

**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**

	<b>Рабочая программа дисциплины (модуля)</b>
<b>по дисциплине:</b>	Обработка сигналов
<b>по направлению:</b>	Информатика и вычислительная техника
<b>профиль подготовки:</b>	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра интеллектуальных систем
<b>курс:</b>	2
<b>квалификация:</b>	магистр

Семестр, формы промежуточной аттестации: 3 (осенний) - Дифференцированный зачет

Аудиторных часов: 45 всего, в том числе:

лекции: 0 час.

семинары: 30 час.

лабораторные занятия: 15 час.

Самостоятельная работа: 45 час.

Всего часов: 90, всего зач. ед.: 2

Программу составил: П.А. Северилов

Программа обсуждена на заседании кафедры интеллектуальных систем 03.04.2024

## Аннотация

Курс "Обработка сигналов" посвящен современным методам обработки, распознавания, генерации речи и смежными задачами обработки аудиоданных. В рамках курса будет рассматриваться историческое развитие моделей глубинного обучения, включая самые недавние исследования в каждой из областей обработки речи. Также на курсе будут разобраны модели, которые используют в работе крупнейшие технические компании.

### 1. Цели и задачи

#### Цель дисциплины

- знакомство с современными методами обработки, распознавания, генерации речи и смежными задачами обработки аудиоданных.

#### Задачи дисциплины

- рассмотреть историческое развитие моделей глубинного обучения, включая самые недавние исследования в каждой из областей обработки речи;
- разобрать модели, которые используют в работе крупнейшие технические компании;
- дать представление о современном состоянии области;
- поделиться практическими навыками построения эффективных моделей глубокого обучения.

### 2. Перечень формируемых компетенций

Освоение дисциплины направлено на формирование следующих компетенций:

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-1 Владеет системой фундаментальных научных знаний в области информатики и вычислительной техники	ОПК-1.1 Знает и способен использовать в профессиональной деятельности фундаментальные научные знания и новые научные принципы и методы исследований в области информатики и вычислительной техники
	ОПК-1.2 Способен обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности
	ОПК-1.3 Понимает междисциплинарные связи в области информатики и вычислительной техники и способен их применять при решении задач профессиональной деятельности
ПК-2 Понимает и способен применить в научно-исследовательской и прикладной деятельности основные законы естествознания, современный математический аппарат и алгоритмы, современные информационно-коммуникационные технологии	ПК-2.1 Знает основы научно-исследовательской деятельности в области информационных технологий, владеет знанием основ философии и методологии науки; знанием методов научных исследований и навыками их проведения
	ПК-2.2 Умеет применять полученные знания в области фундаментальных научных основ теории информации и решать стандартные задачи в собственной научно-исследовательской деятельности
	ПК-2.3 Имеет практический опыт научно-исследовательской деятельности в области информационно-коммуникационных технологий
	ПК-2.4 Владеет методами и алгоритмами решения задач цифровой обработки сигналов, использования сети Интернет, аннотирования, реферирования, библиографического поиска, опыт работы с научными источниками

### 3. Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- современные подходы распознавания, генерации, обработки аудиосигналов.

уметь:

- анализировать аудиосигналы, внутреннее устройство и результаты работы модели глубокого обучения.

владеть:

- навыками реализации базовых алгоритмов обработки, распознавания, генерации речи и смежными задачами обработки аудиоданных.

#### 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	Цифровая обработка сигналов		6	3	9
2	Автоматическое распознавание речи		6	3	9
3	Синтез речи		6	3	9
4	Улучшение качества аудиосигналов		6	3	9
5	Обучение моделей глубинного обучения без учителя с аудиоданными		6	3	9
Итого часов			30	15	45
Подготовка к экзамену		0 час.			
Общая трудоёмкость		90 час., 2 зач.ед.			

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

Семестр: 3 (Осенний)

###### 1. Цифровая обработка сигналов

Звук: типы, характеристики, теорема Найквиста. Спектрограммы: преобразование Фурье, Short-time преобразование Фурье, Mel спектрограммы.

###### 2. Автоматическое распознавание речи

ASR: постановка задачи, метрики, CTC: алгоритм, проблемы, лосс функция, свойства, beam search, Listen, Attend and Spell (LAS), RNN-T: архитектура модели, inference модели, процесс обучения, преимущества над CTC и LAS, Языковые модели для задачи ASR: мотивация, подходы. Byte pair encoding (BPE): мотивация, алгоритм. Whisper.

###### 3. Синтез речи

Пайплайн решения задачи, метрики. Использование идей механизма внимания в задаче TTS. Современные подходы в решении задачи синтеза речи: DeepVoice, FastSpeech, FastSpeech 2, AdaSpeech.

#### 4. Улучшение качества аудиосигналов

Нейронные вокодеры.

- Wavenet: causal+dilated свертка, обуславливание в модели, решение проблемы большого числа таймстемпов для предсказания.
- Parallel WaveGAN: архитектура, лосс
- DiffWave: диффузионные процессы, обучение модели, сэмплирование из модели, преимущества.

#### 5. Обучение моделей глубинного обучения без учителя с аудиоданными

- Обучение без учителя в аудио:
- Авторегрессионные модели (APC)
- Сиамские модели (BYOL, SimSiam)
- InfoNCE
- Contrastive лосс
- CPC.

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

Учебная аудитория, оснащенная мультимедийным оборудованием (проектор или плазменная панель), доской.

