

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ливанов Дмитрий Викторович

Должность: Ректор

Дата подписания: 10.02.2026 12:27:28

Уникальный программный ключ:

c6d909c49c1d2034fa3a0156c4eaa51e72328982

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

ВЫПИСКА ИЗ ПРОТОКОЛА № 4

заседания учебно-методического совета от 24 декабря 2025 года.

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 программу профессиональной переподготовки «Инженер машинного обучения».

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”

Ректор

д-р физ.-мат. наук



Д. В. Ливанов

«24» декабря 2025 г.

**Дополнительная профессиональная программа
Программа профессиональной переподготовки**

«Инженер машинного обучения»

Москва 2025

Оглавление

1. Рабочая группа	2
2. Профессиональные стандарты, учтённые в содержании программы	2
3. Квалификационные требования, учтённые в содержании программы	2
4. Требования федерального государственного стандарта высшего образования, на основании которых была разработана программа	2
5. Характеристика новой квалификации и связанных с ней видов профессиональной деятельности, трудовых функций и (или) уровней квалификации.	2
6. Характеристика компетенций, подлежащих совершенствованию, и (или) перечень компетенций, формирующихся в результате освоения программы	3
7. Цель программы	4
8. Планируемые результаты обучения по программе	4
9. Учебный план	4
10. Календарный учебный график	5
11. Рабочие программы (учебная программа)	5
12. Организационно-педагогические условия реализации дополнительной профессиональной программы	6
12.1. Требования к квалификации педагогических кадров/представителей предприятий и организаций, обеспечивающих реализацию образовательного процесса	6
12.2. Требования к материально-техническим условиям	6
12.3. Требованиям к информационным и учебно-методическим условиям*	6
12.4. Общие требования к организации образовательного процесса*	6
13. Формы аттестации и оценочные материалы по программе	7

1. Рабочая группа

С целью повышения качества дополнительной профессиональной программы (ДПП) с учетом требований профессиональных стандартов (ПС) в группу разработчиков входили*:

Райгородский Андрей Михайлович

Доктор физико-математических наук, директор ФПМИ МФТИ

Ивченко Олег Николаевич

старший преподаватель кафедры АТП ФПМИ

Благодарный Евгений Владимирович

заведующий учебно-методической лабораторией инноватики ФПМИ

Иванова Анастасия Сергеевна

руководитель проектов учебно-методической лаборатории инноватики ФПМИ

2. Профессиональные стандарты, учтённые в содержании программы

Настоящая дополнительная профессиональная программа (ДПП) разработана с учётом соответствующих ей профессиональных стандартов (ПС) из национального реестра профессиональных стандартов. Профессиональные стандарты (ПС), связанные с настоящей ДПП:

ПС 06.042 Специалист по большим данным

3. Квалификационные требования, учтённые в содержании программы

Квалификационные требования, учтённые в содержании программы, указаны в описании должностей единого квалификационного справочника должностей руководителей, специалистов и служащих. Причём учтены требования должностей, указанных в дополнительных характеристиках отобранных для данной программы ОТФ (п. 5. настоящего документа), которые приведены в указанных в п. 2. настоящего документа профессиональных стандартах (ПС) в строке «ЕТКС или ЕКС». Квалификационные требования инкорпорированы в настоящий документ путем отсылки и являются его неотъемлемой частью.

4. Требования федерального государственного стандарта высшего образования, на основании которых была разработана программа

Настоящей дополнительной профессиональной программе соответствует федеральный государственный стандарт высшего образования (ФГОС ВО): *ФГОС 01.04.02 «Прикладная математика и информатика (уровень магистратуры)*

Настоящая дополнительная профессиональная программа разработана на основании требований указанного ФГОС ВО.

Полный перечень требований приведён в тексте указанного ФГОС ВО, инкорпорирован в настоящий документ путём отсылки и является его неотъемлемой частью.

5. Характеристика новой квалификации и связанных с ней видов профессиональной деятельности, трудовых функций и (или) уровней квалификации.

«Программа профессиональной переподготовки направлена на получение компетенции, необходимой для выполнения нового вида профессиональной деятельности, приобретение новой квалификации (часть 5 статьи 76 Федерального закона N 273-ФЗ). Для их определения и может

использоваться ПС. При этом необходимо учесть различия терминологии, используемой в образовании и в профессиональных стандартах, о которых говорилось выше. Вид профессиональной деятельности, квалификация, упоминаемые в цитируемой статье закона, в ПС в большинстве случаев соответствуют ОТФ, иногда - ТФ.»¹,

Новые квалификации (ОТФ/ТФ) и их уровни:

ПС 06.042 *Специалист по большим данным*

Обобщенные трудовые функции			Трудовые функции		
код	наименование	уровень квалификации	наименование	код	уровень (подуровень) квалификации
А	Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры	6	Выявление, формирование и согласование требований к результатам аналитических работ с применением технологий больших данных	А/01.6	6
			Подготовка данных для проведения аналитических работ по исследованию больших данных	А/03.6	6
			Проведение аналитического исследования с применением технологий больших данных в соответствии с требованиями заказчика	А/04.6	6
В	Управление этапами жизненного цикла методологической и технологической	7	Анализ потребности заинтересованных лиц и/или подразделений организации в	В/01.7	7

¹ Письмо Министерства образования и науки РФ от 22 апреля 2015 г. N ВК-1032/06 "О направлении методических рекомендаций"

