет: ставить задачи и разрабатывать новые методы и алгоритмы машинного обучения; определять критерии и метрики оценки результатов моделирования при построении систем искусственного интеллекта в исследуемой области</p> <p>Имеет навыки: по разработке или совершенствованию методов и алгоритмов для решения комплекса задач предметной области</p>	самостоятельной работы, вопросы и задания для зачета с оценкой и экзамена
--	---	---	---

4. ОБЪЕМ И СТРУКТУРА ДИСЦИПЛИНЫ

Трудоемкость дисциплины составляет 8 зачетных единиц, 288 часов

Тематический план форма обучения – очная

№ п/п	Наименование тем и/или разделов/тем дисциплины	Семестр	Неделя семестра	Контактная работа обучающихся с педагогическим работником				Самостоятельная работа	Формы текущего контроля успеваемости, форма промежуточной аттестации (по семестрам)
				Лекции	Практические занятия	Лабораторные работы	в форме практической подготовки		
1	История машинного обучения и базовые понятия. Данные. Методы уменьшения размерности	2	1-2	2		2	1	12	
2	Кластеризация	2	3-4	2		2	1	12	
3	Регрессия. Классификация	2	5-6	2		2	1	12	Рейтинг-контроль №1
4	Ближайшие соседи.	2	7-8	2		2	1	12	

5	Байесовские методы	2	9-10	2		2	1	12	
6	Метод опорных векторов	2	11-12	2		2	1	12	Рейтинг-контроль №2
7	Деревья Решений	2	13-14	2		2	1	12	
8	Ансамблевые методы	2	15-16	2		2	1	12	
9	Применение методов машинного обучения	2	17-18	2		2	1	12	Рейтинг-контроль №3
Всего за 2 семестр:				18		18		108	Зачет с оценкой
1	Основы Компьютерного Зрения. Компьютерное Зрение: Базовые Операции. Выделение Объектов	3	1-2	2	2		1	8	
2	Особые Точки на Изображениях	3	3-4	2	2		1	8	
3	Применение Машинного обучения в Компьютерном Зрении	3	5-6	2	2		1	8	Рейтинг-контроль №1
4	Нейронные Сети. Архитектуры Нейронных сетей для классификации изображений	3	7-8	2	2		1	8	
5	Архитектуры Нейронных Сетей для Поиска Объектов	3	9-10	2	2		1	8	
6	Генеративные модели Нейронных Сетей	3	11-12	2	2		1	8	Рейтинг-контроль №2
7	Введение в обработку естественного языка	3	13-14	2	2		1	8	
8	Вероятностные Модели в обработке естественного языка	3	15-16	2	2		1	8	
9	Применение нейронных сетей для обработки естественного языка	3	17-18	2	2		1	8	Рейтинг-контроль №3
Всего за 3 семестр:				18	18			72	Экзамен, 36
Наличие в дисциплине КП/КР									
Итого по дисциплине				36	18	18		180	Зачет с оценкой, Экзамен (36)

Содержание лекционных занятий по дисциплине

2 семестр:

1. История машинного обучения и базовые понятия

Определение машинного обучения (МО). Развитие МО: основные исторические этапы.

Классификация задач в МО. Базовые понятия в МО.

Данные

Типы данных. Представление данных. Предварительная обработка. Базы данных

Методы уменьшения размерности

Матрица ковариации. Метод Главных Компонент. Сингулярное разложение Матрицы.

2. Кластеризация

Метрики расстояния. Кластеризация k-Средних (k-Means). Иерархическая кластеризация. Кластеризация DBSCAN

3. Регрессия

Линейная Регрессия. Метод наименьших квадратов. Градиентный спуск. Регуляризация. Метрики моделей регрессии.

Классификация

Типы задач классификации. Логистическая регрессия. Метрики классификации. Матрица ошибок

4. Ближайшие соседи.

Классификатор k-ближайших соседей (k-nearest neighbors). Регрессия k-ближайших соседей. Neighborhood Component Analysis. Визуализация данных методом t-SNE

5. Байесовские методы

Теорема Байеса. Наивный Байесовский классификатор. Дискриминантный Анализ. Линейный дискриминант Фишера

6. Метод опорных векторов

Опорные вектора. Зазор (margin). Ядра. Kernel Trick. Применение метода опорных векторов в задачах классификации и регрессии.

7. Деревья Решений

Применение деревьев решений для решения задач классификации и регрессии. Основные элементы деревьев решений.

8. Ансамблевые методы

Методы усреднения. Бэггинг. Случайный Лес (Random Forest). Методы Бустинга. AdaBoost. Градиентный бустинг

9. Применение методов машинного обучения

Получение Данных. Предварительная Обработка. Отбор значимых параметров (feature selection). Выбор Модели. Оценка Модели. Настройка модели (fine-tuning). Анализ Модели

3 семестр:

1. Основы Компьютерного Зрения

Цели и задачи компьютерного зрения. Цвет и его восприятие. Цветовые модели. Регистрация изображений. Библиотека OpenCV/

Компьютерное Зрение: Базовые Операции

Коррекция артефактов изображений. Фильтры и Свертка. Преобразование Фурье.

