

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Беспалов Владимир Александрович

Должность: Ректор МИЭТ

Дата подписания: 01.09.2025 15:20:13

Уникальный программный ключ:

ef5a4fe6ed0ffdf3f1a49d6ad1b49464dc1bf7354175603ce9b0c481288ca01

Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

«Национальный исследовательский университет

«Московский институт электронной техники»



Проректор по УР

И.Г. Игнатова

« 1 »

12

20 20

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

«Методы обработки и анализа больших данных»

Направление подготовки 11.03.04- «Электроника и нанoeлектроника»

Направленность (профиль)- «Автоматизация проектирования изделий нанoeлектроники»

2020 г.

1. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ

Дисциплина участвует в формировании следующих компетенций образовательных программ:

Компетенция ПК-2 «Способен аргументировано выбирать и реализовывать на практике эффективную методику экспериментального исследования параметров и характеристик приборов, схем, устройств и установок электроники и наноэлектроники различного функционального назначения» **сформулирована на основе профессионального стандарта 06.001 «Программист»**

Обобщенная трудовая функция: D – «Разработка требований и проектирование программного обеспечения»

Трудовая функция: D/03.6 – «Проектирование программного обеспечения»

Подкомпетенции, формируемые в дисциплине	Задачи профессиональной деятельности	Индикаторы достижения компетенций/подкомпетенций
ПК-2.МОиАБД Способен аргументировано выбирать и реализовывать на практике эффективную методику анализа данных экспериментальных исследований	Применять методы обработки больших данных, в том числе нейронные сети для анализа изображений, звука, текстов и градиентный бустинг для анализа табличных данных.	Знания методов обработки и анализа больших данных экспериментальных исследований. Умения обрабатывать и анализировать большие данные экспериментальных исследований. Опыт обработки и анализа больших данных в виде изображений, полученных с камер и визуальных датчиков устройств электроники и наноэлектроники различного функционального назначения.

Компетенция ПК-3 «Способен выполнять расчет и проектирование электронных приборов, схем и устройств различного функционального назначения в соответствии с техническим заданием с использованием средств автоматизации проектирования» **сформулирована на основе профессионального стандарта 06.001 «Программист».**

Обобщенная трудовая функция: D – «Разработка требований и проектирование программного обеспечения»

Трудовая функция: D/01.6 – «Анализ требований к программному обеспечению»

Подкомпетенции, формируемые в дисциплине	Задачи профессиональной деятельности	Индикаторы достижения компетенций/подкомпетенций
ПК-3.МОиАБД Способен выполнять	Применять методы обработки больших	Знания принципов конструирования электронных

расчет и проектирование электронных приборов, схем и устройств различного функционального назначения большой размерности	данных, в том числе нейронные сети для решения типовых задач связанных с обработкой больших данных	приборов большой размерности. Умения проводить оценочные расчеты характеристик электронных приборов большой размерности. Опыт использования методов обработки данных большой размерности.
--	--	---

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина входит в часть, формируемую участниками образовательных отношений Блока 1 «Дисциплины (модули)» образовательной программы, изучается на Зкурсе, в 5 семестре (очная форма обучения).

Изучение дисциплины базируется на следующих ранее изучаемых дисциплинах: «Математический анализ», «Дискретная математика», «Теория вероятностей», «Основы объектно-ориентированного программирования».

3. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Курс	Семестр	Общая трудоёмкость (ЗЕ)	Общая трудоёмкость (часы)	Контактная работа			Самостоятельная работа (часы)	Промежуточная аттестация
				Лекции (часы)	Лабораторные работы (часы)	Практические занятия (часы)		
3	5	2	72	16	16	-	40	ЗаО

4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

№ и наименование модуля	Контактная работа			Самостоятельная работа	Формы текущего контроля
	Лекции	Практические занятия	Лабораторные занятия		
1. Введение в методы обработки и анализа больших данных и язык программирования Python	4	-	-	8	Написание самостоятельной работы к лекциям модуля 1
2. Методы обработки табличных данных и	2	-	4	8	Написание самостоятельной работы к лекциям модуля 2

