

Аннотация к рабочей программы дисциплины  
Б1.В.07 «Обработка естественного языка»

**Объем трудоемкости:** 4 зачетных единиц

**Цель дисциплины:**

Формирование у студентов базовых знаний и практических навыков в области обработки естественного языка (NLP) с акцентом на применение глубокого обучения, включая архитектуры LLM и трансферное обучение.

**Задачи дисциплины**

Изучить математические основы NLP (векторизация, вероятностные модели).

Освоить архитектуры нейросетей для NLP (RNN, Transformer, BERT, GPT).

Научиться применять предобученные модели (Hugging Face) для классификации, генерации и анализа текста.

Разрабатывать и оптимизировать NLP-пайплайны (дообучение, RAG, развертывание).

**Место дисциплины в структуре образовательной программы**

Дисциплина «Обработка естественного языка» относится к «Часть, формируемая участниками образовательных отношений» Блока 1 «Дисциплины (модули)» учебного плана.

**Требования к уровню освоения дисциплины**

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

<i>DL-1</i>	<i>Способен применять и (или) разрабатывать архитектуры глубоких нейронных сетей.</i>
DL-1.1	Способен объяснять и применять математические основы нейронных сетей, включая расчет градиентов, методы оптимизации и алгоритм обратного распространения ошибки (backpropagation), для эффективного обучения моделей.
	Знать: математические основы нейронных сетей, включая расчет градиентов, методы оптимизации и алгоритм обратного распространения ошибки Уметь: модифицировать популярные архитектуры графовых нейронных сетей (GCN, GraphSAGE, GIN); применяет графовые сети внимания (GAT). Адаптирует архитектуры под специфические типы графов (гетерогенные, динамические). Разрабатывает кастомные механизмы агрегации. Владеть: методами расчета градиентов и оптимизации нейронных сетей
DL-1.3	Способен применять современные архитектуры глубоких сетей для решения различных задач, понимая их внутреннюю структуру и особенности обучения.
	Знать: современные архитектуры глубоких сетей и их внутреннюю структуру Уметь: применять глубокие сети для решения различных задач. Разрабатывает кастомные архитектуры для специфических комбинаций модальностей. Оптимизирует совместное обучение нескольких задач (multi-task learning). Владеть: методами обучения современных архитектур глубоких сетей. Применяет продвинутые техники (cross-modal distillation, contrastive learning). Создает эффективные пайплайны обработки мультимодальных данных

DL-1.5	Способен разрабатывать, адаптировать и внедрять генеративные нейронные сети для решения практических задач, включая создание новых архитектур, оптимизацию обучения и промышленное развертывание моделей
	<p>Знать: архитектуры генеративных нейронных сетей. Понимает и применяет специфические архитектурные варианты для задач (например, Encoder-Decoder для перевода/суммаризации, Decoder-only для генерации).</p> <p>Уметь: разрабатывать и адаптировать генеративные сети для практических задач. Использует техники для работы с длинными последовательностями: эффективные схемы внимания (LInformer, Performer, Longformer, Sparse Transformers). Проводит детальный анализ ошибок модели, понимает типичные проблемы (bias в данных/модели, переобучение) и применяет методы борьбы с ними (регуляризация, аугментация данных)</p> <p>Владеть: промышленным развертыванием генеративных моделей</p>
DL-1.8	Способен разрабатывать, адаптировать и внедрять трансформерные архитектуры для решения задач обработки последовательностей, включая создание новых моделей, оптимизацию обучения и промышленное развертывание
	<p>Знать: трансформерные архитектуры и их особенности</p> <p>Уметь: адаптировать трансформерные архитектуры для решения задач обработки последовательностей</p> <p>Владеть: промышленным развертыванием трансформерных моделей</p>
<i>DL-2</i>	<i>Способен применять и (или) разрабатывать современные архитектуры генеративных глубоких сетей</i>
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
	<p>Знать: известные архитектуры генеративных глубоких нейронных сетей</p> <p>Уметь: применять генеративные сети для решения прикладных задач. Настраивает параметры генерации под конкретную задачу. Реализует техники контролируемой генерации (guidance scale, prompt engineering). Комбинирует несколько моделей в пайплайн. Проводит domain adaptation для специфических данных. Использует продвинутые методы эффективного обучения (QLoRA, DreamBooth). Оптимизирует процесс обучения (подбор lr, батчей).</p> <p>Владеть: методами дообучения генеративных моделей на наборах данных. Модифицирует архитектуры под специфические требования. Разрабатывает гибридные подходы (например, диффузионные модели + GAN). Оптимизирует архитектуры для целевых аппаратных платформ. Разрабатывает новые методы дообучения для генеративных моделей. Применяет few-shot/zero-shot learning техники. Реализует reinforcement learning для генерации.</p>
DL-2.2	Имплементирует известные архитектуры генеративных сетей, реализует пайплайны их обучения на датасетах и вывод генеративных моделей в продуктивную среду
	<p>Знать: архитектуры генеративных сетей и пайплайны их обучения</p> <p>Уметь: имплементировать архитектуры генеративных сетей. Разрабатывает кастомные реализации генеративных архитектур. Создает high-load генеративные сервисы. Разрабатывает системы кеширования генераций. Реализует сложные А/В тестирования. Написание кастом CUDA-ядер.</p> <p>Владеть: вывода генеративных моделей в продуктивную среду</p>