Выделение Объектов

Бинаризация. Поиск границ. Distance Transform. Сегментация. Поиск по шаблону

2. Особые Точки на Изображениях

Определение особых точек. Детекторы и Детекторы. Детектор Харриса. Scale-Invariant Feature Transform. Speeded-Up Robust Features. Features from Accelerated Segment Test. Binary Robust Independent Elementary Features. Oriented FAST and Rotated BRIEF

3. Применение Машинного обучения в Компьютерном Зрении

Задачи Машинного обучения в Компьютерном зрении. Мешок Слов. Визуальный словарь. Viola–Jones object detection framework.

4. Нейронные Сети

Введение в нейронные сети. Описание TensorFlow. Полносвязные (Dense) нейронные сети. Функции потерь. Оптимизаторы. Регуляризация Нейронных сетей. Сверточные (Convolutional) нейронные сети. Pooling

Архитектуры Нейронных сетей для классификации изображений

LeNet-5. AlexNet. VGG. GoogLeNet. ResNet. MobileNet. EfficientNet. Перенос обучения (Transfer Learning)

5. Архитектуры Нейронных Сетей для Поиска Объектов

Классификация с локализацией. Region Based Convolutional Neural Networks. Архитектуры семейства YOLO. Single Shot MultiBox Detector. CenterNet.

6. Генеративные модели Нейронных Сетей

Генеративно-состязательные сети. Энкодеры. Декодеры. Вариационные Авто-Энкодеры. Перенос Стиля.

7. Введение в обработку естественного языка

Цели и задачи обработки естественного языка. Базовые понятия. Предварительная обработка текста. Sentiment Analysis. Векторные модели.

8. Вероятностные Модели в обработке естественного языка

Автокоррекция слов. Марковские цепи. Автозаполнение. N-граммы. Языковые Модели.

9. Применение нейронных сетей для обработки естественного языка

Последовательные модели. Embedding. Recurrent Neural Network. Gated Recurrent Units. Long Short-Term Memory. Внимание (Attention). Трансформеры

Содержание лабораторных занятий по дисциплине

2 семестр:

1. Векторные и матричные операции с использованием библиотеки NumPy
2. Работа с наборами данных с использованием библиотеки Pandas
3. Разложение матриц с использованием метода главных компонент и матрицы разложения по сингулярным значениям
4. Кластеризация: k-Means, DBSCAN, иерархическая кластеризация
5. Реализация алгоритма линейной регрессии
6. Реализация алгоритма логистической регрессии
7. Классификация и регрессия машины опорных векторов
8. K-Классификация и регрессия ближайших соседей
9. Визуализация данных с использованием анализа компонент окрестностей
10. Визуализация данных T-SNE
11. Классификация данных с использованием наивного байесовского классификатора

12. Классификация данных с использованием дискриминантного анализа
13. Визуализация данных с использованием линейного дискриминанта Фишера
14. Классификация и регрессия данных с использованием Деревьев решений
15. Классификация и регрессия данных с использованием методов ансамбля

Содержание практических занятий по дисциплине

3 семестр:

1. Предварительная обработка изображений и удаление артефактов
2. Обнаружение объектов: бинаризация, сегментация, поиск по шаблону
3. Детекторы и дескрипторы: SIFT, SURF, ORB
4. Методы мешка слов для классификации изображений
5. Обнаружение объектов Виолы–Джонса
6. Плотные нейронные сети в TensorFlow
7. Сверточные нейронные сети в TensorFlow
8. Передача обучения в TensorFlow
9. Обнаружение объектов нейронными сетями в TensorFlow
10. Передача стиля в TensorFlow
11. Векторные модели в обработке естественного языка
12. Модели автоматической коррекции
13. Генерация текста с использованием рекуррентных нейронных сетей
14. Обучение передаче трансформаторов

5. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

5.1. Текущий контроль успеваемости

2 семестр:

Рейтинг-контроль №1

1. Определение машинного обучения (МО).
2. Развитие МО: основные исторические этапы.
3. Классификация задач в МО.
4. Базовые понятия в МО.
5. Типы данных.
6. Представление данных.
7. Предварительная обработка.
8. Базы данных
9. Матрица ковариации.
10. Метод Главных Компонент.
11. Сингулярное разложение Матрицы.
12. Метрики расстояния.
13. Кластеризация k-Средних (k-Means).
14. Иерархическая кластеризация.
15. Кластеризация DBSCAN
16. Линейная Регрессия.
17. Метод наименьших квадратов.
18. Градиентный спуск.
19. Регуляризация.