	инфраструктуры анализа больших данных в организации		исследовании больших данных		
			Разработка и согласование технического проекта методической и технологической инфраструктуры больших данных	В/03.7	7
			Разработка, согласование и управление реализацией рабочего проекта методической и технологической инфраструктуры больших данных	В/04.7	7
			Управление получением, хранением, передачей, обработкой больших данных	В/05.7	7
			Управление качеством больших данных	В/07.7	7
			Управление защитой и обеспечением конфиденциальности больших данных	В/08.7	7
			Управление разработкой продуктов, услуг и решений на основе больших данных	8	Разработка продуктов на основе встроенной аналитики больших данных
Разработка сервисов на основе аналитики больших данных	С/02.8	8			
Разработка инфраструктурных	С/03.8	8			

			решений на основе аналитики больших данных		
D	Разработка и внедрение новых методов и технологий исследования больших данных	8	Совершенствование и разработка новых методов, моделей, алгоритмов, технологий и инструментальных средств работы с большими данными	D/01.8	8
			Проведение испытаний и разработка рекомендаций по внедрению и использованию усовершенствованных или разработанных новых методов, моделей, алгоритмов, технологий и инструментальных средств работы с большими данными	D/02.8	8

Характеристики данных ОТФ/ТФ приведены в профессиональном стандарте (ПС), указанном в п. 2. настоящего документа. инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

Связанные виды профессиональной деятельности (ВПД):

научно-исследовательская;
 проектная;
 производственно-технологическая;
 организационно-управленческая;
 педагогическая.

Характеристика данных ВПД приведена в федеральном государственном стандарте высшего образования (ФГОС ВО), указанном в п. 4. настоящего документа, инкорпорирована в настоящий документ путём отсылки, и является его неотъемлемой частью. Указанные ВПД, освоение которых предусмотрено указанным ФГОС ВО, являются «сквозными» и служат основой овладения выбранной квалификацией (ОТФ/ТФ).

ТФ, связанные с указанными в настоящем пункте ОТФ (в случае их указания), а также их характеристики приведены в профессиональном стандарте (ПС), указанном в п. 2. настоящего документа. инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

6. Характеристика компетенций, подлежащих совершенствованию, и (или) перечень компетенций, формирующихся в результате освоения программы

Учитывая, что слушатели программы уже освоили или осваивают основную профессиональную образовательную программу, в том числе общие компетенции, соответствующего уровня профессионального образования, компетенции формирующиеся в результате освоения настоящей программы и компетенции подлежащие совершенствованию, а также их характеристика, представлены в ФГОС, указанном в п. 4. настоящего документа, инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

Данные компетенции указанного ФГОС ВО служат основой определения результатов обучения по программе.

7. Цель программы

Реализация программы профессиональной переподготовки направлена на получение компетенции, необходимой для выполнения нового вида профессиональной деятельности, приобретение новой квалификации.

В соответствии с Письмом Министерства образования и науки РФ от 22 апреля 2015 г. N ВК-1032/06 «О направлении методических рекомендаций», цель представляет собой осознанное представление (предвосхищение) результата деятельности. Планируемые результаты отражены ниже.

8. Планируемые результаты обучения по программе

Планируемые результаты обучения соответствуют результатам освоения соответствующей основной профессиональной образовательной программы, а также направлены на приобретение новой квалификации, требующей изменение направленности (профиля) или специализации в рамках направления подготовки (специальности) полученного ранее профессионального образования, определены на основе профессиональных компетенций соответствующего федерального государственного образовательного стандарта, указанного в п. 4. настоящего документа.

Имеющаяся квалификация и (или) уровень образования (требования к слушателям):

Высшее и незаконченное высшее образование

Новые виды профессиональной деятельности, и ПК соответствующие им, которые планируются к получению в результате обучения по настоящей программе, а также их характеристики, представлены в ФГОС, указанном в п. 4. настоящего документа, инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

Выпускник также должен обладать прочими компетенциями в соответствии с указанным ФГОС, которые приведены и охарактеризованы в нём, инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

Новые квалификации (ОТФ/ТФ), планируемые к приобретению в результате обучения по настоящей программе, приведены и охарактеризованы в п. 5. настоящего документа.

Приведённая информация о цели и результатах обучения является основой для разработки рабочих программ, оценочных материалов и иных компонентов дополнительной профессиональной программы.

9. Учебный план

№ п/п *	Наименование Модулей, дисциплин	Всего, час.	В том числе:				
			Лекци и	Практически е занятия (семинары)	Самостоятел ьная работа	Контрольны е задания и подготовка к аттестации	Форма контроля
1	Анализ изображений	135	26	26	38	45	12 ДЗ
2	Обработка естественного языка/Анализ текста	100	24	24	10	42	тести рование, 2 ДЗ, проект
3	Машинное обучение на больших объёмах данных	90	32	32	10	16	4 ДЗ
4	Итоговая квалификационн ая работа	35					
	Итого	360					

* Слушатели могут проходить модули в произвольном порядке

Срок освоения настоящей программы профессиональной переподготовки превышает минимально допустимый срок освоения 250 часов (либо другой установленный актуальными нормативными документами соответствующего Министерства срок) или равен ему.