градиентный бустинг					Защита лабораторных работ
3. Методы работы с нейронными сетями для обработки больших данных разного рода	8	-	12	14	Написание самостоятельной работы к лекциям модуля 3
					Защита лабораторных работ
4. Аппаратные реализации нейронных сетей	2	-	-	6	Написание самостоятельной работы к лекциям модуля 4
1,2,3,4	-	-	-	4	Сдача практико-ориентированного задания

4.1. Лекционные занятия

№ модуля дисциплины	№ лекции	Объем занятий (часы)	Краткое содержание
1	1	2	Язык программирования Python и полезные модули. Важность Python для машинного обучения и обработки больших данных. Встроенные типы данных. Основные алгоритмические конструкции. Арифметические операции. Работа со списками, кортежами, словарями, множествами. Функции, исключения, ООП в Python. Итераторы и генераторы. Ввод и вывод данных. Модули и пакеты. Модули numpy, pandas, opencv.
	2	2	Типы больших данных, подготовка данных, метрики качества и методы валидации. Типы и источники больших данных. Известные наборы данных для изображений. Определение машинного обучения и его типы. Модели и метрики качества. Дилемма смещения и дисперсии. Перекрестная проверка (cross-validation). Матрица ошибок (confusionmatrix). Площадь под ROC-кривой. Экстракция признаков. Регуляризация. Ансамбли моделей.
2	3	2	Градиентный бустинг и другие методы для обработки табличных данных. Деревья решений. Коэффициент Джинни. Алгоритм «Случайный лес». Алгоритм «AdaBoost». Алгоритм “Gradient Boost”. Модуль XGBoost.
3	4	2	Введение в нейронные сети. Многослойный персептрон. Методы построения нейронных сетей. Полносвязные слои. Функции активации sigmoid и RELU. Набор данных MNIST. Градиент функции цены. Алгоритм обратного распространения ошибки (Back Propagation). Стохастический градиентный спуск. Вычислительный граф и chain rule.
	5	2	Свёрточные нейронные сети для работы с изображениями.

			Изображение с точки зрения компьютера. Классы задач при работе с изображениями. Известные наборы данных (datasets) для классификации изображений. Различные архитектуры сверточных нейронных сетей от самых первых до настоящего времени. Метод «Transfer Learning». Разбор работы простой свёрточной нейронной сети в деталях. Двухмерная свертка (Conv2D). Слои MaxPooling, RELU, Dropout, BatchNormalization. Активация softmax. Тренировка нейронной сети. Пример тренировки с помощью модуля PyTorch. Аугментации изображений.
6	2		Работа с текстовыми данными большого объема. Определение что такое текст. Традиционный подход к кодированию слов для обработки. Алгоритм Word2Vec и подготовка данных для него. CBOW, Skipgram и NegativeSampling. Алгоритм GloVe. Рекуррентные нейронные сети. Нейронные сети типа “Transformers”. Нейронная сеть BERT. Известные текстовые наборы данных. Детальный пример Fine-Tuning сети BERT для датасета CoLA. Аугментации текста.
7	2		Работа со звуковыми данными. Задачи ObjectDetection и Segmentation для изображений. Представление звука в цифровом виде. Понятие «частота сигнала» или frequency. Дискретизация сигнала. Алиасинг, примеры и как избежать. Быстрое преобразование Фурье. Спектрограммы и мел-спектрограммы. Классификация звука на одномерных свёртках. Узкие места при тренировке нейронных сетей. Известные датасеты со звуком. Аугментации звука. Задача обнаружения объектов на изображении (Object Detection). Типы и устройство моделей для Object Detection. Задача семантической сегментации изображений. Известные архитектуры сетей для сегментации. Проблемы в задачах сегментации.
4	8	2	Реализация нейронных сетей на аппаратном уровне. Режимы работы нейронной сети. Устройства для расчёта на нейронных сетях. Отличия работы CPU, GPU и TPU. Методы ускорения тренировки нейронных сетей. Тренировка в режиме Mixed Precision. Методы оптимизации режима Inference. Методы квантизации нейронных сетей. Кластеризация весов. Алгоритмические оптимизации. Потребление энергии для разных типов операций.

