

МИНОБРНАУКИ РОССИИ



Федеральное государственное автономное образовательное учреждение
высшего образования
«**Российский государственный гуманитарный университет**»
(ФГАОУ ВО «РГГУ»)

ИНСТИТУТ ЛИНГВИСТИКИ
Учебно-научный центр компьютерной лингвистики

Основы глубинного обучения

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

45.04.03 Фундаментальная и прикладная лингвистика

Код и наименование направления подготовки/специальности

Фундаментальная и компьютерная лингвистика

Наименование направленности (профиля)/ специализации

Уровень высшего образования: магистратура

Форма обучения: очная

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2024

Основы глубинного обучения
Рабочая программа дисциплины

Составитель(и):

старший преподаватель А.М.Ивойлова

Ответственный редактор:

к.ф.н, доцент Н.А.Коротаев

УТВЕРЖДЕНО

Протокол заседания УНЦ компьютерной лингвистики

№ 5 от 26 марта 2024 г.

ОГЛАВЛЕНИЕ

1.	Пояснительная записка	
1.1.	Цель и задачи дисциплины	
1.2.	Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций	
1.3.	Место дисциплины в структуре образовательной программы	
2.	Структура дисциплины	
3.	Содержание дисциплины	
4.	Образовательные технологии	
5.	Оценка планируемых результатов обучения	
5.1	Система оценивания	
5.2	Критерии выставления оценки по дисциплине	
5.3	Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине	
6.	Учебно-методическое и информационное обеспечение дисциплины	
6.1	Список источников и литературы	
6.2	Перечень ресурсов информационно-телекоммуникационной сети «Интернет».	
6.3	Профессиональные базы данных и информационно-справочные системы	
7.	Материально-техническое обеспечение дисциплины	
8.	Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов	
9.	Методические материалы	
9.1	Планы семинарских/ практических/ лабораторных занятий	
9.2	Методические рекомендации по подготовке письменных работ	
9.3	Иные материалы	

1. Пояснительная записка

1.1. Цель и задачи дисциплины

Предметом дисциплины является изучение основ глубинного обучения для автоматической обработки естественного языка. В рамках курса студенты знакомятся с основными принципами создания нейронных сетей, с такими понятиями, как нейрон, функция активации, слой нейронной сети; изучаются виды градиентного спуска и функций активации, а также главные типы архитектур нейронных сетей в связи с решаемыми ими задачами. В результате освоения данного курса студенты должны уметь самостоятельно создавать модели нейронных сетей и обучать их для решения задач, связанных с автоматической обработкой естественного языка.

Курс направлен на решение следующих задач:

- познакомить обучающихся с основными понятиями глубинного обучения и с основными архитектурами нейронных сетей, применяемыми для решения лингвистических задач, а также с программными модулями языка Python, реализующими данные методы, в частности, с библиотекой torch;
- познакомить магистрантов с основными задачами автоматической обработки естественного языка и используемыми для их решения архитектурами нейронных сетей;
- познакомить магистрантов с математическими методами, лежащими в основе алгоритмов глубинного обучения;
- научить магистрантов как предварительно выбирать архитектуру нейронной сети для решения для прикладных лингвистических задач, так и дорабатывать выбранную модель в зависимости от специфики задачи и исходных данных;
- научить магистрантов анализировать результаты применения алгоритмов глубинного обучения к лингвистическим данным;
- дать магистрантам знания, позволяющие им квалифицированно читать литературу по специальности, включающую в себя как учебные материалы и научные статьи, так и более специализированные технические материалы, например, программную документацию.

1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-1 Владеет основными методами фонологического, морфологического, синтаксического, дискурсивного и семантического анализа с учетом языковых и экстралингвистических факторов	ПК-1.2 Умеет применять полученные знания в области теории языка для лингвистического анализа с учетом языковых и экстралингвистических факторов; различать основные типы формальных моделей описания естественного языка, формальных	Знать: <ul style="list-style-type: none"> – принципы работы лингвистически ориентированных программных продуктов; Уметь: <ul style="list-style-type: none"> – пользоваться лингвистически ориентированными программными продуктами; Владеть: <ul style="list-style-type: none"> – навыками использования

	грамматик; структурировать и моделировать базовые явления языка	лингвистически ориентированных программных продуктов.
ПК-2 Владеет принципами создания электронных языковых ресурсов (текстовых, речевых и мультимодальных корпусов; словарей, тезаурусов, онтологий; фонетических, лексических, грамматических и иных баз данных и баз знаний) и умением пользоваться такими ресурсами	ПК-2.2 Умеет пользоваться основными методами, способами и средствами получения, хранения, переработки информации; пользоваться лингвистически ориентированными программными продуктами	Знать: – принципы работы лингвистически ориентированных программных продуктов; Уметь: – пользоваться лингвистически ориентированными программными продуктами; Владеть: – навыками использования лингвистически ориентированных программных продуктов.

