

**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:	Дополнительные методы машинного обучения
по направлению:	Прикладная математика и информатика
профиль подготовки:	Комбинаторика и цифровая экономика
	центр дополнительного, дополнительного профессионального и онлайн-образования "Пуск"
	кафедра дискретной математики
курс:	2
квалификация:	магистр

Семестр, формы промежуточной аттестации: 4 (весенний) - Экзамен

Аудиторных часов: 60 всего, в том числе:

лекции: 30 час.

семинары: 30 час.

лабораторные занятия: 0 час.

Самостоятельная работа: 45 час.

Подготовка к экзамену: 30 час.

Всего часов: 135, всего зач. ед.: 3

Программу составил: Д.В. Мусатов, канд. физ.-мат. наук

Программа обсуждена на заседании кафедры дискретной математики 30.01.2025

Аннотация

Целью курса является освоение методов и теории глубокого обучения для решения прикладных задач, в том числе в области компьютерного зрения и обработки естественного языка. Студент после освоения курса будет понимать теорию машинного и глубокого обучения (понятия, аксиомы, методы), современные проблемы соответствующих разделов глубокого обучения, основные библиотеки и программное обеспечение для реализации собственных или типовых проектов, использующих глубокое обучение.

1. Цели и задачи

Цель дисциплины

- освоение методов и теории глубокого обучения для решения прикладных задач, в том числе в области компьютерного зрения и обработки естественного языка.

Задачи дисциплины

- освоение студентами знаний (понятий, концепций, методов и моделей) в области глубокого обучения;
- приобретение теоретических знаний и практических умений и навыков в области глубокого обучения для решения прикладных задач, в том числе и необходимые знания по разработке программного обеспечения;
- оказание консультаций и помощи студентам в проведении собственных исследований в области глубокого обучения.

2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
УК-4 Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия	УК-4.1 Способен вести обмен деловой информацией в устной и письменной формах на государственном языке Российской Федерации и не менее чем на одном иностранном языке
ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности	ОПК-3.2 Владеет исследовательскими методами и способен использовать их при решении новых задач, применяя знания из различных областей науки (техники)
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

3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- теорию машинного и глубокого обучения (понятия, аксиомы, методы);
- современные проблемы соответствующих разделов глубокого обучения;
- основные библиотеки и программное обеспечение для реализации собственных или типовых проектов, использующих глубокое обучение.

уметь:

- понять поставленную задачу и оценить ее корректность;
- использовать свои знания для решения прикладных задач;
- находить алгоритмы решения задач и проводить их анализ;
- анализировать научную литературу по глубокому обучению.

владеть:

навыками освоения большого объема информации и решения задач (в том числе, сложных);
навыками самостоятельной работы и освоения новых дисциплин;
культурой постановки, анализа и решения математических и прикладных задач;
предметным языком глубокого обучения и навыками грамотного описания решения задач и представления полученных результатов.

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	Введение в глубокое обучение	4	4		7
2	Борьба с переобучением в нейронных сетях	4	4		7
3	Свёрточные нейронные сети	4	4		6
4	Рекуррентные нейронные сети	4	4		5
5	Seq2seq модели	4	4		5
6	Механизм внимания	4	4		5
7	Обучение без учителя	3	3		5
8	Генеративные модели	3	3		5
Итого часов		30	30		45
Подготовка к экзамену		30 час.			
Общая трудоёмкость		135 час., 3 зач.ед.			

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

Семестр: 4 (Весенний)

1. Введение в глубокое обучение

Нейронные сети. Обратное распространение ошибки. Стохастический градиентный спуск. Инициализация весов. Реализация с помощью библиотеки Pytorch.

2. Борьба с переобучением в нейронных сетях

Продвинутые методы оптимизации. Оптимизация гиперпараметров. Батч нормализация. Регуляризация. Dropout. Проблема затухания градиента

3. Свёрточные нейронные сети

Задачи с изображениями: классификация, локализация, детектирование, сегментация, семантическая сегментация. Основные архитектуры свёрточных сетей. Transfer learning

4. Рекуррентные нейронные сети

GRU и LSTM. Применение рекуррентных сетей для задач обработки естественного языка. Векторные представления слов. word2vec: CBOW и skip-gram. Языковые модели. Модели GPT-(1,2,3).

5. Seq2seq модели

Seq2seq модели. Задачи перевода, text2speech, распознавание речи.

6. Механизм внимания

Механизм внимания (attention). Трансформеры (Transformer). Модели BERT, RoBERTa.

7. Обучение без учителя

Автоэнкодеры. Вариационный автоэнкодер.

8. Генеративные модели

Generative Adversarial Networks (GAN). Генерация изображений.

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

Учебная аудитория, оснащенная компьютером и мультимедийным оборудованием (проектор, звуковая система).

