

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное
учреждение высшего образования
«Казанский национальный исследовательский технический
университет им. А.Н. Туполева – КАИ»
(КНИТУ – КАИ)

СОГЛАСОВАНО:
Директор ПИШ КАИ

УТВЕРЖДАЮ:
Директор корпоративного института

Шабалин Л.П.

Гимбицкий А.В.

31 мая 2024 г.

31 мая 2024 г.

ДОПОЛНИТЕЛЬНАЯ ПРОФЕССИОНАЛЬНАЯ ПРОГРАММА
ПОВЫШЕНИЯ КВАЛИФИКАЦИИ

«Технологии интеллектуальной обработки данных»

1. Основные характеристики программы

Соответствие профессиональным стандартам	Программа составлена с учетом профессионального стандарта «Специалист по большим данным», утвержденного приказом Министерства труда и социальной защиты Российской Федерации от 06.06.2020 № 405н
Соответствие квалификационным требованиям	Программа составлена с учетом постановления Минтруда России от 21.08.1998 № 37 «Общепрофессиональные квалификационные характеристики должностей работников, занятых на предприятиях, в учреждениях и организациях» и приказа Минздравсоцразвития РФ от 11.01.2011 № 1н «Об утверждении Единого квалификационного справочника должностей руководителей, специалистов и служащих» раздел «Квалификационные характеристики должностей руководителей и специалистов высшего профессионального и дополнительного профессионального образования»
Категория слушателей	Лица, имеющие или получающие высшее или среднее профессиональное образование
Срок обучения	72 часа
Форма обучения	Очная с применением дистанционных образовательных технологий

Программа дополнительного профессионального образования разработана и реализуется в рамках федерального проекта «Передовые инженерные школы. Комплексная авиационная инженерия».

2. Цель реализации программы: приобретение компетенций решения задач в профессиональной деятельности научных сотрудников, преподавателей образовательных организаций, инженеров и технологов с использованием технологий интеллектуальной обработки данных.

3. Требования к результатам обучения

Слушатель, освоивший программу, должен:

3.1. Обладать следующими компетенциями:

- способность использовать знание основных методов интеллектуальной обработки данных для решения профессиональных задач в деятельности научных сотрудников, преподавателей образовательных организаций высшего образования, инженеров, технологов;
- способность разрабатывать оригинальные алгоритмы и программные средства интеллектуальной обработки данных для решения профессиональных задач в деятельности научных сотрудников, преподавателей образовательных организаций высшего образования, инженеров, технологов.

3.2. Знать:

- основы интеллектуальной обработки данных;

- модели машинного обучения, глубокого обучения и обучения с подкреплением;
- методики формирования обучающих, тестовых и контрольных наборов данных;
- методики оценки моделей обучения;
- технологии решения задач интеллектуальной обработки данных на языке программирования Python.

3.3. Уметь:

- формировать обучающие, тестовые и контрольные наборы данных для моделей обучения при решении задач интеллектуальной обработки данных;
- обучать и применять модели машинного обучения, глубокого обучения и обучения с подкреплением для решения задач интеллектуальной обработки данных;
- выполнять оценку эффективности обученных моделей;
- применять технологии решения задач интеллектуальной обработки данных на языке программирования Python.

3.4. Владеть:

- навыками формирования обучающих, тестовых и контрольных наборов данных для моделей обучения при решении задач интеллектуальной обработки данных;
- навыками обучения и применения моделей машинного обучения, глубокого обучения и обучения с подкреплением для решения задач интеллектуальной обработки данных;
- навыками оценки эффективности обученных моделей для решения задач интеллектуальной обработки данных;
- навыками применения технологий решения задач интеллектуальной обработки данных на языке программирования Python.

4. Содержание программы

Календарный учебный график

Образовательный процесс по программе может осуществляться в течение всего календарного года.