1.3. Место дисциплины в структуре образовательной программы

Дисциплина «Основы глубинного обучения» является элективной дисциплиной и относится к части, формируемой участниками образовательных отношений.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин и прохождения практик: Основы языка программирования Python, Инструменты лингвистического анализа в Python, Алгоритмы машинного обучения, Основания математики, Основы математической статистики.

2. Структура дисциплины

Общая трудоёмкость дисциплины составляет 3 з.е., 108 академических часов.

Структура дисциплины для очной формы обучения

Объем дисциплины в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Семестр	Тип учебных занятий	Количество часов
3	Практические занятия	30
Всего:		30

Объем дисциплины (модуля) в форме самостоятельной работы обучающихся составляет 78 академических часов.

3. Содержание дисциплины

№ п/п	Наименование раздела дисциплины	Содержание
1.	Введение. Основные понятия глубинного обучения. Структура нейрона. Функция потерь. Градиентный спуск. Функции активации.	<ul style="list-style-type: none"> • Знакомство с основными понятиями глубинного обучения: нейрон, параметры, свободный коэффициент, функция активации, функция потерь, градиентный спуск. Полносвязный слой. • Знакомство с библиотекой torch и ее базовым функционалом. Тип данных torch.Tensor.
2.	Виды градиентного спуска. Многослойный перцептрон.	<ul style="list-style-type: none"> • Знакомство с видами градиентного спуска: SGD, MBGD, адаптивный градиентный спуск: RMSProp, Adagrad, Adam, AdamW. • Написание простейшей архитектуры MLP с использованием библиотеки torch. Решение задач с табличными данными.
3.	Сверточные нейронные сети. Ядра свертки. Работа с изображениями.	<ul style="list-style-type: none"> • Знакомство с архитектурой CNN и понятием свертки. • Написание архитектуры нейронной сети для классификации изображений.
4.	Сверточные нейронные сети для работы с текстами	<ul style="list-style-type: none"> • Одномерная архитектура CNN и извлечение признаков из текстовых данных. • Написание архитектуры нейронной сети для классификации текстов.
5.	Рекуррентные нейронные сети. RNN, LSTM, GRU.	<ul style="list-style-type: none"> • Знакомство с архитектурой RNN и с ее усовершенствованиями: LSTM и GRU. Достоинства и недостатки рекуррентных нейронных сетей. • Языковые модели на LSTM: ELMo. • Написание модели нейронной сети для генерации текста.
6.	Архитектура Seq2Seq. Механизм внимания	<ul style="list-style-type: none"> • Основные виды задач, решаемых нейронными сетями: one-to-one, one-to-many, many-to-one, many-to-many. Задача машинного перевода и другие задачи, требующие генерации нового текста. • Написание модели нейронной сети Seq2Seq для машинного перевода с английского на русский.
7.	Архитектура Transformer. Механизм самовнимания	<ul style="list-style-type: none"> • Механизм самовнимания и устройство архитектуры Transformer: понятия энкодер, декодер, key, query, value. Визуализация голов внимания (attention heads). • Написание модели нейронной сети Transformer для машинного перевода с английского на русский.
8.	Контекстуальные эмбединги и языковые модели, основанные на архитектуре Transformer. Бертология.	<ul style="list-style-type: none"> • Языковые модели BERT, GPT и их разновидности. Маскированное и каузальное языковое моделирование. Основные типы

		задач, решаемые с помощью BERT и GPT.
9.	Генеративно-состязательные сети. Архитектура GAN и ее применение в NLP.	<ul style="list-style-type: none"> Генерация текста. Алгоритм BeamSearch. Архитектура GAN, ее устройство, дискриминатор и генератор. Использование GAN для решения лингвистических задач.

4. Образовательные технологии

Для проведения учебных занятий по дисциплине используются различные образовательные технологии. Для организации учебного процесса может быть использовано электронное обучение и (или) дистанционные образовательные технологии.