20. Метрики моделей регрессии.
21. Типы задач классификации.
22. Логистическая регрессия.
23. Метрики классификации.
24. Матрица ошибок

Рейтинг-контроль №2

1. Классификатор k-ближайших соседей (k-nearest neighbors).
2. Регрессия k-ближайших соседей.
3. Neighborhood Component Analysis.
4. Визуализация данных методом t-SNE
5. Теорема Байеса.
6. Наивный Байесовский классификатор.
7. Дискриминантный Анализ.
8. Линейный дискриминант Фишера
9. Опорные вектора.
10. Зазор (margin).
11. Ядра.
12. Kernel Trick.
13. Применение метода опорных векторов в задачах классификации и регрессии.

Рейтинг-контроль №3

1. Применение деревьев решений для решения задач классификации и регрессии.
2. Основные элементы деревьев решений.
3. Методы усреднения.
4. Бэггинг.
5. Случайный Лес (Random Forest).
6. Методы Бустинга.
7. AdaBoost.
8. Градиентный бустинг
9. Получение Данных.
10. Предварительная Обработка.
11. Отбор значимых параметров (feature selection).
12. Выбор Модели.
13. Оценка Модели.
14. Настройка модели (fine-tuning).
15. Анализ Модели

3 семестр:

Рейтинг-контроль №1

1. Цели и задачи компьютерного зрения.
2. Цвет и его восприятие.
3. Цветовые модели.
4. Регистрация изображений. Библиотека OpenCV/
5. Коррекция артефактов изображений.
6. Фильтры и Свертка.
7. Преобразование Фурье.
8. Бинаризация.
9. Поиск границ.
10. Distance Transform.

11. Сегментация.
12. Поиск по шаблону
13. Определение особых точек.
14. Детекторы и Дескрипторы.
15. Детектор Харриса.
16. Scale-Invariant Feature Transform.
17. Speeded-Up Robust Features.
18. Features from Accelerated Segment Test.
19. Binary Robust Independent Elementary Features.
20. Oriented FAST and Rotated BRIEF
21. Задачи Машинного обучения в Компьютерном зрении.
22. Мешок Слов.
23. Визуальный словарь.
24. Viola–Jones object detection framework.

Рейтинг-контроль №2

1. Описание TensorFlow.
2. Полносвязные (Dense) нейронные сети.
3. Функции потерь.
4. Оптимизаторы.
5. Регуляризация Нейронных сетей.
6. Сверточные (Convolutional) нейронные сети.
7. Pooling
8. LeNet-5.
9. AlexNet.
10. VGG.
11. GoogLeNet.
12. ResNet.
13. MobileNet.
14. EfficientNet.
15. Перенос обучения (Transfer Learning)
16. Классификация с локализацией.
17. Region Based Convolutional Neural Networks.
18. Архитектуры семейства YOLO.
19. Single Shot MultiBox Detector.
20. CenterNet.
21. Генеративно-состязательные сети.
22. Энкодеры.
23. Декодеры.
24. Вариационные Авто-Энкодеры.
25. Перенос Стиля.

Рейтинг-контроль №3

1. Цели и задачи обработки естественного языка.
2. Базовые понятия обработки естественного языка. Предварительная обработка текста.
3. Sentiment Analysis.
4. Векторные модели.
5. Автокоррекция слов.
6. Марковские цепи.
7. Автозаполнение.
8. N-граммы.

9. Языковые Модели.
10. Последовательные модели.
11. Embedding.
12. Recurrent Neural Network.
13. Gated Recurrent Units.
14. Long Short-Term Memory.
15. Внимание (Attention).
16. Трансформеры

5.2. Промежуточная аттестация по итогам освоения дисциплины (зачет с оценкой, экзамен)

Примерный перечень вопросов к зачету с оценкой (2 семестр):

Билет на зачет с оценкой состоит из 5 вопросов, по одному на тематику.

Тематика	Тема	Вопрос
1 Основные Понятия	1.1 Тип задач машинного обучения	Предоставьте все необходимые определения. Приведите примеры вариантов использования
	1.2 Типы данных	
	1.3 Недостаточная и чрезмерная подгонка	
	1.4 Градиентный спуск	
	1.5 Перекрестная проверка	
	1.6 Матрица ошибок и метрики классификации	
	1.7 Показатели регрессии	
	1.8 Предварительная обработка данных	
2 Кластеризация	2.1 Кластеризация k-means	Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
	2.2 Иерархическая кластеризация	
	2.3 DBSCAN	
3 Уменьшение Размерности	3.1 Анализ основных компонентов	Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
	3.2 Разложение по сингулярным значениям	
	3.3 Анализ компонентов	

