

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»**

УТВЕРЖДЕНО

**Директор физтех-школы
прикладной математики и
информатики**

А.М. Райгородский

	Рабочая программа дисциплины (модуля)
по дисциплине:	Большие языковые модели: обучение и применение
по направлению:	Прикладная математика и информатика
профиль подготовки:	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра машинного обучения и цифровой гуманитаристики
курс:	1
квалификация:	магистр

Семестр, формы промежуточной аттестации: 1 (осенний) - Экзамен

Аудиторных часов: 45 всего, в том числе:

лекции: 30 час.

семинары: 15 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 60 час.

Подготовка к экзамену: 30 час.

Всего часов: 135, всего зач. ед.: 3

Количество контрольных работ, заданий: 2

Программу составил: Р.Г. Нейчев, старший преподаватель

Программа обсуждена на заседании кафедры машинного обучения и цифровой гуманитаристики 26.02.2024

Аннотация

В курсе рассматриваются большие языковые модели (LLM) в контексте современных методов искусственного интеллекта. Рассматриваются архитектурные особенности языковых моделей: как уже ставшие классикой механизм внимания и сам механизм Transformer, так и перспективные State Space Models. Помимо основ построения языковых моделей, отдельное внимание уделяется их дообучению в реальных условиях: с учетом ограниченных данных и вычислительных ресурсов.

Курс предполагает, что слушатели знакомы с основами машинного и глубокого обучения, линейной алгебры, теории вероятностей и программирования. Курс состоит из лекций и практических занятий. В курсе предполагаются как практические задания, так и теоретический экзамен в конце.

1. Цели и задачи

Цель дисциплины

- познакомить с основными свойствами больших языковых моделей (LLM);
- предоставить практический опыт использования LLM для решения задач обработки естественного языка;
- развить навыки применения LLM к реальным прикладным задачам.

Задачи дисциплины

- предоставить набор методов донастройки LLM под конкретные нужды;
- научить выбирать подходящую базовую (foundation) модель для конкретной задачи;
- обеспечить существенный опыт работы с современными LLM.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
	ОПК-2.2 Способен оценить актуальность и практическую значимость прикладных математических исследований в своей профессиональной области
	ОПК-2.3 Владеет профессиональной терминологией, используемой в современной научно-технической литературе, обладает навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации
ПК-1 Готов к включению в профессиональное сообщество; способен проводить под научным руководством локальные исследования на основе существующих методов в конкретной области профессиональной деятельности	ПК-1.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации; владеет навыками подготовки научных обзоров, публикаций рефератов и библиографий по тематике проводимых исследований на русском и английском языке
	ПК-1.2 Умеет решать научные задачи с пониманием существующих подходов к верификации модели программного обеспечения в связи с поставленной целью и в соответствии с выбранной методикой
	ПК-1.3 Имеет практический опыт выступлений на научной аргументации при анализе объекта научной профессиональной деятельности
ПК-2 Понимает и способен применить в научно-исследовательской и прикладной	ПК-2.1 Знает основы научно-исследовательской деятельности в области информационно-коммуникационных технологий; владеет знанием основ философии и методологии науки; знанием методов научных исследований и навыками их проведения

деятельности современный математический аппарат и алгоритмы, основные законы естествознания, современные языки программирования и программное обеспечение; операционные системы и сетевые технологии	ПК-2.2 Имеет практически опыт использования существующих методов и алгоритмов решения задач цифровой обработки сигналов, использования сети Интернет, аннотирования, реферирования, библиографического разыскания и описания, опыт работы с научными источниками
	ПК-2.3 Умеет применять полученные знания в области фундаментальных научных основ теории информации и решать стандартные задачи в собственной научно-исследовательской деятельности

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные подходы к построению языковых моделей;
- способы использования LLM для решения прикладных задач.

уметь:

- формулировать подходящие критерии качества для указанных выше задач;
- оценивать вычислительную сложность используемых подходов на основе LLM;
- решать задачи, требующие совместного использования LLM и методов обучения с подкреплением/с учителем.