10. Календарный учебный график

Модуль 1: Анализ изображений

	Обучение по установленной форме: очная, с применением электронного обучения и дистанционных образовательных технологий
	Количество академических часов: 135

Порядковый № занятия (по горизонтали)	1	2	3	4	5	6	7	8	9	10	11	12	13
Лекции	2	2	2	2	2	2	2	2	2	2	2	2	2
Практические	2	2	2	2	2	2	2	2	2	2	2	2	2
Самостоятельные				2	2	2	2	2	2	2	2	2	2
Контрольные задания			2	4	4	4	4	4	4	3		4	
Итого	4	4	6	10	10	10	10	10	10	9	6	10	6
подготовка к аттестации	2	2	2	2	3	3	2	3	3	2	2	2	2

Модуль 2: Обработка естественного языка/Анализ текста

	Обучение по установленной форме: очная, с применением дистанционных технологий											
	Количество академических часов: 100											
Порядковый № занятия (по горизонтали)	1	2	3	4	5	6	7	8	9	10	11	12
лекции	2	2	2	2	2	2	2	2	2	2	2	2
практические	2	2	2	2	2	2	2	2	2	2	2	2

самостоятель ные	0,5	0,5	0,5	0,5	1	1	1	1	1	1	1	1	1
контрольные задания	1,25	1,25	2,25	3,25	3,25	3,25	3,25	4,25	3,25	6,25	5,25	5,25	
Итого	5,75	5,75	6,75	7,75	8,25	8,25	8,25	9,25	8,25	11,25	10,25	10,25	

Модуль 3: Машинное обучение на больших объёмах данных

	Обучение по установленной форме: очная, с применением электронного обучения и дистанционных образовательных технологий															
	Количество академических часов: 90															
Порядко вый № занятия (по горизонт али)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Лекции	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Практич еские	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Самосто ятельны е											1	1	2	2	2	2
Контрол ьные задания											4	3	4	3	2	

Итого	4	4	4	4	4	4	4	4	4	4	9	8	10	9	8	6
-------	---	---	---	---	---	---	---	---	---	---	---	---	----	---	---	---

11. Рабочие программы (учебная программа)

Модуль 1: Анализ изображений

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем, ак.час.
1.1	Введение в анализ изображений. Основы обработки изображений	14
1.1.1	Лекция. Цифровое изображение	2
	Семинар. Введение в практическую часть курса, Работа с numpy	2
	Самостоятельная работа.	
	Контрольные задания.	
1.1.2.	Лекция. Основы обработки изображений	2
	Семинар. Базовая обработка изображений	2
	Самостоятельная работа	
	Контрольные задания	
1.1.3.	Лекция. Сжатие изображений, Преобразование Фурье	2
	Семинар. Преобразование Фурье	2
	Самостоятельная работа	
	Контрольные задания	2
	Подготовка к итоговой аттестации	6
1.2	Основы машинного обучения. Обнаружение объектов. Введение в нейросети	30
1.2.1	Лекция. Классификация изображений. Введение в нейросети	2
	Семинар. Иерархия абстракций: numpy, pytorch, pytorch-lightning	2
	Самостоятельная работа	2

	Контрольные задания	4
1.2.2.	Лекция. Сверточные нейросетевые архитектуры	2
	Семинар. Нейросетевые задания курса, Базовое обучение нейросетей	2
	Самостоятельная работа	2
	Контрольные задания	4
1.2.3	Лекция. Трансформеры и сверточные нейронные сети с большими ядрами	2
	Семинар. Разбор устройства ViT, Приемы для дообучения нейросетей	2
	Самостоятельная работа	2
	Контрольные задания	4
	Подготовка к итоговой аттестации	8
1.3	Выделение базовых объектов на изображениях. Геометрические модели сопоставления изображений.	30
1.3.1.	Лекция. Поиск похожих изображений	2
	Семинар. Метрическое обучение, Эффективность тензорных вычислений	2
	Самостоятельная работа	2
	Контрольные задания	4
1.3.2.	Лекция. Детекторы объектов	2
	Семинар. Детектирование объектов	2
	Самостоятельная работа	2
	Контрольные задания	4
1.3.3	Лекция. Сегментация изображений	2
	Семинар. Простая нейросеть для сегментации	2
	Самостоятельная работа	2
	Контрольные задания	4
	Подготовка к итоговой аттестации	8
1.4.	Нейросетевые подходы к анализу изображений. Основные задачи и алгоритмы	31
1.4.1.	Лекция. Основы обработки видео	2

	Семинар. Работа с видео	2
	Самостоятельная работа	2
	Контрольные задания	3
1.4.2.	Лекция. Self-supervised learning, foundation models	2
	Семинар. Self-supervised learning	2
	Самостоятельная работа	2
	Контрольные задания	
1.4.3.	Лекция. Перенос стиля, superresolution, GAN, VAE (начало)	2
	Семинар. AE, VAE, GAN	2
	Самостоятельная работа	2
	Контрольные задания	4
1.4.4	Лекция. VAE (продолжение), Diffusion	2
	Семинар. Diffusion	2
	Самостоятельная работа	2
	Контрольные задания	
	Подготовка к итоговой аттестации	8
	Итого	135

Модуль 2: Обработка естественного языка /Анализ текста

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем, ак.час.
2.1	Введение в обработку естественного языка (NLP)	26
2.1.1	Лекция. Введение в обработку естественного языка	2
	Семинар. Применение Python, NumPy и NLTK в задачах обработки естественного языка	2
	Самостоятельная работа.	0,5
	Контрольные задания.	1,25