4.2. Практические занятия

Не предусмотрены

4.3. Лабораторные занятия

№ модуля дисциплины	№ лабораторной работы	Объем занятий (часы)	Краткое содержание
2	1	4	Применение методов градиентного бустинга (модуль XGBoost) для предсказания оттока абонентов. Изучение системы Kaggle In-class. Работа с языком Python в разделе Code онлайн.
3	2	4	Использование свёрточных нейронных сетей для поиска людей на изображениях. Изучение работы с GPU для ускорения вычислений. Работа с модулем PyTorch в Python.
	3	4	Автоматическая классификация рецензий к фильмам на позитивные и негативные на базе нейронных сетей BERT.
	4	4	Автоматическое определение заданного набора слов из отдельных аудио-файлов на базе спектрограмм и свёрточных нейронных сетей.

4.4. Самостоятельная работа студентов

№ модуля дисциплины	Объем занятий (часы)	Вид СРС
1	7	Подготовка к лабораторным работам: изучение базовых методик программирования на языке Python.
2	7	Изучение системы Kaggle In-class и раздела Code для отправки решений лабораторных работ на проверку.
3	7	Изучение необходимой литературы и методических указаний для решения лабораторных работ.
1, 2, 3, 4	11	Выполнение и контроль самостоятельной работы к лекциям модулей 1, 2, 3 и 4.
	4	Выполнение практико-ориентированного задания
	4	Подготовка к зачёту

4.5. Примерная тематика курсовых работ (проектов)

Не предусмотрены

5. ПЕРЕЧЕНЬ УЧЕБНО-МЕТОДИЧЕСКОГО ОБЕСПЕЧЕНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Учебно-методическое обеспечение для самостоятельной работы студентов в составе УМК дисциплины (ОРИОКС, <http://orioks.miet.ru/>):

Методические указания студентам по изучению дисциплины «Теория алгоритмов»: https://orioks.miet.ru/prepare/ir-science?id_science=2434084

Модуль 1 «Введение в методы обработки и анализа больших данных и язык программирования Python»

Методические материалы, перечень литературы, информационных источников для выполнения заданий для самостоятельной работы по тематике модуля 1, требования к выполнению самостоятельной работы и методика её оценивания, а также отражение результатов выполнения самостоятельной работы в НБС содержатся в разделе «Самостоятельная работа студентов» УМК дисциплины.

Модуль 2 «Методы обработки табличных данных и градиентный бустинг»

Методические материалы, перечень литературы, информационных источников для выполнения заданий для самостоятельной работы по тематике модуля 2, требования к выполнению самостоятельной работы и методика её оценивания, а также отражение результатов выполнения самостоятельной работы в НБС содержатся в разделе «Самостоятельная работа студентов» УМК дисциплины.

Модуль 3 «Методы работы с нейронами сетями для обработки больших данных разного рода»

Методические материалы, перечень литературы, информационных источников для выполнения заданий для самостоятельной работы по тематике модуля 3, требования к выполнению самостоятельной работы и методика её оценивания, а также отражение результатов выполнения самостоятельной работы в НБС содержатся в разделе «Самостоятельная работа студентов» УМК дисциплины.

Модуль 4 «Аппаратные реализации нейронных сетей»

Методические материалы, перечень литературы, информационных источников для выполнения заданий для самостоятельной работы по тематике модуля 4, требования к выполнению самостоятельной работы и методика её оценивания, а также отражение результатов выполнения самостоятельной работы в НБС содержатся в разделе «Самостоятельная работа студентов» УМК дисциплины.