	Окрестностей	
	3.4 Визуализация данных с помощью t-SNE	
	3.5 Линейный дискриминантный анализ	
4 Регрессия	4.1 Линейная регрессия	Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
	4.2 Регуляризация линейной регрессии	
	4.3 Регрессия по k-Ближайшим соседям	
	4.4 Регрессия по Деревьям принятия решений	
	4.5 Регрессия с помощью машин опорных векторов	
	4.6 Регрессия по AdaBoost	
	4.7 Регрессия с помощью градиентного спуска	
	4.8 Регрессия по Случайному лесу	
5 Классификация	5.1 Логистическая регрессия	Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
	5.2 Классификация по k-Ближайшим соседям	
	5.3 Наивный байесовский классификатор	
	5.4 Дискриминантный анализ (интерпретация Фишера)	
	5.5 Дискриминантный анализ (байесовская версия)	
	5.6 Классификация по деревьям принятия решений	
	5.7 Классификация с помощью машин опорных векторов	

	(мягкое и жесткое поле)	
	5.8 Классификация с помощью машин опорных векторов (трюк с ядром)	
	5.9 Классификация по AdaBoost	
	5.10 Классификация по градиентному спуску	
	5.11 Классификация по Случайному лесу	

Пример билета:

1. Матрица ошибок и метрики классификации. Предоставьте все необходимые определения. Приведите примеры вариантов использования
2. Иерархическая кластеризация. Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
3. Визуализация данных с помощью t-SNE. Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
4. Линейная регрессия. Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
5. Классификация по AdaBoost. Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы

Примерный перечень вопросов к экзамену (3 семестр):

Билет на экзамен состоит из 5 вопросов, по одному на тематику.

Тематика	Тема	Вопрос
1 Основные концепции	1.1 Тип задач машинного обучения	Предоставьте все необходимые определения. Приведите примеры вариантов использования
	1.2 Предварительная обработка данных	
	1.3 Недостаточная и чрезмерная подгонка	
	1.4 Градиентный спуск	
	1.5 Перекрестная проверка	
	1.6 Матрица ошибок и метрики классификации	
	1.7 Показатели регрессии	

2 Классические алгоритмы машинного обучения	2.1 PCA и SVD	Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
	2.2 Кластеризация k-means	
	2.3 Affinity Propagation	
	2.4 Иерархическая кластеризация	
	2.5 DBSCAN	
	2.6 Линейная регрессия	
	2.7 Логистическая регрессия	
	2.8 k-Nearest Neighbors	
	2.9 t-SNE	
	2.10 Анализ компонент Окрестностей	
	2.11 Наивный байесовский классификатор	
	2.12 Дискриминантный анализ	
	2.13 Схемы принятия решений	
	2.14 Машины опорных векторов	
	2.15 AdaBoost	
	2.16 Градиентный спуск	
	2.17 Случайный лес	
3 Компьютерное зрение	3.1. Цветовые модели	Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
	3.2. Регистрация изображений	
	3.3. Исправление артефактов изображения	
	3.4. Применение фильтров для обработки изображений	
	3.5. Соответствие шаблону	

	3.6. Обнаружение границ на изображении	
	3.7. Сегментация изображений	
	3.8. Характерные точки на изображении и детектор Харриса	
	3.9. Детектор и дескриптор просеивания	
	3.10. Детектор и дескриптор серфинга	
	3.11. КРАТКИЙ детектор и дескриптор	
	3.12. Метод визуального словаря и набора слов	
	3.13. Алгоритм Виолы-Джонса для распознавания лиц.	
4 Нейронные сети	4.1 Основные элементы плотных нейронных сетей: нейроны, функции активации, веса и смещения	Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
	4.2 Основные элементы сверточных нейронных сетей: свертки, заполнение, шаг, объединение	
	4.3 Оптимизаторы нейронных сетей: градиентный спуск, стохастический градиентный спуск, экспоненциально взвешенные средние, RMSProp, Adam	
	4.4 Особенности архитектуры нейронных сетей для классификации изображений	
	4.5 свертка 1x1 и начальная архитектура	

	нейронной сети Google для классификации изображений	
	4.6 Остаточные блоки и архитектура нейронной сети ResNet для классификации изображений	
	4.7 Разделяемая свертка по глубине и архитектура нейронной сети мобильной сети для классификации изображений	
	4.8 Классификация изображений с локализацией	
	4.9 Эволюция архитектур YOLO для обнаружения объектов	
	4.10 Сети R-CNN для обнаружения объектов	
	4.11 Генеративно-Состязательные сети	
	4.12 Автокодеры и Вариационные автокодеры	
	4.13 Передача стиля с использованием нейронных сетей	
5 Обработка естественного Языка	5.1 Предварительная обработка текста на Python	Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
	5.2 Анализ настроений с помощью машинного обучения	
	5.3 Векторная Пространственная Модель	
	5.4 Наивный Машинный Перевод	
	5.5 Автозамена и	

	редактирование расстояния	
	5.6 Маркировка части речи по марковским моделям	
	5.7 Модели автозаполнения и языка N-Грамм	
	5.8 Рекуррентные нейронные сети	
	5.9 Закрытые повторяющиеся единицы и Долговременная Кратковременная память	
	5.10 Трансформаторы: внимание, многоголовое внимание, Кодеры и декодеры	
	5.11 Передача обучения в NLP и языковых моделях	

