

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ливанов Дмитрий Викторович

Должность: Ректор

Дата подписания: 17.06.2025 16:37:23

Уникальный программный ключ:

c6d909c49c1d2034fa3a0156c4eaa51e7237a7a2

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

ВЫПИСКА ИЗ ПРОТОКОЛА № 6

заседания учебно-методического совета от 27 января 2022 г.

ПОВЕСТКА:

Рассмотрение дополнительных общеобразовательных и профессиональных программ.

Проректор по учебной работе А. А. Воронов

СЛУШАЛИ: Директора Физтех-школы прикладной математики и информатики А.М. Райгородского.

ПОСТАНОВИЛИ:

Рекомендовать к утверждению в установленном порядке дополнительную общеобразовательную программу «Transformer: работа с текстом».

Решение принято единогласно заочным голосованием.

Форма проведения заседания: заочная

Председатель УМС МФТИ

А.А. Воронов

Ученый секретарь УМС МФТИ

М.В. Березникова

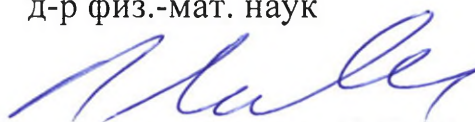
Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение высшего
образования

«Московский физико-технический институт
(национальный исследовательский университет)»

УТВЕРЖДАЮ

Исполняющий обязанности ректора

д-р физ.-мат. наук



Д. В. Ливанов

«27» января 2022 г.

Программа дополнительного образования

«Transformers: работа с текстом»

Москва 2022

1. Общая характеристика программы

1.1.Целью реализации дополнительной общеобразовательной программы «Transformers: работа с текстом» (далее – программа) является приобретение и совершенствование профессиональных, общепрофессиональных, универсальных и общекультурных компетенций слушателей для работы с классификацией текста с использованием Transformers

1.2.Категории слушателей, на обучение которых рассчитана программа: опытные ИТ-специалисты в области больших данных, которые хотят расширить свои знания и навыки в работе с Transformer с помощью практических заданий, основанных на реальных примерах

1.3.Нормативный срок освоения программы – 16 академических часов.

1.4.Форма обучения – очная (с применением дистанционных образовательных технологий).

1.5.Режим обучения: дистанционно – 2 дня (8 ак. часов в день).

2. Планируемые результаты обучения

В результате освоения программы слушатель должен:

Владеть навыками:

- использования модели на основе Transformer для задач бизнеса

Знать:

- основные тренды Natural Language Processing (NLP) 2018-2021 годов
- как устроена архитектура в ее реализации с нуля на pytorch

Уметь:

- строить нейронные сети на основе архитектуры Transformer

3. Структура программы

Программа предусматривает изучение следующих модулей:

Модуль 1. «Transformers: работа с текстом»

Таблица 1

№ п/п	Модуль	Тема	Колич ество часов	В том числе	
				ауд. заняти я	самос т. работа
1.	«Transformers: работа с текстом»	Трансформеры	8	4	4
2.		NLP архитектуры трансформеров.	8	4	4
ИТОГО			16	8	8

4. Содержание программы

4.1. Учебно-тематический план программы

Таблица 2

№ п/п	Наименование модуля, разделов и тем	Содержание обучения, наименование и тематика практических занятий (семинаров), самостоятельной работы	Всего, час
1	Transformers: работа с текстом		16
1.1	Трансформеры		8
1.1.1	Лекция	Трансформеры: Ресурсы: seq2seq Высокоуровневое понимание работы трансформера Self-attention Multi-head attention Positional encoding Skip connection & layer norm Decoder и masking Elmo OpenAI transformer	2

		Bert Токенизаторы	
1.1.2	Практическое занятие.		2
1.1.3	Самостоятельная работа		4
1.2	NLP архитектуры трансформеров.		8
1.2.1	Лекция	Анализ контекстуальности и на что смотрим attention Архитектуры	2
1.2.2	Практическое занятие	Устройство архитектуры в ее реализации с нуля на pytorch	2
1.2.3	Самостоятельная работа	Создаем модель классификации текста с использованием трансформеров	4
Итого:			16

4.2. Учебная программа по модулям

Таблица 3

№ п/п	Наименование модуля, разделов и тем	Содержание обучения и тематика практических занятий (семинаров), самостоятельной работы, используемых образовательных технологий
1.	Transformers: работа с текстом	Трансформеры. NLP архитектуры трансформеров. Анализ контекстуальности и на что смотрим attention. Архитектуры. Устройство архитектуры в ее реализации с нуля на pytorch. Создаем модель классификации текста с использованием трансформеров

4.3. Список рекомендуемой литературы

1. Гудфеллоу Я., Бенджио И., Курвилль А. – Глубокое обучение, 2017
2. Jurafsky, Martin. Speech and Language Processing, 2020
3. Rothman. Transformers for Natural Language Processing, 2021
4. Vajjala и др. Practical Natural Language Processing, 2020

1. Материально-технические условия реализации программы

Таблица 4

Наименование специализированных аудиторий, кабинетов, лабораторий	Вид занятий	Наименование оборудования, программного обеспечения
Информационно-коммуникационная платформа дистанционных лекций и семинаров	Лекции, практические занятия (дистанционные семинары)	Слушателю необходимо наличие доступа в сеть интернет, компьютер. Преподавателю курса необходимо оборудование для проведения дистанционных лекций, семинаров (вебинаров), качественный отказоустойчивый доступ в сеть интернет.
Информационно-коммуникационная платформа	Самостоятельная работа	Наличие компьютера и доступа в сеть интернет.
Информационно-коммуникационная платформа дистанционных лекций и семинаров	Промежуточный контроль	Наличие компьютера и доступа в сеть интернет.