5. Оценка планируемых результатов обучения

5.1 Система оценивания

Форма контроля	Макс. количество баллов	
	За одну работу	Всего
Текущий контроль:		
- домашние задания	10 баллов	60 баллов
Промежуточная аттестация – экзамен		40 баллов
Итого за семестр		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55			E
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

5.2 Критерии выставления оценки по дисциплине

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ A,B	отлично/ зачтено	<p>Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.</p> <p>Свободно ориентируется в учебной и профессиональной литературе.</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».
82-68/ С	хорошо/ зачтено	Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей. Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами. Достаточно хорошо ориентируется в учебной и профессиональной литературе. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».
67-50/ D,E	удовлетво- рительно/ зачтено	Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами. Демонстрирует достаточный уровень знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».
49-0/ F,FX	неудовлет- ворительно/ не зачтено	Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами. Демонстрирует фрагментарные знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.

5.3 Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

В качестве домашних заданий предлагаются задания следующих типов

- Д31. Решение задачи классификации на табличных данных с применением архитектуры MLP и библиотеки torch. Ручная реализация градиентного спуска для примитивной архитектуры.
 Д32. Решение задачи классификации изображений с использованием архитектуры CNN.
 Д33. Решение задачи классификации текстов с использованием архитектуры 1D CNN.
 Д34. Обучение примитивной модели для генерации текста (RNN).
 Д35. Решение задачи автоматической морфологической разметки (POS-tagging) с использованием языковой модели BERT и библиотеки transformers.
 Д36. Использование архитектуры GAN для обогащения эмбедингов семантической информацией.

Экзамен ориентирован на следующие контрольные вопросы

Основные понятия глубинного обучения: нейрон, слой, функция потерь, градиентный спуск, функция активации.
 Полносвязный слой. Многослойный перцептрон.
 Виды градиентного спуска. Адаптивный градиентный спуск.
 Архитектура CNN. Ядра свертки.
 Архитектура RNN и основные связанные с ней понятия. Архитектуры LSTM и GRU и разница между ними.
 Задача машинного перевода и архитектура Seq2Seq. Механизм внимания.
 Архитектура Transformer и механизм самовнимания.
 Языковая модель BERT. Языковая модель GPT. Особенности архитектуры, основные различия.
 Виды языкового моделирования в глубинном обучении.
 Архитектура GAN.
 Основные типы решаемых задач автоматической обработки естественного языка и архитектуры, используемые для их решения.

6. Учебно-методическое и информационное обеспечение дисциплины

6.1 Список источников и литературы

Основная литература

1. Хайкин С. Нейронные сети. Полный курс. — М., 2006. — 1104 с.
2. Jurafsky, Daniel, and James H. Martin. "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition." URL: <https://web.stanford.edu/~jurafsky/slp3/>
3. Документация библиотеки torch. URL: <https://pytorch.org/docs/stable/index.html>
4. Документация библиотеки transformers. URL: <https://huggingface.co/docs/transformers/en/index>

Рекомендованная литература

1. Сейновски Т. Антология машинного обучения. Важнейшие исследования в области ИИ за последние 60 лет. — М.: Эксмо, 2022.
2. Mattman Chris A. Machine Learning with TensorFlow. — 2021.
3. Medsker, Larry, and Lakhmi C. Jain, eds. Recurrent neural networks: design and applications. CRC press, 1999.
4. Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017).
5. Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).
6. Conneau, Alexis, and Guillaume Lample. "Cross-lingual language model pretraining." Advances in neural information processing systems 32 (2019).
7. Conneau, Alexis, et al. "Unsupervised cross-lingual representation learning at scale." arXiv preprint arXiv:1911.02116 (2019).
8. Floridi, Luciano, and Massimo Chiriatti. "GPT-3: Its nature, scope, limits, and consequences." Minds and Machines 30 (2020): 681-694.
9. Pikuliak, Matúš, Marián Šimko, and Mária Bieliková. "Cross-lingual learning for text processing: A survey." Expert Systems with Applications 165 (2021): 113765.
10. Jozefowicz, R., Vinyals, O., Schuster, M., Shazeer, N., & Wu, Y. (2016). Exploring the limits of language modeling. arXiv preprint arXiv:1602.02410.

11. Wang, P., Qian, Y., Soong, F. K., He, L., & Zhao, H. (2015). Part-of-speech tagging with bidirectional long short-term memory recurrent neural network. arXiv preprint arXiv:1510.06168.
12. Zhang, Jiajun, and Chengqing Zong. "Deep Neural Networks in Machine Translation: An Overview." IEEE Intell. Syst. 30.5 (2015): 16-25.