Пример билета

1. Матрица ошибок и метрики классификации. Предоставьте все необходимые определения. Приведите примеры вариантов использования
2. Градиентный спуск. Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
3. Детектор и дескриптор просеивания. Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
4. Остаточные блоки и архитектура нейронной сети ResNet для классификации изображений. Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы
5. Наивный Машинный Перевод. Перечислите основные этапы применения метода, укажите его сильные и слабые стороны, возможности и проблемы

5.3. Самостоятельная работа обучающегося

Самостоятельная работа обучающихся заключается в самостоятельном изучении отдельных тем, практической реализации заданий самостоятельных работ по этим темам, выполнении контрольных работ. Контроль выполнения самостоятельной работы проводится при текущих контрольных мероприятиях и на промежуточной аттестации по итогам освоения дисциплины. Учебно-методическое обеспечение самостоятельной работы – основная литература [1-5], дополнительная литература [1-2].

2 семестр:

Контрольная работа №1

1. История машинного обучения и основные концепции

Опишите разницу между подходом машинного обучения и традиционным программированием.

Опишите разницу между обучением под наблюдением и обучением без присмотра

2. Данные

Приведите пару примеров непрерывных и дискретных данных.

У вас есть три матрицы A, B, C: A имеет размеры 5×4 , B имеет размеры 4×6 , C имеет размеры 3×5 . Пожалуйста, напишите все возможные матрицы, которые можно умножить, и запишите размеры результата (-ов)

3. Методы матричного разложения

Объясните, почему необходимо выполнить стандартизацию данных перед применением анализа основных компонентов

В принципе Компонентный анализ что означают собственные значения и собственные векторы ковариационной матрицы?

При разложении по сингулярным значениям что означают матрицы U, S и V?

4. Кластеризация

Какова связь между евклидовым расстоянием и расстоянием Минковского?

Какой гиперпараметр вам нужно настроить для алгоритма кластеризации k-средних?

Может ли коэффициент силуэта быть равен отрицательному числу? Если "да" - в каких случаях, если "нет" - почему?

Иерархическая кластеризация: в чем разница между различными связями?

DBSCAN: какие точки считаются Шумовыми, Пограничными, Основными?

5. Регрессия

Каковы основные различия между обычными наименьшими квадратами и градиентным спуском для нахождения коэффициентов регрессии?

В каком случае Среднеквадратичная логарифмическая ошибка является более подходящей метрикой, чем Среднеквадратичная ошибка?

Может ли коэффициент детерминации (оценка R^2) быть равен отрицательному числу? Если "да" - в каких случаях, если "нет" - почему?

Почему регуляризация L1 может привести к уменьшению функций (в отличие от регуляризации L2)?

6. Классификация

В чем основное различие между задачами классификации и задачами регрессии?

Предположим, что результаты теста мистера С. пришли на коронавирус. Тест дал положительный ответ, хотя на самом деле у мистера С. нет коронавируса. Какую классификационную ошибку допустил тест?

Допустим, есть два классификатора: первый классификатор имеет Точность 95%, Чувствительность 99%, Специфичность 50%; второй классификатор имеет Точность 87%, Чувствительность 84%, Специфичность 94%. Что вы можете сказать о данных, используемых для классификации? Какой из этих классификаторов более надежен?

Как можно построить поверхность принятия решений для логистической регрессии?

7. Nearest Neighbors

В чем основное различие в использовании k-Ближайших соседей для классификации и регрессии

Как можно найти оптимальное значение гиперпараметра k для методов ближайших соседей?

Как можно уменьшить размерность с помощью анализа ближайших компонент?
Какой гиперпараметр в реализации t-SNE связан с балансом между локальными и глобальными аспектами структуры данных?

8. Байесовские методы

Какую информацию вам необходимо получить, чтобы использовать теорему Байеса?
Что означает "Наивный" в Наивном байесовском классификаторе?
Почему линейный дискриминантный анализ может быть использован в качестве метода уменьшения размерности?
В чем разница между Линейным и Квадратичным дискриминантным анализом?

9. Машины опорных векторов

Какие точки считаются опорными векторами (для задач классификации и регрессии)?
В чем разница между SVM с жесткой маржей и SVM с мягкой маржей?
Почему уловка ядра помогает улучшить результаты SVM?

10. Деревья принятия решений

Каковы основные элементы Дерева решений?
Как выбираются наиболее оптимальные узлы принятия решений?
В чем разница между использованием Деревьев решений для классификации и регрессии?

11. Ensemble methods

В чем разница между методами повышения и методами усреднения?
Что такое "слабая оценка" в контексте методов ансамбля?
В чем разница в объединении деревьев для Случайного леса, Градиентного спуска и AdaBoost?

12. Лучшие практики применения машинного обучения

Какова основная идея методов фильтрации для выбора объектов?
Каков основной принцип, лежащий в основе методов обертки для выбора функций?
Каковы необходимые концепции, которые вам необходимы для успешного применения Генетического алгоритма для выбора признаков?