6.Перечень рекомендуемой литературы

Основная литература

1. Мерков А.Б. “Распознавание образов: введение в методы статистического обучения”. Москва, URSS, 2010.
 2. Bishop C. “Pattern Recognition and Machine Learning”. Springer, 2007.
- Фонд литературы базовой кафедры

Дополнительная литература

1. Хайкин С. “Нейронные сети”. Москва, ИД “Вильямс”, 2006.
- Фонд литературы базовой кафедры

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

1. <http://web.stanford.edu/class/cs224n/>
2. <https://pytorch.org/>

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

Для части занятий потребуется Zoom. Google Drive для доступа к материалам курса. Приветствуется наличие во время занятий смартфонов/ноутбуков для участия в интерактивных упражнениях.

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 дисциплины, уметь применять полученные знания для решения различных задач.

Успешное освоение курса требует:

- посещения всех занятий, предусмотренных учебным планом по дисциплине;
- ведения конспекта занятий;
- напряжённой самостоятельной работы студента.

Самостоятельная работа включает в себя:

- чтение рекомендованной литературы;
- проработку учебного материала, подготовку ответов на вопросы, предназначенных для самостоятельного изучения;
- решение задач, предлагаемых студентам на занятиях;
- подготовку к выполнению заданий текущей и промежуточной аттестации.

Показателем владения материалом служит умение без конспекта отвечать на вопросы по темам дисциплины.

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

Возможен промежуточный контроль знаний студентов в виде решения задач в соответствии с тематикой занятий.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

по направлению: Прикладная математика и информатика
профиль подготовки: Комбинаторика и цифровая экономика
центр дополнительного, дополнительного профессионального и
онлайн-образования "Пуск"
кафедра дискретной математики
курс: 2
квалификация: магистр
Семестр, формы промежуточной аттестации: 4 (весенний) - Экзамен
Разработчик: Д.В. Мусатов, канд. физ.-мат. наук

1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
УК-4 Способен применять современные коммуникативные технологии, в том числе на иностранном(ых) языке(ах), для академического и профессионального взаимодействия	УК-4.1 Способен вести обмен деловой информацией в устной и письменной формах на государственном языке Российской Федерации и не менее чем на одном иностранном языке
ОПК-3 Способен разрабатывать математические модели и проводить их анализ при решении задач в области профессиональной деятельности	ОПК-3.2 Владеет исследовательскими методами и способен использовать их при решении новых задач, применяя знания из различных областей науки (техники)
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

2. Показатели оценивания компетенций

В результате изучения дисциплины «Дополнительные методы машинного обучения» обучающийся должен:

знать:

теорию машинного и глубокого обучения (понятия, аксиомы, методы);
современные проблемы соответствующих разделов глубокого обучения;
основные библиотеки и программное обеспечение для реализации собственных или типовых проектов, использующих глубокое обучение.

уметь:

понять поставленную задачу и оценить ее корректность;
использовать свои знания для решения прикладных задач;
находить алгоритмы решения задач и проводить их анализ;
анализировать научную литературу по глубокому обучению.

владеть:

навыками освоения большого объема информации и решения задач (в том числе, сложных);
навыками самостоятельной работы и освоения новых дисциплин;
культурой постановки, анализа и решения математических и прикладных задач;
предметным языком глубокого обучения и навыками грамотного описания решения задач и представления полученных результатов.

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

Во время текущего контроля студент должен уметь ответить на следующие вопросы:

1. Метод обратного распространения ошибки для обучения нейронных сетей.
2. Реализация нейронных сетей с помощью Pytorch.
3. Реализация операции свертки.
4. Задача классификации/ сегментации изображений.
5. Задача классификации текстов.

Во время занятий могут проходить интерактивные обсуждения в чатах курса, что будет являться домашним заданием. Возможно выполнение патентного поиска в качестве самостоятельной задачи. Успешное выполнение всех заданий по курсу и выполнение контрольных срезов знаний дает преимущество на дифференцированном зачете.

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

1. Метод обратного распространения ошибки для обучения нейронных сетей.
2. Реализация нейронных сетей с помощью Pytorch.
- 3 Реализация операции свертки.
4. Задача классификации/ сегментации изображений.
5. Задача классификации текстов.
6. Батч нормализация.
7. Проблема затухания градиента.
8. Задачи с изображениями: классификация, локализация, детектирование, сегментация, семантическая сегментация.
- 9 Применение рекуррентных сетей для задач обработки естественного языка.
11. Модели BERT, RoBERTa.
12. Генеративные модели.

Критерии оценивания

Оценка отлично (10 баллов) - 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 баллов) - 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 баллов) - выставляется студенту, показавшему всесторонние, 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 баллов) - 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 баллов) - 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 баллов) - 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 балла) - 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 балла) - 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 балла) - 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 балл) - 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

При проведении устной аттестации студенту дается 45 минут на подготовку. Опрос студента по билету на устном дифференцированном зачете не должен превышать одного астрономического часа.