2.1.2.	Лекция. Базовые принципы машинного обучения и их применение для классификации текстов	2
	Семинар. Практическая реализация логистической регрессии на PyTorch	2
	Самостоятельная работа	0,5
	Контрольные задания	1,25
2.1.3.	Лекция. Эмбединги слов: как нейросети понимают смысл	2
	Семинар. Строим свои эмбединги: Word2Vec для слов, Doc2Vec для текстов	2
	Самостоятельная работа	0,5
	Контрольные задания	2,25
2.1.4	Лекция. Сверточные нейронные сети (CNN): архитектура и принципы	2
	Семинар. Классификация текстов с использованием сверточных сетей	2
	Самостоятельная работа	0,5
	Контрольные задания	3,25
2.2.	Фундаментальные и современные подходы в NLP	34
2.2.1.	Лекция. Рекуррентные нейронные сети (RNN) для обработки последовательностей	2
	Семинар. Тематическое моделирование: поиск и визуализация скрытых тем в текстах	2
	Самостоятельная работа	1
	Контрольные задания	3,25
2.2.2	Лекция. Статистический машинный перевод (SMT): классический подход	2
	Семинар. Настройка процесса обучения: dropout, регуляризация и инициализация	2
	Самостоятельная работа	1
	Контрольные задания	3,25
2.2.3	Лекция. Архитектура Transformer: внимание как новая парадигма	2
	Семинар. Работа с предобученными Transformer-моделями	2

	Самостоятельная работа	1
	Контрольные задания	3,25
2.2.4	Лекция. Большие языковые модели (LLM): архитектура, обучение, возможности	2
	Семинар. Прикладное использование LLM в NLP	2
	Самостоятельная работа	1
	Контрольные задания	4,25
2.3.	Прикладные техники обработки естественного языка	40
2.3.1	Лекция. Консультация по проектам	2
	Семинар. Named Entity Recognition (NER) с помощью BiLSTM + CRF	2
	Самостоятельная работа	1
	Контрольные задания	3,25
2.3.2	Лекция. Классические вероятностные модели: скрытые марковские модели (HMM) для разметки текста	2
	Семинар. Современное тематическое моделирование: методы на стыке статистики и нейросетей	2
	Самостоятельная работа	1
	Контрольные задания	6,25
2.3.3	Лекция. Выявление скрытых тем в текстовых данных (Topic Modeling)	2
	Семинар. RAG (Retrieval-Augmented Generation): как использовать это понимание для точных ответов	2
	Самостоятельная работа	1
	Контрольные задания	5,25
2.3.4	Лекция. Диалоговый искусственный интеллект (Conversational AI): архитектура и применение	2
	Семинар. Языковые модели для программирования (Code LLMs): генерация, анализ и сопровождение кода	2
	Самостоятельная работа	1
	Контрольные задания	5,25

	Итого	100
--	--------------	------------

Модуль 3: Машинное обучение на больших объёмах данных

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем, ак.час.
3.1	Рекомендательные системы	40
3.1.1	Лекция. Вводное занятие. Модуль рекомендательные системы	2
	Семинар. Изучаем практический фреймворк курса	2
	Самостоятельная работа.	
	Контрольные задания.	
3.1.2.	Лекция. Рекомендательные сервисы в продакшене	2
	Семинар. А/В тест эвристического рекомендера	2
	Самостоятельная работа	
	Контрольные задания	
3.1.3	Лекция. Метрики и базовые подходы	2
	Семинар. Сравниваем бейзлайн-рекомендеры офлайн и онлайн	2
	Самостоятельная работа	
	Контрольные задания	
3.1.4	Лекция. Классические алгоритмы I	2
	Семинар. Используем user2user подход	2
	Самостоятельная работа	
	Контрольные задания	
3.1.5.	Лекция. Классические алгоритмы II	2
	Семинар. Используем lightfm модель для коллаборативных рекомендаций	2
	Самостоятельная работа	

	Контрольные задания	
3.1.6	Лекция. Нейросетевые рекоммендеры I	2
	Семинар. Пробуем нейросетевой отборщик кандидатов (lightfm torch)	2
	Самостоятельная работа	
	Контрольные задания	
3.1.7	Лекция. Нейросетевые рекоммендеры II	2
	Семинар. Пробуем нейросетевой отборщик ранкер	2
	Самостоятельная работа	
	Контрольные задания	
3.1.8	Лекция. Графовые нейросети для рекомендаций	2
	Семинар. Пробуем графовую нейросеть	2
	Самостоятельная работа	
	Контрольные задания	
3.1.9	Лекция. Нерешенные проблемы и новые направления	2
	Семинар. Пробуем разнообразить рекомендации	2
	Самостоятельная работа	
	Контрольные задания	
3.1.10	Лекция. Рекомендации и Reinforcement Learning	2
	Семинар. AMA сессия	2
	Самостоятельная работа	
	Контрольные задания	
3.2.	Графовые сети	36
3.2.1	Лекция. Графовые сети для задач классификации, препроцессинг данных, семплинг, распределенное обучение	2
	Семинар. MT GNN -- GraphConv, NNConv,	2
	Самостоятельная работа	1
	Контрольные задания	4