Самостоятельная работа №1. Обработка и визуализация данных в библиотеке Pandas
Примерные задания:

- выполните предварительную обработку данных обучающего набора данных
- выполните базовую визуализацию обучающего набора данных

Самостоятельная работа №2. Алгоритмы кластеризации и визуализации данных в библиотеке sklearn

Примерные задания:

- кластеризуйте обучающий набор данных с помощью метода k-средних
- кластеризуйте обучающий набор данных с помощью метода DBSCAN
- кластеризуйте обучающий набор данных с помощью метода иерархической кластеризации
- визуализируйте обучающий набор данных с помощью Neighborhood Component Analysis
- визуализируйте обучающий набор данных с помощью метода t-SNE

Самостоятельная работа №3. Алгоритмы регрессии в библиотеке sklearn

Примерные задания:

- a. проанализировать регрессии для набора данных обучения с использованием линейной регрессии
- b. проанализировать регрессии для набора данных обучения, используя метод ближайших соседей
- c. проанализировать регрессии для набора данных обучения с использованием опорных векторов
- d. проанализировать регрессии для набора данных обучения с использованием деревьев решений
- e. проанализировать регрессии для набора данных обучения, используя ensemble methods

Самостоятельная работа №4. Алгоритмы классификации в библиотеке sklearn

Примерные задания:

- a. проанализировать классификации для набора данных обучения с использованием логистической регрессии
- b. проанализировать классификации для набора данных обучения с использованием метода ближайшего соседа
- c. проанализировать классификацию для набора обучающих данных с помощью метода опорных векторов
- d. проанализировать классификации для набора данных обучения с использованием деревьев решений
- e. проанализировать классификации для набора данных обучения, используя ensemble methods

3 семестр:

Контрольная работа №2

1. Основы компьютерного зрения

В чем принципиальная разница между цветовыми моделями RGB и CMYK?

В чем удобство вспомогательных моделей (HSI, HSV) по сравнению с RGB?

В какой цветовой модели библиотека OpenCV (cv2) открывает изображения по умолчанию?

2. Основные операции компьютерного зрения

Какой из фильтров может быть представлен в виде свертки, а какой нет?

Каково максимальное значение яркости для 4-битного изображения?

3. Обнаружение объектов

Какой метод бинаризации следует использовать, чтобы получить исходный объект на черном фоне?

Как вы можете получить общий градиент изображения с помощью вертикальных и горизонтальных фильтров?

Почему алгоритм Кэнни использует гистерезис (верхний и нижний пороги)?

Что такое Преобразование расстояния по существу?

На какой идее основан алгоритм сегментации водораздела?

4. Ключевые точки на изображениях

В чем разница между детектором и дескриптором?

Как можно использовать вторую производную гауссова для прогнозирования масштаба большого двоичного объекта?

Какие функции используются в качестве дескриптора в методе SIFT?

Какие методы объединил метод ORB? Почему их нельзя полностью использовать по отдельности?

5. Применение машинного обучения в задачах компьютерного зрения

Какие проблемы компьютерного зрения можно решить с помощью методов машинного обучения?

Опишите основные этапы применения метода Пакета слов к классификации изображений? Какие задачи машинного обучения выполняются на разных этапах?

Каковы ключевые моменты в алгоритме Виолы Джонс?

6. Плотные Нейронные Сети

В чем основное отличие нейросетевого подхода от "классического" машинного обучения?

Каковы ключевые моменты алгоритма обратного распространения при обучении нейронной сети?

Почему необходимо использовать функции активации в нейронных сетях?

Почему при обучении нейронной сети рекомендуется разбивать исходную выборку на мини-пакеты?

Что означает прямая связь в контексте нейронных сетей?

7. Сверточные Нейронные Сети

Какие ключевые методы используются, чтобы избежать переобучения в нейронных сетях?

В чем заключается идея оптимизации RMSProp?

Каковы обучаемые параметры при использовании сверточных слоев?

Для чего используются слои объединения?

Каков размер вывода нейронной сети, если исходное изображение было 30x30, оно подается на сверточный слой с ядром 3x3, объединяется с ядром 2x2 и другим сверточным слоем 3x3?

8. Архитектуры нейронных сетей для классификации изображений

Что делает свертка 1x1 и каковы веса, которые изучаются во время обучения нейронной сети?

Какая особенность архитектуры ResNet позволила преодолеть проблему исчезающего градиента?

Как количество весов и операций сохраняется в архитектуре MobileNet?

В каком модуле библиотеки tensorflow находятся готовые предварительно обученные архитектуры нейронных сетей?

9. Архитектуры нейронных сетей для обнаружения объектов

В чем особенность Классификации изображений с Локализацией?

Что показывает параметр IOU (Пересечение через объединение)?

Как работает немаксимальное подавление в алгоритме YOLO?

Почему YOLO 9000 был лучше, быстрее, сильнее?