2. Оценка качества освоения программы

Оценка качества освоения программы осуществляется в процессе промежуточной аттестации. Формы и методы промежуточного контроля представлены в таблице 5.

Таблица 5

Наименование модулей	Основные показатели оценки	Формы и методы контроля и оценки
Transformers: работа с текстом	зачет/незачет	Лабораторная работа

3. Составители программы

Райгородский Андрей Михайлович

Доктор физико-математических наук, директор ФПМИ МФТИ

Благодарный Евгений Владимирович

заведующий учебно-методической лабораторией инноватики ФПМИ

Тихонов Илья

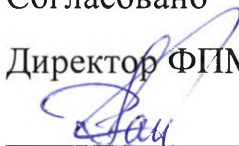
составитель и преподаватель программы

Иванова Анастасия Сергеевна

руководитель проектов учебно-методической лаборатории инноватики ФПМИ

Согласовано

Директор ФПМИ, д.ф.-м.н.

 А.М. Райгородский

« 27 » января 2022 г.

**КВАЛИФИКАЦИЯ И ОПЫТ ПРИВЛЕКАЕМОГО ПРЕПОДАВАТЕЛЬСКОГО СОСТАВА ДЛЯ РЕАЛИЗАЦИИ ПРОГРАММЫ
ДОПОЛНИТЕЛЬНОЙ ОБЩЕОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ «TRANSFORMER: РАБОТА С ТЕКСТОМ»**

Ф.И.О. лектора, год рождения	Информация об образовании, полученном в соответствии с образовательными программами высшего профессионального образования, дополнительного профессионального образования (в т.ч. о наличии званий и ученых степеней) и т.д.	Место работы, занимаемая должность в настоящий момент, общий трудовой стаж	Опыт преподавания и консультирования по предмету, согласующемуся с направлением лота (перечислить), педагогический стаж	Наличие опыта практической работы в отечественных и зарубежных организациях в сфере деятельности, совпадающей с направлением преподавания
Тихонов Илья	высшее образование (МФТИ) Аспирант кафедры ЭВМ МФТИ, научный руководитель д.ф.- м.н.Жданов А.А.	Преподаватель курсов МФТИ по data science, data analysis и data engineering для обучения разработчиков партнера МФТИ - Сбербанка.	Педагогический стаж: 3 года.	Участвовал в создании и проведении курса по основам машинного обучения для Вконтакте и Мфти. Разработчик аналитических продукты и систем с применением алгоритмов машинного обучения для задач тематического моделирования и кластеризации (АО "Интеллект") для компаний на заказ и в рамках работы в Лаборатории Инноватики МФТИ.

**Пояснительная записка
к разработке и реализации дополнительной общеобразовательной программы
“Transformers: работа с текстом”**

Данная программа дает слушателям понимание сути Transformer, из чего состоит модель и какие функции выполняет, где и как применяется. Программа позволит сформировать навыки работы с Transformer - самой продвинутой архитектурой в области обработки естественной речи (NLP) на сегодняшний день.

Трансформеры представляют собой тысячи предварительно обученных моделей для выполнения задач в различных модальностях, таких как текст, зрение и звук. Они используются для таких задач как:

- классификация текста, извлечение информации, ответы на вопросы, обобщение, перевод. создание текста на более чем 100 языках.
- классификация изображений, обнаружение объектов и сегментация
- распознавание речи и классификация аудио
- классификация видео и визуальные ответы на вопросы.

Задачи курса “Transformers: работа с текстом”

- 1- Познакомить слушателей с основными трендами Natural Language Processing (NLP) 2018–2021 годов, и научить слушателей работать архитектурой Transformer
- 2- Реализовывать архитектуру Transformer с нуля на практике с применением моделей для задач бизнеса (на примере дообучения Bert для классификации, многоклассовой классификации, распознавания именованных сущностей)

Программа рассчитана на 16 академических часа. Занятия состоят из лекций и семинаров, для поддержки обучающихся будут созданы чаты с преподавателями. В качестве контроля знаний слушатели выполняют лабораторную работу. После успешного прохождения программы слушатели получают сертификат о прохождении программы.

Составители программы:

Райгородский Андрей Михайлович

Доктор физико-математических наук, директор ФПМИ МФТИ

Благодарный Евгений Владимирович

заведующий учебно-методической лабораторией инноватики ФПМИ

Тихонов Илья

составитель и преподаватель программы

Иванова Анастасия Сергеевна

руководитель проектов учебно-методической лаборатории инноватики ФПМИ

Директор ФПМИ, д.ф.-м.н.

А.М. Райгородский