3.2.2	Лекция. Поиск аномалий на графах (Graph Outlier Detection)	2
	Семинар. PyGOD over BOND -- SOTA 2023	2
	Самостоятельная работа	1
	Контрольные задания	3
3.2.3	Лекция. Динамические GNN с вниманием для предсказания траффика	2
	Семинар. TGCN, A3TGCN	2
	Самостоятельная работа	2
	Контрольные задания	4
3.2.4	Лекция. Графовые сети -- ограничения в применении и устойчивость к атакам	2
	Семинар. Атаки на графовую сеть классикации	2
	Самостоятельная работа	2
	Контрольные задания	3
3.3	Большие языковые модели	14
3.3.1	Лекция. Обучение LLM.	2
	Семинар. Обучение LLM.	2
	Самостоятельная работа	2
	Контрольные задания	2
3.3.2	Лекция. Длинный контекст, позиционное кодирование. Архитектурные модификации Transformer	2
	Семинар. Transformer	2
	Самостоятельная работа	2
	Контрольные задания	
	Итого	90

Содержание дополнительной профессиональной программы направлено на достижение результатов её целей (планируемых результатов).

12. Организационно-педагогические условия реализации дополнительной профессиональной программы

12.1. Требования к квалификации педагогических кадров/представителей предприятий и организаций, обеспечивающих реализацию образовательного процесса

Требования к образованию педагогических и иных работников, а также (при наличии) требования к освоению ими дополнительных профессиональных программ, опыту работы в области профессиональной деятельности, соответствующей направленности ДПП: *высшее образование, трудовой и (или) педагогический стаж от полугода, наличие опыта практической работы в сфере деятельности, совпадающей с направлением преподавания.*

12.2. Требования к материально-техническим условиям

Перечень кабинетов, лабораторий, мастерских, тренажеров и др., обеспечивающих проведение всех предусмотренных программой видов занятий:

Наименование специализированных аудиторий, кабинетов, лабораторий	Вид занятий	Наименование оборудования, программного обеспечения
Система дистанционного обучения провайдера массовых открытых онлайн курсов	Лекции	Слушателю необходимо наличие доступа в сеть интернет, компьютер. Минимально рекомендуемые требования к ПК: Процессор Intel Core i3-12100 OEM Материнская плата MSI PRO H610M-E DDR4 Видеокарта MSI GeForce 210 Оперативная память ADATA XPG SPECTRIX D35G RGB 16 ГБ Накопитель 480 ГБ 2.5" SATA накопитель Aсacer AS340 PANTHER Преподавателю необходимо оборудование для проведения дистанционных семинаров (вебинаров), качественный отказоустойчивый доступ в сеть интернет.

Информационно-коммуникационная платформа дистанционных семинаров	Практические занятия (дистанционные семинары)	Слушателю необходимо наличие доступа в сеть интернет, компьютер. Преподавателю необходимо оборудование для проведения дистанционных семинаров (вебинаров), качественный отказоустойчивый доступ в сеть интернет.
Система дистанционного обучения провайдера массовых открытых онлайн курсов	Самостоятельная работа	Наличие компьютера и доступа в сеть интернет. Минимально рекомендуемые требования к ПК: Процессор Intel Core i3-12100 OEM Материнская плата MSI PRO H610M-E DDR4 Видеокарта MSI GeForce 210 Оперативная память ADATA XPG SPECTRIX D35G RGB 16 ГБ Накопитель 480 ГБ 2.5" SATA накопитель Aсacer AS340 PANTHER
Система дистанционного обучения провайдера массовых открытых онлайн курсов	Промежуточный, рубежный контроль, Итоговая аттестация	Наличие компьютера и доступа в сеть интернет. Минимально рекомендуемые требования к ПК: Процессор Intel Core i3-12100 OEM Материнская плата MSI PRO H610M-E DDR4 Видеокарта MSI GeForce 210 Оперативная память ADATA XPG SPECTRIX D35G RGB 16 ГБ Накопитель 480 ГБ 2.5" SATA накопитель Aсacer AS340 PANTHER

12.3. Требованиям к информационным и учебно-методическим условиям

Модуль 1: Анализ изображений

Основная литература:

Компьютерное зрение. Современный подход, Дэвид Форсайт, Жан Понс, Вильямс, 2004
Computer vision. algorithms and applications, Richard Szeliski, Springer, 2011

Computer vision: models, learning, and inference, Simon Prince, Cambridge University Press, 2012
Глубокое обучение, Ян Гудфеллоу, Йошуа Бенджио, Аарон Курвилль, ДМК Пресс, 2017

Дополнительная литература:

Multiple view geometry in computer vision, Richard Hartley, Andrew Zisserman, Cambridge University Press, 2004

Компьютерное зрение [Текст] : учеб. пособие для вузов / Л. Шапиро, Дж. Стокман ; пер. с англ. А. А. Богуславского ; под ред. С. М. Соколова .— М. : БИНОМ. Лаб. знаний, 2006 .— 752 с.

Методические указания для обучающихся по освоению дисциплины (модуля)

1. По каждой теме курса рекомендуется обзорно изучить материал из списка основной литературы
2. Для практического закрепления навыков решения задач анализа изображений следует разбирать примеры скриптов на языке Python с занятий и решать данные на занятиях задачи
3. При желании приобрести дополнительные теоретические знания по темам, смежным с темой курса, можно обратиться к списку дополнительной литературы.
4. При готовности приобрести дополнительные практические знания по теме курса можно обратиться к любым доступным в интернете материалам прочих курсов по анализу изображений, либо получить индивидуальное задание у лектора.