10. Модели Генеративных Нейронных Сетей

В чем разница между Генеративной и Дискриминационной моделями машинного обучения?

Каков основной принцип генеративных состязательных сетей (GAN)?

Каковы основные блоки автокодеров?

В чем особенность Вариационных автокодеров по сравнению с базовыми

автокодерами?

11. Введение в обработку естественного языка

В чем заключается идея использования стемминга в качестве метода предварительной обработки?

В чем заключается обоснование наивного машинного перевода с использованием векторных моделей?

Каковы основные идеи модели векторного пространства Word2Vec?

12. Вероятностные модели в обработке естественного языка

Как алгоритм автозамены балансирует между Кандидатом, которому требуется меньше правок, и Кандидатом, который чаще встречается в Корпусе?

В чем разница между простыми цепочками Маркова и Скрытыми цепочками Маркова в задачах обработки естественного языка?

Почему увеличение N в N -граммовых моделях помогает улучшить качество генерируемого текста? Как вы можете это измерить?

13. Применение нейронных сетей в обработке естественного языка

Что на самом деле делает слой встраивания в библиотеке Tensorflow?

Какова основная идея Рекуррентных нейронных сетей?

Каковы основные улучшения в Закрытых периодических единицах?

Каковы ключевые моменты в слоях долговременной кратковременной памяти?

Каковы Ключи, Значения и Запросы в слоях внимания?

Почему Многоголовое внимание улучшает результаты Трансформаторов?

Самостоятельная работа №5. Машинное обучение для задач компьютерного зрения

Примерные задания:

- выполните предварительную обработку изображения различными методами
- выполните обнаружение объектов с помощью базовых инструментов компьютерного зрения
- выполните обнаружение ключевых точек
- примените машинное обучение для классификации изображений

Самостоятельная работа №6. Применение нейронных сетей в рамках TensorFlow

Примерные задания:

- выполнить классификацию изображений с использованием плотных нейронных сетей
- выполнить классификацию изображений с использованием сверточных нейронных сетей
- применить обучение передаче предварительно обученных сверточных нейронных сетей для классификации изображений
- применить обучение передаче предварительно обученных сетей для обнаружения объектов
- выполнить передачу стиля для изображений

Самостоятельная работа №7. Задачи обработки естественного языка

Примерные задания в составе СРС:

- выполнить предварительную обработку текста
- выполнить анализ настроений учебного корпуса с помощью классических методов машинного обучения
- применить вероятностные модели для исправления опечаток

- d. выполнить анализ настроений учебного корпуса с помощью нейронных сетей
- e. выполнять генерацию текста с помощью рекуррентных нейронных сетей
- f. применять обучение передаче предварительно обученных сетей трансформаторов для обработки естественного языка

Фонд оценочных материалов (ФОМ) для проведения аттестации уровня сформированности компетенций обучающихся по дисциплине оформляется отдельным документом.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. Книгообеспеченность

Наименование литературы: автор, название, вид издания, издательство	Год издания	КНИГООБЕСПЕЧЕННОСТЬ
		Наличие в электронном каталоге ЭБС
Основная литература		
1. Цитульский Антон Максимович, Иванников Александр Владимирович, Рогов Илья Сергеевич NLP - Обработка естественных языков // StudNet. 2020. №6	2020	https://cyberleninka.ru/article/n/nlp-obrabotka-estestvennyh-yazykov
2. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / Флах П. - Москва : ДМК Пресс, 2015. - 400 с. - ISBN 978-5-97060-273-7. - Текст : электронный. Режим доступа : по подписке.	2015	https://www.studentlibrary.ru/book/ISBN9785970602737.html
3. Паттерсон, Дж. , Гибсон А. Глубокое обучение с точки зрения практика / Паттерсон Дж. , Гибсон А. , пер. с англ. А. А. Слинкина. - Москва : ДМК Пресс, 2018. - 418 с. - ISBN 978-5-97060-481-6. - Текст : электронный. Режим доступа : по подписке.	2018	https://www.studentlibrary.ru/book/ISBN9785970604816.html
4. Шапиро, Л. Компьютерное зрение : учебное пособие / Л. Шапиро, Д. Стокман ; под редакцией С. М. Соколова ; перевод с английского А. А. Богуславского. — 4-е изд. — Москва : Лаборатория знаний, 2020. — 763 с. — ISBN 978-5-00101-696-0. — Текст : электронный. Режим доступа: для авториз. пользователей.	2020	https://e.lanbook.com/book/135496
5. Селянкин, В. В. Компьютерное зрение. Анализ и обработка изображений : учебник для вузов / В. В. Селянкин. — 2-е изд., стер. — Санкт-Петербург : Лань, 2021. — 152 с. — ISBN 978-5-8114-8259-7. — Текст : электронный. Режим доступа: для авториз. пользователей.	2021	https://e.lanbook.com/book/173806
Дополнительная литература		
1. Дэвенпорт, Т. Внедрение искусственного	2021	https://www.studentlibrary.ru/book/ISBN