владеть:

- основными программными системами для работы с LLM и построения систем с их использованием.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины (модуля) и трудоемкости по видам учебных занятий

№	Тема (раздел) дисциплины	Трудоемкость по видам учебных занятий, включая самостоятельную работу, час.			
		Лекции	Семинары	Лаборат. работы	Самост. работа
1	Задача языкового моделирования	6	3		12
2	Механизм внимания, архитектура Transformer	6	3		12
3	Модели из семейства GPT. Decoder-only модели	6	3		12
4	Доснастройка больших языковых моделей на своих данных	6	3		12
5	Методы обучения с подкреплением в задачах обработки естественного языка	6	3		12
Итого часов		30	15		60
Подготовка к экзамену		30 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

4.2. Содержание дисциплины (модуля), структурированное по темам (разделам)

Семестр: 1 (Осенний)

1. Задача языкового моделирования

Марковское свойство. Кросс-энтропия.

2. Механизм внимания, архитектура Transformer

Связь операции свертки и механизма внимания. Обзор архитектуры Transformer. Задача машинного перевода.

3. Модели из семейства GPT. Decoder-only модели

Эффект масштабирования моделей. Дообучение моделей для следования инструкциям.

4. Доснастройка больших языковых моделей на своих данных

Few-shot и zero-shot подходы. Промптинг моделей. Использование низкоранговых приближений матрицы весов (low-rank approximation). P-tuning. PEFT (parameter-efficient fine-tuning)

5. Методы обучения с подкреплением в задачах обработки естественного языка

RLHF (Reinforcement learning with human feedback). DPO (Direct preference optimization).

5. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Стандартная аудитория, оснащенная мультимедийным оборудованием (проектор, звуковая система).

6. Перечень рекомендуемой литературы

Основная литература

1. Python и машинное обучение [Текст], крайне необходимое издание по новейшей предсказательной аналитике для более глубокого понимания методологии машинного обучения/С. Рашка, -М., ДМК Пресс, 2017
2. Python и машинное обучение, Электронная версия печатной публикации / С. Рашка. — Москва, ДМК Пресс, 2017
3. Машинное обучение [Текст]/Х. Бринк, Дж. Ричардс, М. Феверолф, Real-World Machine Learning, -СПб., Питер, 2017

Дополнительная литература

1. Математические основы машинного обучения и прогнозирования [Текст] / В. В. Вьюгин ; Моск. физ.-техн. ин-т (гос. ун-т), Лаб. структурных методов анализа данных в предсказательном моделировании (ПреМоЛаб), Ин-т проблем передачи информации им. А. А. Харкевича РАН, М., МЦНМО, 2013
2. Математические основы машинного обучения и прогнозирования, Электронная версия печатной публикации / В. В. Вьюгин. — Москва, МЦНМО, 2014

7. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

<http://github.com/girafe-ai>
https://lena-voita.github.io/nlp_course.html

8. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)

На лекционных занятиях используются мультимедийные технологии, включая демонстрацию презентаций.

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Студент, изучающий дисциплину, должен с одной стороны, овладеть общим понятийным аппаратом, а с другой стороны, должен научиться применять теоретические знания на практике.

В результате изучения дисциплины студент должен знать основные определения, понятия, аксиомы.

Успешное освоение курса требует напряжённой самостоятельной работы студента. В программе курса приведено минимально необходимое время для работы студента над темой. Самостоятельная работа включает в себя:

- чтение и конспектирование рекомендованной литературы;
- проработку учебного материала (по учебной и научной литературе), подготовку ответов на вопросы, предназначенных для самостоятельного изучения, доказательство отдельных утверждений, свойств;
- подготовку к экзамену.

Руководство и контроль за самостоятельной работой студента осуществляется в форме индивидуальных консультаций.