Перечень ресурсов информационно-телекоммуникационной сети “Интернет”, необходимых для освоения дисциплины (модуля):

github.com

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

Установленное ПО на каждом компьютере: интерпретатор python 3.8.6 и модулями pipru, opencv-python 4.4, pytorch, matplotlib, scikit-learn, pytest

Модуль 2: Обработка естественного языка /Анализ текста

Основная

1. Jurafsky, D., & Martin, J. H. *Speech and Language Processing (3rd ed.)*. Pearson, 2023.
2. Manning, C. D., Raghavan, P., & Schütze, H. *Introduction to Information Retrieval*. Cambridge University Press, 2008.
3. Eisenstein, J. *Introduction to Natural Language Processing*. MIT Press, 2019.
4. Goldberg, Y. *Neural Network Methods for Natural Language Processing*. Morgan & Claypool, 2017
5. Бирюков С. И. *Обработка естественного языка в Python*. ДМК Пресс, 2020.
6. Лопатина М. Г., Соколова Е. Г. *Машинное обучение для обработки текстов*. Питер, 2021.
7. Мельников А. В. *Глубокое обучение для NLP*. СПб.: БХВ-Петербург, 2022.

Статьи и научные работы

8. Vaswani, A., et al. *Attention Is All You Need (Transformer)*. NeurIPS 2017.

9. Devlin, J., et al. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 2019.
10. Mikolov, T., et al. *Efficient Estimation of Word Representations in Vector Space (Word2Vec)*. 2013.
11. *Обзор современных методов NLP // Открытые системы. СУБД, 2021.*
12. Горбань А. Н. *Нейросетевые модели в лингвистике // Труды ИСИ СО РАН, 2020.*

Перечень ресурсов информационно-телекоммуникационной сети “Интернет”, необходимых для освоения дисциплины (модуля):

<https://ods.ai/>

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

Модуль 3: Машинное обучение на больших объёмах данных

Основная литература

1. Методы и средства вычислений с объектами. Аппликативные вычислительные системы [Текст] : [учеб. пособие для вузов] / В. Э. Вольфенгаген ; Ин-т актуального образования <ЮрИнфоР-МГУ>, Каф. перспективных компьютерных исследований и информационных технологий .— М. : JurInfoR, 2004 .— 789 с. — (Компьютерные науки и информационные технологии). - 2000 экз. - ISBN 5-89158-100-0 (в пер.) .
2. Комбинаторная логика в программировании. Вычисления с объектами в примерах и задачах [Текст] : [учеб. пособие для вузов] / В. Э. Вольфенгаген ; НОУ Ин-т Актуального образования "ЮрИнфоР-МГУ, Каф. перспективных компьт. исслед. и информ. технологий .— 3-е изд., доп. и перераб. — М. : Ин-т "ЮрИнфоР-МГУ, 2008 .— 384 с.
3. Нейронные сети [Текст] : полный курс / С. Хайкин ; пер. с англ. Н. Н. Куссуль, А. Ю. Шелестова ; под ред. Н. Н. Куссуль .— 2-е изд., испр. — М. : Вильямс, 2006 .— 1103 с.

Дополнительная литература

1. Прикладная статистика. Принципы и примеры [Текст] : [учеб. пособие для вузов] / Д. Кокс, Э. Снелл ; пер. с англ. Е. В. Чепурина ; под ред. Ю. К. Беляева .— М. : Мир, 1984 .— 200 с.
2. Параллельное программирование многопоточных систем с разделяемой памятью [Текст] : учеб. пособие для вузов / А. Г. Торماسов .— М : Физматкнига, 2014 .— 208 с.

Перечень ресурсов информационно-телекоммуникационной сети “Интернет”, необходимых для освоения дисциплины (модуля):

<https://www.coursera.org/specializations/big-data-engineering> - специализация из 5 курсов, посвящённая тематике обработки больших данных

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

12.4. Общие требования к организации образовательного процесса*

Форма обучения – очная, с применением дистанционных технологий

Режим обучения – 1 раз(а) в неделю по 2 академических часа.

Преподаватель проводит практические занятия: дистанционно в форме вебинаров с использованием платформы ZOOM (или аналогичной).

Самостоятельная работа выполняется слушателем: в удобном для слушателя режиме

В Таблице ниже описаны образовательные технологии.

№ п/п	Вид занятия	Форма проведения занятий	Цель
1	Лекция	Изложение материала посредством лекций, обсуждение общих вопросов по тематике курса.	Ознакомление слушателей с базовым материалом по тематике курса.
2	Практическое занятие	Выполнение практических заданий, получение обратной связи от преподавателя.	Практическое освоение теоретических знаний, разбор ситуаций, а также углубление знаний по курсу
3	Самостоятельная работа	Самостоятельное изучение дополнительных материалов и литературы. Выполнение тренировочных тестов и заданий.	Углубление знаний по курсу. Применение знаний к своей организации / проекту.
4	Выполнение контрольных заданий	Выполнение тестов, проверочных заданий.	Практическое освоение теоретических знаний, контроль освоения материалов.
5	Итоговая аттестация	Подготовка итоговой квалификационной работы.	Практическое освоение теоретических знаний, контроль освоения материалов. Получение практического результата в своих организациях / проектах.