Форма обучения	Ауд. часов в день	Общая продолжительность программы (дней, недель, месяцев)
Очная с применением дистанционных образовательных технологий	4	2 месяца

УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН

программы повышения квалификации
«Технологии интеллектуальной обработки данных»

№	Наименование разделов, дисциплин и тем	Всего часов	В том числе			Формы контроля
			Лекции	Практиче- ские и лабо- раторные занятия	Самостоя- тельная работа	
1	2	3	4	5	6	7
1	Модуль 1 «Основы интеллектуальной обработки данных»	22	4	8	10	
	Тема 1. Основные понятия, методология и средства интеллектуальной обработки данных	10	2	4	4	
	Тема 2. Модели и методы машинного обучения	12	2	4	6	
2	Модуль 2 «Глубокое обучение»	24	4	8	12	
	Тема 1. Глубокие нейронные сети	12	2	4	6	
	Тема 2. Модели глубокого обучения для решения прикладных задач	12	2	4	6	
3	Модуль 3 «Обучение с подкреплением»	24	4	8	12	
	Тема 1. Введение в обучение с подкреплением.	12	2	4	6	
	Тема 2. Глубокое обучение с подкреплением.	12	2	4	6	
	Итоговая аттестация	2		2		Зачет
	Итого	72	12	26	34	

РАБОЧАЯ ПРОГРАММА
программы повышения квалификации
«Технологии интеллектуальной обработки данных»

Раздел 1. Основы интеллектуальной обработки данных (22 ч.)

1. Основные понятия, методология и средства интеллектуальной обработки данных.
 - Данные.
 - Задачи интеллектуальной обработки данных.
 - Примеры систем интеллектуальной обработки данных.
 - Математические основы интеллектуальной обработки данных.
 - Программные средства интеллектуальной обработки данных.
 - Основные понятия и методология интеллектуальной обработки данных.
 - Средства решения задач интеллектуальной обработки данных на языке Python. Основы языка Python.
 - Основные библиотеки Python для анализа данных.
 - Основные этапы интеллектуальной обработки данных.
 - Методики оценки эффективности систем интеллектуальной обработки данных.
 - Методики формирования обучающих, тестовых и контрольных наборов данных.
 - Аугментация данных. Предобработка данных.
2. Модели и методы машинного обучения.
 - Машинное обучение.
 - Модели машинного обучения с учителем.
 - Модели машинного обучения без учителя.
 - Модели машинного обучения с частичным привлечением учителя.
 - Задачи классификации, кластеризации и регрессии.
 - Методы решения задач классификации: байесовский классификатор, метод k-ближайших соседей, логистическая регрессия, машины опорных векторов, деревья решений.
 - Методы решения задач кластеризации: метод k-средних, DBSCAN, иерархическая кластеризация.
 - Методы решения задач регрессии: линейная регрессия, полиномиальная регрессия, гребневая регрессия.

Раздел 2. Глубокое обучение (24 ч.)

1. Глубокие нейронные сети.
 - Искусственный интеллект и нейронные сети.
 - Принципы работы нейронных сетей.
 - Классы нейронных сетей.
 - Нейронные сети для решений задач интеллектуальной обработки данных: модели нейронных сетей для классификации, модели нейронных сетей для кластеризации, модели нейронных сетей для регрессии.
 - Обучение нейронных сетей.
 - Глубокие нейронные сети.
2. Модели глубокого обучения для решения прикладных задач.
 - Основы цифровой обработки изображений.
 - Цифровая обработка изображений с использованием нейронных сетей.
 - Сверточные нейронные сети.
 - Архитектуры глубоких сверточных нейронных сетей: AlexNet, VGGNet, DeconvNet, SegNet.
 - Модели глубокого обучения для обнаружения объектов на изображениях: R-CNN, SPP-Net, Fast R-CNN, Faster R-CNN, SSD, YOLO.
 - Основы обработки естественного языка с использованием нейронных сетей.

- Рекуррентные нейронные сети.
- Трансформеры.
- GPT.

Раздел 3. Обучение с подкреплением (24 ч.)

1. Введение в обучение с подкреплением.
 - Основные понятия обучения с подкреплением (понятия агента, среды, состояния, действий и награды).
 - Функция ценности состояния (Value function) и функция качества действия (Qfunction).
 - Оптимизация стратегии с помощью максимизации функций ценности и качества.
 - Q-обучение.
2. Глубокое обучение с подкреплением.
 - Марковский процесс принятия решений.
 - Policy-based методы в обучении с подкреплением, глубокое обучение с подкреплением.
 - Deep Q-Networks, Actor-critic подходы.

