

АННОТАЦИЯ рабочей программы дисциплины

Б1.В.10 Глубокое обучение

Объем трудоемкости: 2 з.е.

Цели дисциплины: определены государственным образовательным стандартом высшего образования и соотнесены с общими целями ООП ВО по направлениям подготовки, в рамках которых преподается дисциплина. Дисциплина посвящена овладению современными архитектурами и алгоритмами глубокого обучения, обеспечивающими эффективное решение широкого спектра задач компьютерного зрения, обработки естественного языка и синтеза/моделирования данных.

Задачи дисциплины: в соответствии с поставленной целью состоят в следующем:

- Ознакомить студентов с основами глубокого обучения (архитектуры, типы задач, борьба с переобучением, техника transfer learning, fine-tuning нейронной сети и т.д.).
- Освоить создание, настройку и оценку моделей глубокого обучения на практике.
- Научить работать с библиотеками и фреймворками Keras, PyTorch и Tensorflow для целей глубокого обучения.
- Подготовить специалистов, способных применить глубокое обучение в профессиональных проектах и исследованиях.

Место дисциплины в структуре ООП ВО

Дисциплина «Глубокое обучение» относится к части, формируемой участниками образовательных отношений Блока 1 «Дисциплины (модули)» учебного плана. Данная дисциплина в качестве прerreквизитов требует освоения следующих дисциплин: «Математические модели нейронных сетей» (основные алгоритмы обучения, функции активации, оптимизаторы, начальные знания об интерпретируемости моделей, работа в jupyter notebook и google colab, библиотека scikit-learn), «Нейросетевые технологии» (аугментация, ансамблевые методы, автокодировщики, простые архитектуры для работы с текстами и изображениями, метрики качества, начальные сведения о работе tensorflow и keras), «Технологии компьютерного зрения» (начальные сведения о работе с изображениями, фильтрации и задачах в CV, работа с gitlab, библиотека opencv), «Математический анализ» (основы дифференциального исчисления, интегральное исчисление и ряды, теория функций нескольких переменных), «Обработка данных на Python» (основы Python, библиотеки numpy, pandas, matplotlib, seaborn), «Курс теории вероятностей» (распределения, матожидание, дисперсия и др. важные понятия), «Алгебра и геометрия» (операции с векторами и матрицами). Постреквизитами выступают дисциплины, где продолжается применение принципов глубокого обучения: «Технологии обработки языка, звуковых данных, включая распознавание и синтез речи» (архитектуры для работы с текстами и аудио через SpeechRecognition, torchaudio, librosa, Hugging Face Transformers, FastSpeech), «Генеративный искусственный интеллект» (более глубокое освоение трансформеров и GAN архитектур, работа с несколькими модальностями). Смежными дисциплинами, изучаемыми параллельно или логически близко, являются «MLOps&DevOps» (внедрение нейросетей), «Технологии обработки больших данных» (работа с большими датасетами), «Рекомендательные системы» (нейросети в задачах предоставления рекомендаций).

Требования к уровню освоения дисциплины

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

Код, уровень и формулировка компетенции	Индикаторы	Уровни освоения индикаторов компетенции
DL-1 Способен применять и (или) разрабатывать архитектуры глубоких нейронных сетей	DL-1.3 Способен применять современные архитектуры глубоких сетей для решения различных задач, понимая их внутреннюю структуру и особенности обучения.	(С) Применяет принцип построения вычислительного блока Google Inception; Применяет принцип работы остаточного блока в ResNet; разрабатывает решения с применением базовых (backbone) сетей; знает отличия и способен применять нейронные сети для отслеживания объектов (семейство R-CNN, YOLO)
	DL-1.5 Разрабатывает и оптимизирует специализированные архитектуры для работы с последовательностями, учитывая их уникальные	(С) Понимает принцип работы блока единичной задержки в рекуррентных нейронных сетях; применяет блоки RNN, GRU, LSTM для решения задач
	DL-1.6 Разрабатывает, адаптирует и внедряет генеративные нейронные сети для решения практических задач, включая создание новых архитектур, оптимизацию обучения и промышленное развертывание моделей	(С) Модифицирует архитектуры под специфические требования; разрабатывает комплексные пайплайны обучения; применяет продвинутое техники - перенос обучения, генерация по нескольким примерам
	DL-1.7 Разрабатывает, оптимизирует и применяет автоэнкодеры (АЕ) и вариационные автоэнкодеры (VAE) для решения задач снижения размерности, генерации данных и обнаружения аномалий, включая создание архитектур, обучение моделей и их внедрение в продуктивную среду	(С) Применяет различные автокодировщики для решения задач; настраивает сложные архитектуры (глубокие АЕ, условные VAE); оптимизирует функции потерь, такие как комбинация СКО + KL-дивергенция. Работает с разными типами данных (изображения, временные ряды, графы).
	DL-1.9 Разрабатывает, адаптирует и внедряет трансформерные архитектуры для решения задач обработки последовательностей, включая создание новых моделей, оптимизацию обучения и промышленное развертывание	(С) Понимает и применяет специфические архитектурные варианты для задач (например, вариант "энкодер-декодер" для перевода/суммаризации, вариант "только энкодер" для генерации); использует для работы с длинными последовательностями эффективные схемы внимания (Linformer, Performer, Longformer, Sparse Transformers). Проводит детальный анализ ошибок модели, понимает типичные