13. Формы аттестации и оценочные материалы по программе

Оценка качества освоения программы проводится по пятибалльной системе: «отлично», «хорошо», «удовлетворительно» по результатам промежуточного контроля (тестирование,

проверочные задания), контроля посещаемости практических занятий (вебинаров) и результатам итоговой аттестации.

После аттестации слушатель получает оценку “отлично”, если набрано не меньше 80% баллов от возможного максимума, “хорошо”, если набрано не меньше 65%, “удовлетворительно”, если набрано не меньше 40%, в противном случае слушатель курс не сдает.

Результат тестирования, решения проверочных заданий и написания кода проверяется автоматически системой на образовательной платформе. Итоговая аттестационная работа проверяется преподавателем

Составляющие процесса обучения, которые оцениваются в ходе обучения, и их вклад в итоговую оценку представлены в Таблице ниже.

Таблица – Составляющие процесса обучения

№ п/п	Основные показатели оценки	Вклад в итоговую оценку
1	Основной курс обучения на образовательной платформе	30%
2	Практические занятия	50%
3	Итоговая аттестационная работа	20%

Оценочные материалы:

Пример тестового задания к модулю 1: Анализ изображений

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

Письменная работа:
Распознавание лиц

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

1. Операция свёртки и её свойства. Фильтры, представимые в виде свертки
2. Базовая обработка изображений: тональная коррекция, шумоподавление
3. Нелинейные фильтры. Математическая морфология
4. Преобразование Хафа и БПХ.
5. Поиск краев на изображении. Детектор Canny
6. Сегментация и бинаризация изображения. Текстура
7. Детектирование углов. Детектор Харриса. Детектор Харриса-Лапласа
8. Детектирование блоков. Лапласиан гауссианы и разность гауссиан
9. Дескриптор SIFT. Гистограмма ориентированных градиентов
10. Геометрические модели. Метод наименьших квадратов, его связь с сингулярным разложением. Робастные методы оценки параметров модели
11. Основы машинного обучения. Градиентный спуск
12. Линейные классификаторы. Метод опорных векторов
13. Деревья принятия решений, градиентный бустинг, AdaBoost
14. Метод Виолы-Джонса и дальнейшее его развитие
15. Устройство нейронных сетей. Виды слоев. Регуляризация
16. Нейронные сети. Алгоритм обратного распространения ошибки

17. Сверточные нейронные сети. Блоки нейронных сетей. Локальная нормализация, батч-нормализация
18. Основные архитектуры сверточных нейронных сетей. Inception, VGG, ResNet
19. Нейронные сети для распознавания лиц

Пример тестового задания к модулю 2: Анализ текста

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

1. Методы распознавания именованных сущностей в текстах
2. Автоматическое выявление терминов и терминологических связей в тексте
3. Исследование современных языковых моделей для задач семантической классификации
4. Исследование методов создания вопросно-ответных систем: IR based QA vs KB based QA

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

1. Приведите примеры использования технологий обработки текстов, с которыми вы сталкиваетесь в повседневной жизни
2. Напишите регулярное выражение для извлечения дат
3. Сформулируйте законы Хипса и Ципфа
4. Дайте определение скрытой Марковской цепи и покажите, как она может быть использована для разрешения морфологической омонимии
5. Перечислите несколько коэффициентов считаемости биграм
6. Почему возникает необходимость в снижении размерности в векторной модели
7. Как устроен алгоритм обучения word2vec?
8. Что такое латентное размещение Дирихле?
9. Методы классификации текстов
10. Как формулируется задача классификации последовательности?
11. Как сгенерировать текст с помощью счетной языковой модели?
12. Как сгенерировать текст с помощью нейронной вероятностной языковой модели?
13. Как формулируется задача заполнения слотов для чат-ботов?
14. Связь модели зашумленного канала и языковой модели
15. Назовите несколько мер релевантности в задаче поиска

Пример тестового задания к модулю 3: Машинное обучение на больших объемах данных

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

1. Какие способы построения рекомендательных систем Вы знаете? Кратко опишите суть каждого из них.
2. Обработка естественного языка: примеры задач, этапы решения задач естественного языка.

3. Обработка естественного языка: применение регулярных, контекстно-свободных и контекстно-зависимых грамматик.

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

1) Методы оптимизации и линейные модели

1. Машинное обучение с учителем на больших данных. Закон Ципфа. Тематическое моделирование.
2. Метод стохастического градиента. Постановка задачи. Оптимизации обучения на больших данных: градиентный спуск, стохастический градиент.
3. Признаки. Пространства признаков, веса признаков, нормализация признаков. Генерация и хеширование признаков.
4. Онлайн обучение линейных моделей. Метод стохастического градиента: выбор функции потерь. Оценка качества метода стохастического градиента. Алгоритм Бутстрап.
5. Хеширование, чувствительное к расстоянию (LSH). Меры сходства: расстояние Жаккара, Хемминга, косинусное расстояние, Евклидово расстояние.
6. Оптимизация и тестирование гиперпараметров. Симплекс-метод.