Перечень практических и лабораторных занятий

Раздел	Номер темы	Наименование практического или лабораторного занятия	Кол-во часов
1	2	3	4
1	1	Первичная обработка, кластеризация и визуализация данных	4
	2	Классификация данных	4
2	1	Обнаружение объектов на изображениях	4
	2	Обработка запросов на естественном языке	4
3	1	Управление модулями мобильных платформ	4
	2	Формирование рекомендаций	4

5. Материально-технические условия реализации программы

Наименование специализированных аудиторий кабинетов, лабораторий (адрес)	Вид занятий	Наименование оборудования/программного обеспечения
1	2	3
Лаборатории 332, 341, 343, 345, 347, 351, 353 учебного здания № 7 (кафедра автоматизированных систем обработки информации и управления), ул. Б.Красная, 55	Лекции	Проектор, персональный компьютер для обеспечения работы проектора, проекционный экран, доска
Лаборатории 332, 341, 343, 345, 347, 351, 353 учебного здания № 7 (кафедра автоматизированных систем обработки информации и управления), ул. Б.Красная, 55	Практические и лабораторные занятия	Персональные компьютеры под управлением операционной системы Microsoft Windows или Linux, язык программирования Python, среда программирования для языка Python, программные библиотеки OpenCV, PyTorch, TensorFlow, Keras, OpenAI Gym.

6. Учебно-методическое и информационное обеспечение программе

6.1 Основная и дополнительная учебная литература

1. Агарвал, Ч. Нейронные сети и глубокое обучение: учебный курс / Ч. Агарвал. – СПб.: ООО «Диалектика», 2020. – 752 с.
2. Бишоп, К.М. Распознавание образов и машинное обучение / К.М. Бишоп. – СПб.: ООО «Диалектика», 2020. – 960 с.
3. Бурков, А. Машинное обучение без лишних слов / А. Бурков. – СПб.: Питер, 2020. – 192 с.
4. Грессер, Л. Глубокое обучение с подкреплением: теория и практика на языке Python / Л. Грессер, Л.В. Кенг. – СПб.: Питер, 2022. – 416 с.
5. Груздев, А.В. Прогнозное моделирование в IBM SPSS Statistics, R и Python. Метод деревьев решений и случайный лес. Цветное издание / А.В. Груздев. – М.: ДМК Пресс, 2017. – 642 с.
6. Гудфеллоу, Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвиль. – М.: ДМК Пресс, 2017. – 652 с.
7. Джоши, П. Искусственный интеллект с примерами на Python / П. Джоши. – СПб.: ООО «Диалектика», 2019. – 448 с.
8. Компьютерное зрение. Современные методы и перспективы развития / Под ред. Р. Дэвиса, М. Терка. – М.: ДМК Пресс, 2022. – 690 с.
9. Кэлер, А. Изучаем OpenCV 3 / А. Кэлер, Г. Брэдки. – М.: ДМК Пресс, 2017. – 826 с.
10. Лю, Ю. Обучение с подкреплением на PyTorch: сборник рецептов / Ю. Лю. – М.: ДМК Пресс, 2020. – 282 с.
11. Мэрфи, К.П. Вероятностное машинное обучение: введение / К.П. Мэрфи. – М.: ДМК Пресс, 2022. – 940 с.
12. Николенко, С. Глубокое обучение / С. Николенко, А. Кадури, Е. Архангельская. – СПб.: Питер, 2018. – 480 с.
13. Орельен, Ж. Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем, 2-е изд. / Ж. Орельен. – СПб.: ООО «Альфа-книга», 2020. – 1040 с.
14. Осовский, С. Нейронные сети для обработки информации / С. Осовский. – М.: Горячая линия-Телеком, 2017. – 448 с.
15. Протодяконов, А.В. Алгоритмы Data Science и их практическая реализация на Python: учебное пособие / А.В. Протодяконов, П.А. Пылов, В.Е. Садовников. – Москва; Вологда: Инфра-Инженерия, 2022. – 392 с.
16. Рашка, С. Python и машинное обучение: машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow 2, 3-е изд. / С. Рашка, В. Мирджалили. – СПб.: ООО «Диалектика», 2020. – 848 с.
17. Ростовцев, В. С. Искусственные нейронные сети : учебник для вузов / В. С. Ростовцев. – 2-е изд., стер. – СПб.: Лань, 2021. – 216 с.
18. Саваш, Й. Осваиваем архитектуру Transformer. Разработка современных моделей с помощью передовых методов обработки естественного языка / Й. Саваш, А. Мейсам. – М.: ДМК Пресс, 2022. – 320 с.
19. Саттон. Р.С. Обучение с подкреплением: Введение. 2-е изд. / Р.С. Саттон, Э.Дж. Барто. – М.: ДМК Пресс, 2020. – 552 с.
20. Уиндер, Ф. Обучение с подкреплением для реальных задач / Ф. Уиндер. – СПб.: БХВ-Петербург, 2023. – 400 с.