6. ПЕРЕЧЕНЬ УЧЕБНОЙ ЛИТЕРАТУРЫ

Литература

1. Рутковская, Д. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер.с польск.И.Д.Рудинского : учебное пособие / Д. Рутковская, М. Пилиньский, Л. Рутковский. — 2-е изд. — Москва : Горячая линия-Телеком, 2013. — 384 с. — ISBN 978-5-9912-0320-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/11843> (дата обращения: 12.03.2021). — Режим доступа: для авториз. пользователей

2. Галушкин, А. И. Нейронные сети: основы теории / А. И. Галушкин. — Москва : Горячая линия-Телеком, 2017. — 496 с. — ISBN 978-5-9912-0082-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/111043> (дата обращения: 12.03.2021). — Режим доступа: для авториз. пользователей.
3. Обработка изображений с помощью OpenCV / Б. Г. Глория, Д. С. Оскар, Л. Э. Хосе, С. Г. Исмаэль. — Москва : ДМК Пресс, 2016. — 210 с. — ISBN 978-5-97060-387-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/90116> (дата обращения: 12.03.2021). — Режим доступа: для авториз. пользователей

Периодические издания

1. Известия вузов. Электроника : Научно-технический журнал / М-во образования и науки РФ; МИЭТ; Гл. ред. Ю.А. Чаплыгин. - М. : МИЭТ, 1996 - . - ISSN 1561 – 5405
2. IEEETRANSACTIONSONCOMPUTERAIDEDDESIGNOFINTERGRATEDCIRCUITS&SYSTEMS. - USA : IEEE, [б.г.]. – URL:<http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=43>. – Режим доступа: из локальной сети НИУ МИЭТ в рамках проекта "Национальная подписка"

7. ПЕРЕЧЕНЬ ПРОФЕССИОНАЛЬНЫХ БАЗ ДАННЫХ, ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ

1. Лань: электронно-библиотечная система. - Санкт-Петербург, 2011 - URL: <https://e.lanbook.com/> (дата обращения: 30.10.2020). - Режим доступа: для авторизованных пользователей МИЭТ.
2. SCOPUS: Библиографическая и реферативная база данных научной периодики: сайт. – URL: www.scopus.com/ (дата обращения: 30.09.2019). - Режим доступа: для авториз. пользователей МИЭТ
3. eLIBRARY.RU: научная электронная библиотека: сайт. – Москва, 2000. – URL: <https://elibrary.ru> (дата обращения: 30.10.2020). – Режим доступа: для зарегистрированных пользователей.
4. IEEE/IET Electronic Library (IEL) = IEEE Xplore : электронная библиотека. - USA ; UK, 1998 - . - URL: <https://ieeexplore.ieee.org/Xplore/home.jsp> (дата обращения: 28.10.2020). - Режим доступа: из локальной сети НИУ МИЭТ в рамках проекта "Национальная подписка"

8. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В ходе реализации обучения используется смешанное обучение, применяется «расширенная виртуальная модель», которая предполагает обязательное присутствие студентов на очных учебных занятиях (лекциях и лабораторных работах) с последующим самостоятельным выполнением индивидуального задания (индивидуальные практические задания к лабораторным работам).

Обучение может реализовываться с применением дистанционных образовательных технологий.

Освоение образовательной программы обеспечивается ресурсами электронной информационно-образовательной среды ОРИОКС.

Для взаимодействия студентов с преподавателем могут использоваться сервисы обратной связи, такие как электронная почта, система видеоконференций Zoom.