		проблемы (смещение в данных/модели, переобучение) и применяет методы борьбы с ними (регуляризация, аугментация данных)
DL-2 Способен применять и (или) разрабатывать современные архитектуры генеративных глубоких сетей	DL-2.1 Применяет известные архитектуры генеративных глубоких нейронных сетей для решения прикладной задачи (генерация текста, генерация изображений по тексту, синтез речи и т.д.), при необходимости проводя дообучение на наборах данных	(С) Настраивает параметры генерации под конкретную задачу. Реализует техники контролируемой генерации (CFG Scale, промпт-инжиниринг). Комбинирует несколько моделей в пайплайн. Проводит доменную адаптацию для специфических данных. Использует продвинутые методы эффективного обучения (QLoRA, DreamBooth). Оптимизирует процесс обучения (подбор скорости обучения и пакетов).
	DL-2.2 Имплементирует известные архитектуры генеративных сетей (VAE, GAN, диффузионные модели)	(С) Кастомизирует архитектуры под задачи (изменение латентного пространства в GAN). Реализует сложные пайплайны обучения (многостадийное обучение). Применяет продвинутые техники дообучения (LoRA, DreamBooth). Настраивает распределенное обучение (DDP, DeepSpeed). Оптимизирует память и скорость обучения (контрольные точки градиента).
DL-3 Способен применять и (или) разрабатывать алгоритмы, методы и технологии компьютерного зрения	DL-3.1 Применяет (проводя выбор и эксперименты) известные алгоритмы и библиотеки компьютерного зрения, предобученные глубокие нейросетевые модели для прикладных задач анализа изображений и видеопотока, при необходимости дообучая и валидируя на собственных наборах данных	(С) Сравнивает разные предобученные модели под конкретную задачу. Проводит перенос обучения на своих данных. Оптимизирует гиперпараметры для улучшения качества. Создает сложные пайплайны аугментации (с использованием библиотеки Albumentations). Умеет работать с видео: извлечение кадров, обработка временных последовательностей путём применения CNN+RNN, 3D CNN.
LLM-2 Способен дообучать, адаптировать и	LLM-2.1 Понимает принципы дообучения	(С) Применяет дообучение к предобученным моделям на новых датасетах

оптимизировать генеративные модели под специфические задачи и условия применения	LLM-2.2 Создаёт обучающие наборы данных	(С) Адаптирует и валидирует датасеты под разные типы задач
--	---	--

Содержание дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

№	Наименование разделов (тем)	Всего	Количество часов			
			Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	СРС
1	2	3	4	5	6	7
1.	Введение в глубокое обучение	3	2			1
2.	Архитектуры глубоких нейронных сетей	11	4		4	3
3.	Алгоритмы обучения глубоких сетей	12,8	4		6	2,8
4.	Глубокие сети в компьютерном зрении	20	2		12	6
5.	Глубокие сети в обработке естественного языка	9	2		4	3
6.	Оптимизация и повышение производительности	14	2		6	6
ИТОГО по разделам дисциплины		69,8	16		32	21.8
Контроль самостоятельной работы (КСР)		2				
Промежуточная аттестация (ИКР)		0.2				
Подготовка к текущему контролю						
Общая трудоемкость по дисциплине		72				

Курсовые работы: не предусмотрены

Форма проведения аттестации по дисциплине: зачет.

Автор: Кадури А.А. – доцент КАДИИ; Казаковцева Е.В. – к. ф.-м. н., доцент КАДИИ