6.2 Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

№ п/п	Наименование
1	Международные реферативные наукометрические БД, доступные в рамках национальной подписки в 2020 г. Web of Science Scopus
2	Профессиональные полнотекстовые БД, доступные в рамках национальной подписки в 2020 г. Журналы Cambridge University Press ProQuest Dissertation & Theses Global SAGE Journals Журналы Taylor and Francis
3	Профессиональные полнотекстовые БД JSTOR Издания по общественным и гуманитарным наукам Электронная библиотека Grebennikon.ru

6.3 Профессиональные базы данных и информационно-справочные системы

Доступ к профессиональным базам данных: <https://liber.rsuh.ru/ru/bases>

Документация библиотеки torch: <https://pytorch.org/docs/stable/index.html>

Документация библиотеки transformers: <https://huggingface.co/docs/transformers/en/index>

7. Материально-техническое обеспечение дисциплины

№п /п	Наименование ПО	Производитель	Способ распространения (лицензионное или свободно распространяемое)
1	Microsoft Share Point 2010	Microsoft	лицензионное
2	Windows 10 Pro	Microsoft	лицензионное
3	Kaspersky Endpoint Security	Kaspersky	лицензионное
4	Microsoft Office 2016	Microsoft	лицензионное
5	Zoom	Zoom	лицензионное
6	Python 3.9	Python	Свободно распространяемое
7	Visual Studio Code Community	Microsoft	Свободно распространяемое

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением или могут быть заменены устным ответом; обеспечивается индивидуальное равномерное освещение не менее 300 люкс; для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств; письменные задания оформляются увеличенным шрифтом; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

- для глухих и слабослышащих: лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования; письменные задания выполняются на компьютере в письменной форме; экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.

- для лиц с нарушениями опорно-двигательного аппарата: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих: в печатной форме увеличенным шрифтом, в форме электронного документа, в форме аудиофайла.

- для глухих и слабослышащих: в печатной форме, в форме электронного документа.

- для обучающихся с нарушениями опорно-двигательного аппарата: в печатной форме, в форме электронного документа, в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих: устройством для сканирования и чтения с камерой SARA SE; дисплеем Брайля PAC Mate 20; принтером Брайля EmBraille ViewPlus;

- для глухих и слабослышащих: автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих; акустический усилитель и колонки;

- для обучающихся с нарушениями опорно-двигательного аппарата: передвижными, регулируемые эргономическими партами СИ-1; компьютерной техникой со специальным программным обеспечением.

9. Методические материалы

9.1 Планы семинарских/ практических/ лабораторных занятий

Тема 1.

Знакомство с основными понятиями глубинного обучения: нейрон, параметры, свободный коэффициент, функция активации, функция потерь, градиентный спуск. Полносвязный слой.

Знакомство с библиотекой torch и ее базовым функционалом. Тип данных torch.Tensor.

Тема 2.

Знакомство с видами градиентного спуска: SGD, MBGD, адаптивный градиентный спуск: RMSProp, Adagrad, Adam, AdamW.

Написание простейшей архитектуры MLP с использованием библиотеки torch. Решение задач с табличными данными.

Тема 3.

Знакомство с архитектурой CNN и понятием свертки.

Написание архитектуры нейронной сети для классификации изображений.

Тема 4.

Одномерная архитектура CNN и извлечение признаков из текстовых данных.

Написание архитектуры нейронной сети для классификации текстов.

Тема 5.

Знакомство с архитектурой RNN и с ее усовершенствованиями: LSTM и GRU.

Достоинства и недостатки рекуррентных нейронных сетей.

Языковые модели на LSTM: ELMo.

Написание модели нейронной сети для генерации текста.

Тема 6.

Основные виды задач, решаемых нейронными сетями: one-to-one, one-to-many, many-to-one, many-to-many. Задача машинного перевода и другие задачи, требующие генерации нового текста.

Написание модели нейронной сети Seq2Seq для машинного перевода с английского на русский.

Тема 7.

Механизм самовнимания и устройство архитектуры Transformer: понятия энкодер, декодер, key, query, value. Визуализация голов внимания (attention heads).

Написание модели нейронной сети Transformer для машинного перевода с английского на русский.

Тема 8.

Языковые модели BERT, GPT и их разновидности. Маскированное и каузальное языковое моделирование. Основные типы задач, решаемые с помощью BERT и GPT.

Генерация текста. Алгоритм BeamSearch.

Тема 9.

Архитектура GAN, ее устройство, дискриминатор и генератор. Использование GAN для решения лингвистических задач.

9.2 Иные материалы

Все необходимые для обучения материалы публикуются по адресу <https://github.com/rsuh-python/> в соответствующих репозиториях.