интеллекта в бизнес-практику. Преимущества и сложности / Т. Дэвенпорт. - Москва : Альпина Паблишер, 2021. - 316 с. - ISBN 978-5-9614-3952-6. - Текст : электронный. Режим доступа : по подписке.		9785961439526.html
2. Берджесс, Э. Искусственный интеллект - для вашего бизнеса : Руководство по оценке и применению / Э. Берджесс. - Москва : Интеллектуальная Литература, 2021. - 232 с. - ISBN 9-785-907274-81-5. - Текст : электронный. Режим доступа : по подписке.	2021	https://www.studentlibrary.ru/book/ISBN9785907274815.html

6.2. Периодические издания

1. Вестник компьютерных и информационных технологий ISSN 1810-7206.
2. Цифровая библиотека научно-технических изданий Института инженеров по электротехнике и радиоэлектронике (Institute of Electrical and Electronic Engineers (IEEE)) на английском языке – <http://www.ieee.org/ieeexplore>

6.3. Интернет-ресурсы

1. Academic Search Ultimate EBSCO publishing – <http://search.ebscohost.com>
2. eBook Collections Springer Nature – <https://link.springer.com/>
3. Гугл Академия – <https://scholar.google.ru/>
4. Электронно-библиотечная система «Лань» – <https://e.lanbook.com/>
5. Университетская библиотека ONLINE – <https://biblioclub.ru/>
6. Электронно-библиотечная система "Библиокомплектатор" (IPRbooks) <http://www.bibliocomplectator.ru/available>
7. Электронные информационные ресурсы Российской государственной библиотеки <https://www.rsl.ru/>
8. Научная электронная библиотека «КиберЛенинка» <https://cyberleninka.ru/>
9. Портал российского образования www.edu.ru
10. Портал российских электронных библиотек www.elbib.ru
11. Научная электронная библиотека www.eLibrary.ru
12. Научная библиотека ВлГУ library.vlsu.ru
13. Электронная библиотечная система ВлГУ <https://vlsu.bibliotech.ru/>
14. Web of Science Core Collection – <http://apps.webofknowledge.com/>
15. Курс Methods of Machine Learning <https://elearn.urfu.ru/enrol/index.php?id=5960>

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Для реализации данной дисциплины имеются специальные помещения для проведения занятий: занятий лекционного и лабораторного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы. Лабораторные работы проводятся в компьютерном классе, оборудованном мультимедийным проектором с экраном и обеспеченным доступом в Интернет.

Перечень используемого лицензионного программного обеспечения:

- Операционная система Microsoft Windows 10
- Офисный пакет Microsoft Office 2016
- Браузеры Google Chrome и Mozilla Firefox

Рабочую программу составил Куликов К.В. зав. каф. ВТиСУ
(ФИО, должность, подпись)



Рецензент

(представитель работодателя) _____ Генеральный директор ООО "Диagramma" Протягов И.В.



Программа рассмотрена и одобрена на заседании кафедры ВТ и СУ
Протокол № 1 от 29 августа 2022 года
Заведующий кафедрой Куликов К.В. _____



Рабочая программа рассмотрена и одобрена
на заседании учебно-методической комиссии направления 09.04.01 информатика и
вычислительная техника
Протокол № 1 от 29 августа 2022 года
Председатель комиссии Куликов К.В. зав. каф. ВТиСУ _____



**ЛИСТ ПЕРЕУТВЕРЖДЕНИЯ
РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ**

Рабочая программа одобрена на 20____ / 20____ учебный года

Протокол заседания кафедры № _____ от _____ года

Заведующий кафедрой _____

Рабочая программа одобрена на 20____ / 20____ учебный года

Протокол заседания кафедры № _____ от _____ года

Заведующий кафедрой _____

Рабочая программа одобрена на 20____ / 20____ учебный года

Протокол заседания кафедры № _____ от _____ года

Заведующий кафедрой _____

Рабочая программа одобрена на 20____ / 20____ учебный года

Протокол заседания кафедры № _____ от _____ года

Заведующий кафедрой _____

Рабочая программа одобрена на 20____ / 20____ учебный года

Протокол заседания кафедры № _____ от _____ года

Заведующий кафедрой _____

Рабочая программа одобрена на 20____ / 20____ учебный года

Протокол заседания кафедры № _____ от _____ года

Заведующий кафедрой _____

Рабочая программа одобрена на 20____ / 20____ учебный года

Протокол заседания кафедры № _____ от _____ года

Заведующий кафедрой _____

ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ

в рабочую программу дисциплины

Машинное обучение

образовательной программы направления подготовки *09.04.01 «Информатика и вычислительная техника»*, направленность: *Инженерия искусственного интеллекта (магистратура)*

Номер изменения	Внесены изменения в части/разделы рабочей программы	Исполнитель ФИО	Основание (номер и дата протокола заседания кафедры)

Заведующий кафедрой _____ / _____