В процессе обучения при проведении занятий и для самостоятельной работы используются внутренние электронные ресурсы в формах материалов в системе ОРИОКС:
URL: https://orioks.miet.ru/prepare/ir-science?id_science=2434084

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень программного обеспечения
Учебная аудитория	Мультимедийное оборудование	Microsoft (Azure), Microsoft Office
Вычислительный класс каф. ПКИМС, ауд. 4131.	ПЭВМ Intel LGA1156 Core i5-661 с мониторами Iiyama и ViewSonic.	Microsoft (Azure) LibreOffice Python
Помещение для самостоятельной работы	Компьютерная техника с возможностью подключения к сети «Интернет» и обеспечением доступа в ОРИОКС	Microsoft (Azure), браузер Google Chrome

10. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕРКИ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ

1. ФОС по компетенции/подкомпетенции **ПК-2.МОиАБД** «Способен аргументировано выбирать и реализовывать на практике эффективную методику анализа данных экспериментальных исследований».
2. ФОС по компетенции/подкомпетенции **ПК-3.МОиАБД** «Способен выполнять расчет и проектирование электронных приборов, схем и устройств различного функционального назначения большой размерности»

Фонды оценочных средств представлены отдельными документами и размещены в составе УМК дисциплины электронной информационной образовательной среды ОРИОКС// URL: <http://orioks.miet.ru/>.

11. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

11.1. Особенности организации процесса обучения

Студенты, изучающие дисциплину, обязаны:

- посетить лекции по предмету;
- выполнить лабораторные работы (подтверждается сдачей каждой лабораторной работы);
- принять участие в дискуссиях во время лекций и лабораторных работ;
- выполнить задание проектного типа на опыт деятельности.

В процессе изучения курса предполагается самостоятельная работа студента при подготовке к лекционным занятиям, лабораторным работам, использование литературы, интернет-ресурсов.

С целью качественной организации самостоятельной работы студентов проводятся разъяснения материала. Вводное разъяснение проводится лектором дисциплины в начале первой лекции и включает: информацию о структуре и графике контрольных мероприятий, содержании и порядке проведения контрольных мероприятий, правилах оценивания согласно НБС МИЭТ, учебной литературе и дополнительных информационных источниках, основных требованиях по оценке качества освоения дисциплины, самостоятельной работе студентов, организации и назначении консультаций. На каждой из последующих лекций проводится разъяснение по выполнению заданий для СРС.

Для студентов проводятся консультации согласно расписанию. Студентам рекомендуется активно пользоваться консультациями преподавателя: это единственная возможность обучаться индивидуально и выяснить все возникшие вопросы. Кроме этого, на консультациях можно защитить лабораторную работу, если не успели на занятии.

По завершению изучения дисциплины предусмотрено публичное представление результатов заданий для СРС на опыт деятельности и заданий проектного типа и промежуточная аттестация в виде зачета с оценкой.

11.2. Система контроля и оценивания

Для оценки успеваемости студентов по дисциплине используется накопительная балльная система (НБС).

Баллами оцениваются: выполнение каждого контрольного мероприятия в семестре (в сумме максимум 80 баллов), и сдача зачёта с оценкой (максимум 20 баллов).

По сумме баллов выставляется итоговая оценка по предмету. Структура и график контрольных мероприятий приведены ниже в таблице (см. также журнал успеваемости на ОРИОКС, <http://orioks.miet.ru/>).

РАЗРАБОТЧИК:

Профессор кафедры ПКИМС, д.т.н.



/Р.А. Соловьёв/

Рабочая программа дисциплины «Методы обработки и анализа больших данных» по направлению подготовки 11.03.04 «Электроника и нанoeлектроника», направленности (профилю) «Автоматизация проектирования изделий нанoeлектроники» разработана на кафедре ПКИМС и утверждена на заседании кафедры 27 ноября 2020 года, протокол № 8

Заведующий кафедрой ПКИМС

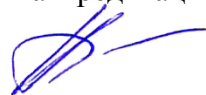


/С.В. Гаврилов/

ЛИСТ СОГЛАСОВАНИЯ

Рабочая программа согласована с центром подготовки к аккредитации и независимой оценки качества

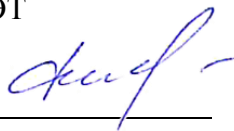
Начальник АНОК



/И.М. Никулина/

Рабочая программа согласована с библиотекой МИЭТ

Директор библиотеки



/ Т.П. Филиппова/