<b>DL-4</b>	<b><i>Способен применять и (или) разрабатывать алгоритмы, методы и технологии обработки естественного языка</i></b>
DL-4.1	Имплементирует известные алгоритмы, архитектуры и модели обработки естественного языка на реальных данных, строит пайплайны обучения моделей и развертывания NLP-сервисов в продуктивной среде
	Знать: алгоритмы, архитектуры и модели обработки естественного языка Уметь: имплементировать NLP-модели на реальных данных. Владеет инструментами грамматического разбора структурированных и слабо-структурированных текстов, способен написать свой парсер. Владеет инструментами разметки текстовых данных и формирования словарей. Владеть: развертыванием NLP-сервисов в продуктивной среде
DL-4.2	Определяет стек технологий, методов и алгоритмов для построения продуктов с обработкой естественного языка (диалоговые системы, вопросно-ответные системы, рекомендательные системы и т.д.)
	Знать: основные архитектуры сетей, использующиеся для векторизации текстовых данных: Word2Vec, Doc2Vec, Glove, FastText, рекуррентные нейронные сети и сети-трансформеры (энкодеры). Уметь: применять простейшие классические модели для извлечения полезных признаков из текстов: мешок слов, TF/IDF, латентный семантический анализ. Владеть: самостоятельно находит подходящую модель для векторизации текстовых данных в открытых источниках и применить её для практической задачи.
DL-4.3	Имплементирует известные алгоритмы, архитектуры и модели обработки естественного языка на реальных данных, строит пайплайны обучения моделей и развертывания NLP-сервисов в продуктивной среде
	Знать: классические алгоритмы тематического моделирования и способен применить их в практической задаче: LSA, pLSA, BigARTM. Уметь: обучать модель классификации текстовых данных с использованием векторизованного представления текста - разреженного в случае применения классических статистических алгоритмов ML для классификации и плотного для применения в нейросетевых моделях) Владеть: архитектурой и применяет рекуррентные нейронные сети и сети типа "трансформер" для downstream задач, требующих языкового моделирования (контекстуализированные векторные представления, генерация текстов, аннотирование и реферирование текстов и т.п.)
DL-4.4	Разрабатывает новые алгоритмы и библиотеки обработки естественного языка, новые архитектуры глубоких нейронных сетей и методы их обучения для задач анализа текста
	Знать: архитектуру рекуррентных нейронных сетей и сетей типа "трансформер" для downstream задач Уметь: применяет рекуррентные нейронные сети и сети типа "трансформер" для downstream задач, требующих языкового моделирования (контекстуализированные векторные представления, генерация текстов, аннотирование и реферирование текстов и т.п.) Владеть: Использует предобученные модели без модификации архитектуры.
<b>LLM-1</b>	<b><i>Способен применять и (или) разрабатывать генеративные модели и БЯМ</i></b>
LLM-1.1	Знает архитектуры генеративных моделей.
	Знать: архитектуры генеративных моделей Уметь: различать архитектуры генеративных моделей. Сравнивает архитектуры и выбирает подходящую под задачу

	Владеть: знаниями об архитектурах генеративных моделей
<b>FC-2</b>	<b>Способен проводить фронтальные исследования в области фундаментальных и генеративных моделей</b>
FC-2.3	Исследует и создает мульти-модальные большие языковые модели (LLM)
	Знать: архитектуры мульти-модальных больших языковых моделей Уметь: исследовать и создавать мульти-модальные LLM Владеть: созданием мульти-модальных языковых моделей

### Содержание дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.  
Разделы (темы) дисциплины, изучаемые в 6 семестре

№	Наименование разделов (тем)	Количество часов				Внеаудиторная работа СРС
		Всего	Аудиторная работа			
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1.	Введение в NLP	5	2		2	1
2.	Лингвистические основы NLP	5	2		2	1
3.	Предобработка текста	6	2		2	2
4.	Векторизация текста: TF-IDF, Word2Vec	8	2		4	2
5.	Рекуррентные сети (RNN, LSTM)	8	2		4	2
6.	Механизм внимания и Transformer	8	2		4	2
7.	BERT и трансферное обучение	7	2		4	1
8.	Генеративные модели (GPT, T5)	8	2		4	2
9.	Оценка качества моделей (BLEU, ROUGE)	6	2		2	2
10	Дообучение и адаптация LLM	7	2		4	1
11	RAG (Retrieval-Augmented Generation)	7	2		4	1
12	Мультимодальные NLP-модели	8	2		4	2
13	Этика и bias в NLP	5	2		2	1
14	Промышленное внедрение NLP-моделей	5	2		2	1
15	Оптимизация NLP-моделей (квантование, дистилляция)	6	2		2	2
16	Итоговый обзор и перспективы NLP	5	2		2	1
<b>ИТОГО по разделам дисциплины</b>		<b>104</b>	<b>32</b>		<b>48</b>	<b>24</b>
Контроль самостоятельной работы (КСР)		4				
Промежуточная аттестация (ИКР)		0,3				
Подготовка к текущему контролю		35,7				
<b>Общая трудоемкость по дисциплине</b>		<b>144</b>				

**Курсовые работы:** не предусмотрена

**Форма проведения аттестации по дисциплине:** экзамен

Автор

Добровольская Наталья Юрьевна