2) Алгоритмы работы с графами большого размера

1. Графы, их виды. Стохастический граф.
2. Представление графа: матрицы смежности, инцидентности, достижимости. Списки смежности. Алгоритмы перевода из одного представления в другое.
3. Социальный граф. Задача поиска общих друзей в социальном графе. Язык DSL.
4. Граф пользовательских предпочтений.
5. Использование подхода BigData в анализе графов.

3) Информационный поиск

1. Постановка ранжирования. Основные подходы к решению задачи ранжирования.
2. Метрики измерения точности ранжирования. Кликовые модели.
3. Тематическое моделирование и его связь с ранжированием.
4. Проблемы тематического моделирования при больших данных. AD-LDA, его недостатки, Y!LDA, Mr. LDA. ARTM. Архитектура библиотеки BigARTM. Online LDA и его применение в Vowpal Wabbit.

4) Рекомендательные системы

1. Постановка задачи рекомендаций. Неперсонализованные рекомендательные системы. Content-Based системы. Использование подхода BigData в рекомендательных системах.
2. Коллаборативная фильтрация и матричная факторизация.
3. Проблема “холодного старта”. Методы понижения размерности.

Директор ФПМИ, д.ф.-м.н.

A handwritten signature in blue ink, appearing to be 'А.М. Райгородский', written over a horizontal line.

А.М. Райгородский

КВАЛИФИКАЦИЯ И ОПЫТ ПРИВЛЕКАЕМОГО ПРЕПОДАВАТЕЛЬНОГО СОСТАВА ДЛЯ РЕАЛИЗАЦИИ ДОПОЛНИТЕЛЬНОЙ ПРОФЕССИОНАЛЬНОЙ ПРОГРАММЫ ПРОФЕССИОНАЛЬНОЙ ПЕРЕПОДГОТОВКИ «Инженер машинного обучения»

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, педагогический стаж	Опыт преподавания и консультирования по предмету, согласующемуся с направлением лота (перечислить), преподавательский стаж	Наличие опыта практической работы в отечественных и зарубежных организациях в сфере деятельности, совпадающей с направлением преподавания
Ивченко Олег Николаевич	<p>Высшее образование - специалист, магистратура, специальность: Прикладная математика и информатика, квалификация: Магистр, ФГБОУ ВПО "Московский физико-технический институт (национальный исследовательский университет)"</p> <p>Дополнительное профессиональное образование, курс: Математическое моделирование для анализа данных, ФГБОУ ВПО "Московский физико-технический институт (национальный исследовательский университет)"</p>	<p>Должность: Старший преподаватель Общий стаж: 4 года Педагогический стаж: 4 года.</p>	<p>Преподавательский стаж: 4 года</p>	<p>Разработчик системы NJudge - системы автоматизированного тестирования Nadoor-приложений, системный администратор инфраструктуры больших данных в группе Яндекс-CERN</p>

Пояснительная записка
к разработке и реализации дополнительной программы
повышения квалификации
«Инженер машинного обучения»

Согласно данным ведущих аналитических агентств профессии Data Scientist и инженер машинного обучения стабильно входят в топ-5 самых востребованных и высокооплачиваемых IT-специальностей. При этом рынок испытывает острый дефицит квалифицированных кадров, обладающих не только теоретическими знаниями, но и практическим опытом построения end-to-end решений. Особенно высок спрос на специалистов, способных работать с неструктурированными данными — изображениями и текстами, которые составляют свыше 80% всей цифровой информации.

Данная программа призвана заполнить нишу практико-ориентированной подготовки, где фокус смещен на решение реальных задач: от формулировки гипотезы и подготовки данных до развертывания и анализа работающей модели. Развитие компетенций в области ИИ и обработки данных напрямую соответствует целям национального проекта «Цифровая экономика» и «Стратегии развития искусственного интеллекта», направленным на обеспечение технологического суверенитета и подготовку кадров для высокотехнологичных отраслей.

В программу входит 3 модуля:

1. Анализ изображений
2. Обработка естественного языка / Анализ текстов
- 3.. Машинное обучение на больших объемах данных

Структура программы и наполненность соответствует текущим запросам рынка труда и технологическим трендам, что обеспечит высокую востребованность выпускников.

Программа будет построена по очному принципу с применением дистанционных технологий (с применением электронного обучения и дистанционных образовательных технологий) в формате вебинаров, которые займут 164 ак.ч.

Преподавателями программы будут являться преподаватели МФТИ, имеющие опыт преподавания и консультирования по заявленным модулям, а также опыт практической работы в отечественных и/или зарубежных организациях в сфере деятельности, совпадающей с направлением преподавания.

Взаимодействие со слушателями курса будет осуществляться дистанционно с помощью системы управления образовательным процессом (далее СДО – система дистанционного образования). Посредством СДО слушатели будут получать доступ к интерактивным лекциям в формате видеолекций. Контроль изучения этих материалов будет организован, как автоматически с помощью средств СДО, так и через тестовые задания. Дистанционный цикл будет проводиться одновременно для всех участников. Задания с открытым вариантом ответа будут проверяться преподавателями курса.

Составители программы:

Райгородский Андрей Михайлович

Доктор физико-математических наук, директор ФПМИ МФТИ

Благодарный Евгений Владимирович

заведующий учебно-методической лабораторией инноватики

Иванова Анастасия Сергеевна

руководитель проектов учебно-методической лаборатории инноватики ФПМИ

Директор ФПМИ, д.ф.-м.н.



А.М. Райгородский