Важно добиться понимания изучаемого материала, а не механического его запоминания. При затруднении изучения отдельных тем, вопросов, следует обращаться за консультациями к лектору.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению:	Прикладная математика и информатика
профиль подготовки:	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра машинного обучения и цифровой гуманитаристики
курс:	<u>1</u>
квалификация:	магистр
Семестр, формы промежуточной аттестации: 1 (осенний) - Экзамен	
Разработчик:	Р.Г. Нейчев, старший преподаватель

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-2 Способен совершенствовать и реализовывать новые математические методы решения прикладных задач	ОПК-2.1 Имеет представление о современном состоянии математических исследований в рамках тематической области своей профессиональной деятельности
	ОПК-2.2 Способен оценить актуальность и практическую значимость прикладных математических исследований в своей профессиональной области
	ОПК-2.3 Владеет профессиональной терминологией, используемой в современной научно-технической литературе, обладает навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации
ПК-1 Готов к включению в профессиональное сообщество; способен проводить под научным руководством локальные исследования на основе существующих методов в конкретной области профессиональной деятельности	ПК-1.1 Знает принципы построения научной работы, методы сбора и анализа полученного материала, способы аргументации; владеет навыками подготовки научных обзоров, публикаций рефератов и библиографий по тематике проводимых исследований на русском и английском языке
	ПК-1.2 Умеет решать научные задачи с пониманием существующих подходов к верификации модели программного обеспечения в связи с поставленной целью и в соответствии с выбранной методикой
	ПК-1.3 Имеет практический опыт выступлений на научной аргументации при анализе объекта научной профессиональной деятельности
ПК-2 Понимает и способен применить в научно-исследовательской и прикладной деятельности современный математический аппарат и алгоритмы, основные законы естествознания, современные языки программирования и программное обеспечение; операционные системы и сетевые технологии	ПК-2.1 Знает основы научно-исследовательской деятельности в области информационно-коммуникационных технологий; владеет знанием основ философии и методологии науки; знанием методов научных исследований и навыками их проведения
	ПК-2.2 Имеет практический опыт использования существующих методов и алгоритмов решения задач цифровой обработки сигналов, использования сети Интернет, аннотирования, реферирования, библиографического разыскания и описания, опыт работы с научными источниками
	ПК-2.3 Умеет применять полученные знания в области фундаментальных научных основ теории информации и решать стандартные задачи в собственной научно-исследовательской деятельности

2. Показатели оценивания компетенций

В результате изучения дисциплины «Большие языковые модели: обучение и применение» обучающийся должен:

знать:

- основные подходы к построению языковых моделей;
- способы использования LLM для решения прикладных задач.

уметь:

- формулировать подходящие критерии качества для указанных выше задач;
- оценивать вычислительную сложность используемых подходов на основе LLM;
- решать задачи, требующие совместного использования LLM и методов обучения с подкреплением/с учителем.

владеть:

- основными программными системами для работы с LLM и построения систем с их использованием.

3. Перечень типовых (примерных) вопросов, заданий, тем для подготовки к текущему контролю

С целью контроля освоения обучающимися учебного материала проводится устный опрос в начале занятия по теме прошлой лекции или в конце занятия по пройденной теме.

4. Перечень типовых (примерных) вопросов и тем для проведения промежуточной аттестации обучающихся

Перечень контрольных вопросов для сдачи экзамена:

1. Назовите основные свойства рекуррентных нейронных сетей.
2. Что такое механизм внимания? Какая у него сложность по памяти, по времени?
3. Какие основные блоки включает в себя архитектура Transformer?
4. Как стоит дообучать модель, если количество доступных данных составляет сотни примеров, а модель насчитывает миллиарды параметров?
5. Почему при минимизации кросс-энтропии задача машинного перевода может решаться неоптимальным способом?
6. Каковы различия между CBOW и Skip-gram моделями в контексте Word2Vec?
7. В чем особенности использования предобученных векторных представлений слов по сравнению с обучением эмбеддингов с нуля?
8. Какие стратегии можно применить для уменьшения переобучения в нейронных сетях на задачах NLP?
9. В чем заключается концепция дистилляции моделей (model distillation) и в каких случаях она может быть полезна?
10. Какие базовые векторные представления слов (word embeddings) вы знаете, и в чем заключаются их ключевые отличия? (word2vec, glove, fasttext)
11. В чем принципиальное отличие между моделями на основе правил (rule-based models) и статистическими моделями в NLP?
12. Каковы примеры задач, где предпочтительно использовать однонаправленные RNN, и почему?
13. Что такое переобучение (overfitting) и недообучение (underfitting) в контексте машинного обучения, и как их можно обнаружить при работе с NLP-моделями?
14. Какие методы существуют для оценки качества генерируемого текста в задачах NLP, и в чем их особенности?
15. Что такое гроккинг (grokking) языковых моделей и в каких случаях он наблюдается?
16. Каковы преимущества использования подхода трансферного обучения в NLP?
17. В чем разница между стеммингом и лемматизацией, и как это влияет на предобработку текстовых данных?
18. Какие существуют подходы к решению проблемы неоднозначности слов (word sense disambiguation) в NLP?
19. В чем заключается задача определения тональности (sentiment analysis), и какие методы могут использоваться для её решения?
20. Какое влияние на процесс обучения нейронной сети оказывает размер мини-батча, и как определить оптимальный размер?
21. Как решается проблема длинных зависимостей в LSTM сетях по сравнению с обычными RNN?
22. Что такое BERT и для каких задач обработки естественного языка он наиболее эффективен?

Примеры билетов для экзамена:

Билет 1.

1. Назовите основные свойства рекуррентных нейронных сетей.
2. Опишите основные проблемы механизма внимания.

Билет 2.

1. Каким образом можно применить градиентные методы для задач с недифференцируемой целевой функцией?
2. Почему при минимизации кросс-энтропии задача машинного перевода может решаться неоптимальным способом?

Критерии оценивания

Оценка отлично (10) выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины, проявляющему интерес к данной предметной области, продемонстрировавшему умение уверенно и творчески применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично (9) выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

Оценка отлично (8) выставляется студенту, показавшему систематизированные, глубокие знания учебной программы дисциплины и умение уверенно применять их на практике при решении конкретных задач, правильное обоснование принятых решений, с некоторыми недочетами.

Оценка хорошо (7) выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но недостаточно грамотно обосновывает полученные результаты.

Оценка хорошо (6) выставляется студенту, если он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач некоторые неточности.

Оценка хорошо (5) выставляется студенту, если он в основном знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает в ответе или в решении задач достаточно большое количество неточностей.

Оценка удовлетворительно (4) выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала, но при этом он освоил основные разделы учебной программы, необходимые для дальнейшего обучения, и может применять полученные знания по образцу в стандартной ситуации.

Оценка удовлетворительно (3) выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, допускающему ошибки в формулировках базовых понятий, нарушения логической последовательности в изложении программного материала, слабо владеет основными разделами учебной программы, необходимыми для дальнейшего обучения и с трудом применяет полученные знания даже в стандартной ситуации.

Оценка неудовлетворительно (2) выставляется студенту, который не знает большей части основного содержания учебной программы дисциплины, допускает грубые ошибки в формулировках основных принципов и не умеет использовать полученные знания при решении типовых задач.

Оценка неудовлетворительно (1) выставляется студенту, который не знает основного содержания учебной программы дисциплины, допускает грубейшие ошибки в формулировках базовых понятий дисциплины и вообще не имеет навыков решения типовых практических задач.

5. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Экзамен проводится в устной форме.

При проведении устного экзамена обучающемуся предоставляется 1 час на подготовку. Опрос обучающегося по билету на устном экзамене не должен превышать двух астрономических часов. Во время проведения экзамен обучающиеся не могут пользоваться программой дисциплины.