### 6. Перечень рекомендуемой литературы

Основная литература

1. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль . — Москва, ДМК Пресс, 2018.— URL: <https://e.lanbook.com/book/107901> (дата обращения: 29.01.2021). - Полный текст (Режим доступа : из сети МФТИ / Удаленный доступ)

Дополнительная литература

1. Моттль В.В., Мучник И.Б.. Скрытые марковские модели в структурном анализе сигналов. М.: Наука, Физматлит, 1999, 351 с.
2. Виленкин С.Я. Статистическая обработка результатов исследования случайных функций. М.: Энергия, 1979.
3. Goodfellow I., Bengio Y., Courville A. Deep learning. – MIT press, 2016.
4. MacKay D. J. C. et al. Information theory, inference and learning algorithms. – Cambridge university press, 2003.

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

<http://www.machinelearning.ru/wiki/index.php>

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

Программное обеспечение и информационные технологии не требуются.

### 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 дифференцированному зачету.

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

Важно добиться понимания изучаемого материала, а не механического его запоминания.

В курсе выполняются 4 лабораторные работы.

Первая работа – состоит из реализации алгоритма обработки и классификации аудиоданных с помощью нейросетевых архитектур,

Вторая работа – состоит из реализации классического метода распознавания речи нейросетевым подходом, обучение алгоритма СТС и анализа архитектуры модели глубинного обучения и результатов ее работы.

Третья работа – состоит из реализации современного метода распознавания речи, анализа архитектуры модели глубинного обучения и результатов ее работы.

Четвертая работа – состоит из реализации современного метода генерации речи, анализа архитектуры модели глубинного обучения и результатов ее работы.

Каждое задание и зачет оцениваются по двухбалльной шкале.

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)**

<b>по направлению:</b>	Информатика и вычислительная техника
<b>профиль подготовки:</b>	Прикладная математика и информатика Физтех-школа Прикладной Математики и Информатики кафедра интеллектуальных систем
<b>курс:</b>	<u>2</u>
<b>квалификация:</b>	магистр

Семестр, формы промежуточной аттестации: 3 (осенний) - Дифференцированный зачет

**Разработчик:** П.А. Северилов

## 1. Компетенции, формируемые в процессе изучения дисциплины

Код и наименование компетенции	Индикаторы достижения компетенции
ОПК-1 Владеет системой фундаментальных научных знаний в области информатики и вычислительной техники	ОПК-1.1 Знает и способен использовать в профессиональной деятельности фундаментальные научные знания и новые научные принципы и методы исследований в области информатики и вычислительной техники
	ОПК-1.2 Способен обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности
	ОПК-1.3 Понимает междисциплинарные связи в области информатики и вычислительной техники и способен их применять при решении задач профессиональной деятельности
ПК-2 Понимает и способен применить в научно-исследовательской и прикладной деятельности основные законы естествознания, современный математический аппарат и алгоритмы, современные информационно-коммуникационные технологии	ПК-2.1 Знает основы научно-исследовательской деятельности в области информационных технологий, владеет знанием основ философии и методологии науки; знанием методов научных исследований и навыками их проведения
	ПК-2.2 Умеет применять полученные знания в области фундаментальных научных основ теории информации и решать стандартные задачи в собственной научно-исследовательской деятельности
	ПК-2.3 Имеет практический опыт научно-исследовательской деятельности в области информационно-коммуникационных технологий
	ПК-2.4 Владеет методами и алгоритмами решения задач цифровой обработки сигналов, использования сети Интернет, аннотирования, реферирования, библиографического поиска, опыт работы с научными источниками

## 2. Показатели оценивания компетенций

В результате изучения дисциплины «Обработка сигналов» обучающийся должен:

### знать:

- современные подходы распознавания, генерации, обработки аудиосигналов.

### уметь:

- анализировать аудиосигналы, внутреннее устройство и результаты работы модели глубокого обучения.

### владеть:

- навыками реализации базовых алгоритмов обработки, распознавания, генерации речи и смежными задачами обработки аудиоданных.

## 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 прошлого занятия.

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

1. Звук: типы, характеристики, теорема Найквиста. Спектрограммы: преобразование Фурье, Short-time преобразование Фурье, Mel спектрограммы

2. Автоматическое распознавание речи (ASR): постановка задачи, метрики. CTC: алгоритм, проблемы, лосс функция, свойства, beam search. LAS
3. RNN-T: архитектура модели, inference модели, процесс обучения, преимущества над CTC и LAS
4. Языковые модели для задачи ASR: мотивация, различные подходы. Byte pair encoding (BPE): мотивация, алгоритм.
5. Keyword spotting (KWS): метрики, DNN/CNN-based, модели KWS на основе механизма внимания, субспектральная нормализация, проблемы реализации KWS на практике и их решения.
6. Text-to-speech (TTS): пайплайн решения задачи, метрики. Использование идей механизма внимания в задаче TTS
7. Современные подходы в решении задачи TTS: DeepVoice, FastSpeech, FastSpeech 2, AdaSpeech
8. Wavenet: causal+dilated свертка, решение проблемы большого числа таймстемпов для предсказания.
9. Parallel WaveGAN: архитектура, лосс. DiffWave: диффузионные процессы, обучение модели, сэмплирование из модели, преимущества 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**

Во время приёма дифференцированного зачёта (он проводится в письменной форме) при подготовке билета можно пользоваться любыми материалами. При непосредственном ответе ничем пользоваться нельзя. Необходимо обратить внимание на теоретический минимум. Незнание ответов на вопросы из теоретического минимума автоматически влечёт неудовлетворительную оценку.