6.2 Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

1. Афонин В., Макушкин В. Интеллектуальные робототехнические системы [Электронный ресурс]. – URL: <https://www.intuit.ru/studies/courses/46/46/info> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.

2. Введение в разработку мультимедийных приложений с использованием библиотек OpenCV и IPP [Электронный ресурс]. – URL: <http://www.intuit.ru/studies/courses/10621/1105/info> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.
3. Воронцов К.В. Машинное обучение [Электронный ресурс]. – URL: <http://www.machinelearning.ru> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.
4. Местецкий Л.М. Математические методы распознавания образов [Электронный ресурс]. – URL: <https://www.intuit.ru/studies/courses/2265/243/info> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.
5. Официальный сайт библиотеки OpenCV [Электронный ресурс]. – URL: <https://opencv.org/> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.
6. Сотник С. Проектирование систем искусственного интеллекта [Электронный ресурс]. – URL: <https://www.intuit.ru/studies/courses/1122/167/info> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.
7. Технологии анализа данных [Электронный ресурс]. – URL: <https://basegroup.ru/> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.
8. Open Machine Learning Course [Электронный ресурс]. – URL: <https://mlcourse.ai> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.
9. Введение в Data Science и машинное обучение [Электронный ресурс]. – URL: <https://stepik.org/course/4852/promo> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.
10. Level up with the largest AI & ML community [Электронный ресурс]. – URL: <https://www.kaggle.com/> (дата обращения: 17.05.24). – Текст электронный. Режим доступа: свободный.

6.3 Перечень информационных технологий, используемых при осуществлении образовательных процессов по программе

1. Программно-аппаратные средства подготовки и демонстрации презентаций.
2. Программно-аппаратное обеспечение лабораторных работ.
3. Образовательные ресурсы КНИТУ-КАИ.
4. Образовательные ресурсы сети Интернет.

7. Оценка качества освоения программы

Контроль освоения программы осуществляется в виде итоговой аттестации по результатам тестирования во время зачета. Для получения оценки «Зачтено» должны быть правильно решены не менее 50% тестовых заданий. В противном случае будет получена оценка «Не зачтено».

Примеры тестовых заданий:

1. Укажите непараметрические методы обучения.
 - + Метод k ближайших соседей;
 - + Метод на основе ядерных моделей плотности вероятности;
 - Метод SVM;
 - Метод байесовской классификации.
2. Как называется отношение истинно положительных прогнозов к общему количеству положительных прогнозов?
 - + Точность;
 - Полнота;
 - Правильность;
 - Специфика.

8. Кадровые условия реализации программы

В реализации программы принимает участие профессорско-преподавательский состав, научные сотрудники и инженерно-технические работники кафедры автоматизированных систем обработки информации и управления КНИТУ-КАИ.

9. Разработчики и составители программы

1. Андреянов Н.В., старший преподаватель кафедры автоматизированных систем обработки информации и управления;
2. Евдокимова Т.С., ассистент кафедры автоматизированных систем обработки информации и управления;
3. Павлов А.Д., старший преподаватель кафедры автоматизированных систем обработки информации и управления;
4. Шлеймович М.П., заведующий кафедрой автоматизированных систем обработки информации и управления.